



Friend or Therapist? A Systematic Literature Review on Chatbots and Mental Health in Young People

Downloaded from: <https://research.chalmers.se>, 2026-07-17 18:15 UTC

Citation for the original published paper (version of record):

Hashmati, N., Skoog, T., Obaid, M. et al (2026). Friend or Therapist? A Systematic Literature Review on Chatbots and Mental Health in Young People. Proceedings of the 1st International Conference on Human-Computer Interaction in the Alps. <http://dx.doi.org/10.1145/3780045.3780052>

N.B. When citing this work, cite the original published paper.

Friend or Therapist? A Systematic Literature Review on Chatbots and Mental Health in Young People

Negin Hashmati

Chalmers University of Technology and University of
Gothenburg
Gothenburg, Sweden
hashmati@chalmers.se

Mohammad Obaid

Chalmers University of Technology and University of
Gothenburg
Gothenburg, Sweden
mobaid@chalmers.se

Therése Skoog

University of Gothenburg
Gothenburg, Sweden
therese.skoog@psy.gu.se

Thommy Eriksson

Chalmers University of Technology and University of
Gothenburg
Gothenburg, Sweden
thommy@chalmers.se

Abstract

In this systematic literature review, we analyze 58 papers from the Human-Computer Interaction and Psychology literature over the last decade to understand how researchers have studied chatbots and their impact on mental health in young people (aged 16-25), covering mental health conditions such as anxiety and depression. Many studies target general mental health conditions and often utilize self-help methods. Concerning types of relationships, we found that friendship and therapeutic relationships were sometimes brought up, but in many cases, the relationship aspect was not considered. Our findings highlight the need for continued interdisciplinary research on how young people use chatbots, including how avatar design may reinforce or reduce biases and stereotypes. We also suggest recommendations for chatbot development and research that can help cover the research gaps we have observed.

CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI)**.

Keywords

Literature Review, Chatbots, Mental Health, Well-being, Young People, Human-Computer Interaction

ACM Reference Format:

Negin Hashmati, Therése Skoog, Mohammad Obaid, and Thommy Eriksson. 2026. Friend or Therapist? A Systematic Literature Review on Chatbots and Mental Health in Young People. In *1st International Conference on Human-Computer Interaction in the Alps (AlpCHI 2026)*, March 01–05, 2026, Ascona, Switzerland. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3780045.3780052>



This work is licensed under a Creative Commons Attribution 4.0 International License. *AlpCHI 2026*, Ascona, Switzerland
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1948-6/26/03
<https://doi.org/10.1145/3780045.3780052>

1 Introduction

Young people are sometimes referred to as "digital natives" [54], meaning that they have had easy access to, and have relied on, digital platforms through most of their lives. In parallel, the mental health among young people is deteriorating, with conditions such as depression and anxiety increasing [86]. In fact, levels of mental health problems are disproportionately high among young people (i.e., adolescents and young adults) compared to other age groups. As a response to this, there is a large focus on chatbots for therapy [1, 41, 47, 75, 76], something that is a result of the growing demand for therapy in general across the globe, together with there not being enough therapists available, and people living too far away from the closest therapist [5], such as in rural or village areas. Juxtaposing these aspects led us to investigate if chatbots are used for mental health support, how they are used, and what the future challenges are regarding these technologies. Our aim with this review is to investigate how researchers have utilized chatbots for young people (aged 16-25) concerning their mental health. We are also interested in chatbots that serve as companions to users, who consider them friends, partners, or someone with whom they can discuss anything, and investigating how that affects their mental health. Our rationale for conducting a literature review on this topic at this time is that the most recent comprehensive review we identified is three years old [47]. Given the rapid developments in this field, a new examination of the current state of research is both timely and warranted.

In this paper, we use the concept **chatbot**, which overlaps with concepts such as: *virtual companion*, *artificial companion*, *intelligent agent*, *virtual agent*, *intelligent virtual agent*, *AI companion*, *AI agent*, *conversational agent*. By chatbot, we mean a computer program that engages users in conversations, often giving the illusion of humanness [58]. We are especially interested in whether the chatbots have *memory* (remembering previous conversations) and *initiative* (initiating conversations with the user), as they have been suggested as features that can facilitate social interaction and social bonding over time [9, 11].

Our contributions are 1) a comprehensive review of the past decade of research on the use of chatbots in the context of mental health in young people, 2) highlighting existing research gaps, and

3) recommendations for chatbot research and development. We used the following research questions to guide this study:

- RQ1: What aspects of user mental health have been addressed in the research on chatbots, and what therapeutic tools are used for them?
- RQ2: What study designs have been used, and what are the most common research methods used in the studies?
- RQ3: What are the most common relationship types with chatbots in the studies?
- RQ4: What characteristics and interaction modalities of chatbots can be found in the studies?

2 Related Work

Looking at previous literature reviews [1, 2, 7, 41, 47, 50, 70, 75, 76], there are a number of conclusions that appear consistently. A rationale for using chatbots in mental healthcare is the increasing issues with mental health conditions such as anxiety and depression, and the lack of therapeutic intervention resources [1, 2, 75, 76]. There are estimates that 20-40% of the population worldwide will need mental healthcare at some point during their lives [2, 75].

At the same time, there are potential problems with providing counseling through chatbots. Several ethical concerns, such as risks of privacy breaches, counseling practices that are unsupported by research, biases, and unclear limits between therapeutic alliance relationships and more friendship-oriented relations between the care seeker and the chatbot, have been raised [1, 2, 7, 41, 47, 70, 75, 76]. Previous literature reviews call for further research specifically on these and other ethical issues [1, 2, 41, 47, 70, 75, 76]. Studies report low levels of engagement from care seekers, and it seems common that interactions with chatbots are quickly abandoned [1, 7, 76]. A common suggestion to mitigate these problems is to use therapist-guided chatbot interventions, a form of collaboration between the chatbot and a human counselor [1, 2, 75, 76]. However, this would require integration between the chatbot and existing healthcare infrastructure (such as patient medical records), and such integration is not in place [76].

While the number of commercial chatbots marketed as therapists, friends, and romantic partners has increased drastically in the last five years, there is little research on the usage or efficacy of these [7, 70]. At the same time, the most popular of these chatbots – such as Replika¹ and Character.AI² – have millions of active users [7, 50], and there are some indications that these chatbots actually can mitigate, for example, suicidal ideation [50].

There are several theoretical frameworks in the literature that can be used to describe and understand user-chatbot relationships. For instance, according to the Computers Are Social Actors (CASA) theory [57], users often treat and think of chatbots as if they were social, intelligent beings. Similarly, the social presence theory [13] can be taken to suggest that users can feel like chatbots are real people and partners. As a result, and according to attachment theory, it can be expected that users develop what can be thought of as attachment "bonds" and attachment styles vis-a-vis chatbots [30]. Consequently, the relationships may feel very real, important, and human-like to users.

¹<https://replika.com/>

²<https://character.ai/about>

In sum, further research on chatbots and their utilization in mitigating mental health conditions is justified. The area is complex and requires interdisciplinary approaches [1, 70, 75], which already seems to be the case in many studies [41]. These calls for further research stem from literature reviews published several years ago [47]. Given the rapid evolution of AI chatbot technologies in recent years, the landscape has shifted considerably, warranting updated examination. The focus we have chosen in this paper – chatbots for mental health support in young people – has not been thoroughly examined in any previous systematic literature review to our knowledge. This lack of knowledge needs to be addressed, given that mid-adolescence to young adulthood (age 16–25 years) is a developmental period marked by vulnerability to mental health difficulties [8], low rates of formal help-seeking [62, 68], and high engagement with digital technologies [82]. Chatbots have the potential to be scalable, accessible, and stigma-reducing tools that could bridge service gaps for this population, yet it remains unclear how they are used and experienced by young people. A systematic review is essential to identify and map the existing evidence base, identify opportunities and limitations, and inform the development of adequate and effective policy and practice in the field.

3 Method

3.1 Procedure

We followed a systematic literature review process [77], conducted following the PRISMA format [56] (see the supplementary material for a figure of the process).

3.1.1 Search Query Keywords. The databases we used were the ACM Digital Library and Web of Science, prominent digital libraries used in earlier systematic literature reviews [41, 70, 75]. Including both databases ensured both HCI and psychology perspectives. The search was done on both databases on January 20th, 2025, and included literature from 2015 to 2025.

The query keywords used were: *chatbot*, *virtual companion*, *virtual agent*, *intelligent agent*, *artificial companion*, *conversational agent*, *intelligent virtual agent*, *AI agent*, *AI companion*, *agent*, *mental health*, *teenager**, *adolescent**, *young adult**, *child**, *student**. We included the terms *teenager*, *adolescent*, and *young adult* as they represent the age range (16-25 years old) we were interested in. We include *children* and *students* since these terms might include participants within our age range.

This search gave 790 results from ACM Digital Library and 379 results from Web of Science. We identified three papers as duplicates and removed them, giving us a total of 1166 papers. The first author read the titles and abstracts to determine if they were within the scope of our study. The papers were excluded if their abstract did not mention mental health, young people, and chatbots or any variations of these terms. This left 241 papers that two authors divided between themselves and did a full-text reading on based on the eligibility criteria described below. Papers that were unclear whether they should be included or not were discussed between the same two authors. After this full-text screening, 65 papers remained to code. During the coding process, an additional seven papers were removed when they at closer scrutiny were found not to fulfill the

eligibility criteria. We coded the final set of 58 papers which are included in our results.

3.1.2 Eligibility Criteria. We used the following eligibility criteria. If a criterion was not fulfilled, then that paper was excluded.

Type of dissemination: The papers had to be either full conference papers or journal articles.

Age: This criterion could be fulfilled in three ways: 1) the reported mean, median or average ages of the participants were within 16-25 years, 2) the reported age span was within 16-25 years for at least 70% of the participants (which meant that studies could have a limited number of older participants), or 3) the participants were categorized as teenager, adolescent, young adult, or student, (which means that if a paper did not report any age at all but only these categories, we assumed that a majority of the participants would be within our age span of 16-25). We omitted studies that used *children* to describe their participants, as they could be very young users and thus outside the focus of this study. If a study did not involve participants, there had to be an explicit focus on users categorized as teenagers, adolescents, young adults, or students.

Mental health conditions: We included studies that addressed mental health either broadly or specifically referring to common and general mental health conditions or symptoms, such as depression, anxiety, social anxiety, and loneliness. We excluded studies that covered chatbots used for specific disorders such as Attention-Deficit/Hyperactivity Disorder (ADHD), Post-Traumatic Stress Disorder (PTSD), autism, or substance use disorder (SUD), to name a few examples. This is because we were interested in what research has been conducted on the most common mental health conditions.

Chatbot type: We excluded all studies covering physical social robots, and rule-based virtual assistants such as Amazon's *Alexa*.

3.2 Coding Scheme

The coding scheme was based on the research questions and data available in the included papers. For each paper, we extracted: type of research paper (conference paper or journal article), name of conference/journal, publication year, what mental health conditions or symptoms the participants in the studies had, how the application or the study intended to help the participants, what research tools and research methods were used, how the relationship between the user and the chatbot was described, whether it was an in-lab, field, or design study, age of the participants in the studies, whether the chatbot used in the study was commercial or developed by the researchers themselves (and if it was a commercial chatbot, which one they used), number of participants, what hardware platform (desktop, smartphone, etc) was used for the chatbot, what software platform (Facebook Messenger, WhatsApp, etc) was used for the chatbot, what interaction modality the chatbot had (text, audio, etc), what the design of the chatbot was like (chat interface, an avatar, videos, etc), if the chatbot had a memory, if the chatbot had initiative (i.e., if the chatbot reached out to the user first). The full coding scheme is provided in the supplementary material.

For the relationship code, we had the options *friendship*, *therapeutic* (e.g., therapist, counselor, or coach role), *romantic*, or *not mentioned*. For each of the first three codes, we added the qualifying sub-codes *mentioned* (the relationship was only briefly mentioned

by the researchers) and *in focus* (the relationship was discussed in detail and/or used to characterize the chatbot they used).

Two authors coded 20% of the papers in parallel and compared them to check rating agreement. Cohen's Kappa was calculated for the items that were not text-based (such as number of participants). The mean unweighted Kappa was 0.66, which indicates a substantial agreement. Discrepancies were discussed until all of the disagreements were resolved. The remaining papers were divided between the two authors and coded separately.

4 Results

Twenty-two of the papers were ACM conference papers (e.g. [45, 83, 89, 91]). Eight were published in the CHI Conference on Human Factors in Computing Systems [3, 6, 10, 35, 36, 42, 45, 90], and the remaining 50 were spread out over different conferences. Thirty-six papers were journal articles, over half of them retrieved in Web of Science (e.g. [17, 32, 51]). There was an increase in the amount of literature published from 2021 to 2024, with a peak in 2023 and 2024, when 14 papers were published each year.

We were interested in whether the chatbots were developed by the researchers or if they were commercially and publicly available. Fourteen studies used commercial chatbots [3, 14, 23, 24, 29, 34, 37, 39, 40, 46, 55, 60, 65, 80]. Of these, six used Woebot³ [3, 14, 23, 34, 60, 65], and the others all used different commercial ones. Self-developed chatbots were used in 29 studies (e.g. [42, 49, 53, 73]). Two studies had two chatbots present; both their self-developed one and a commercial one [27, 33].

Based on our eligibility criteria, we were only interested in studies where at least 70% the participants were between the ages of 16 and 25 years old. The papers all reported on the ages of the participants differently, whereby some only provided a mean/average age (such as [27]), while others reported only an age span (such as [3]), and some reported on both an age span and a mean/average age (such as [18]). Especially in the cases of the studies where they reported age spans, there are some studies ([16–18, 28, 33, 42, 48, 49, 51, 60, 80]) that include participants outside our age range, but we decided to include these when at least 70% of their participants were between the ages 16 and 25 or the mean/average age was within our age range.

Smartphones, desktops, and a combination of smartphones and desktops were the most common hardware platforms. We were surprised by the lack of studies that utilized virtual reality (VR) and augmented reality (AR) technologies for their chatbots. We also noted the average age of participants in the studies respective to these, and found that the studies using smartphones had slightly younger participants (19.8 years old).

In 13 studies [16–18, 31, 33, 37, 42, 43, 45, 49, 73, 81, 85], Facebook Messenger or WhatsApp was used to deploy self-developed chatbots. The researchers used these common messaging apps because it is easy to deploy chatbots to them, and they are also applications that a lot of people have access to and use.

4.1 User Mental Health and Therapeutic Tools

We were interested in what aspects of users' mental health were studied, as well as what therapeutic tools were used. There was

³<https://woebothealth.com/>

a wide variety of both mental health conditions/symptoms and therapeutic tools used to target them. Studies targeting general mental health and well-being, depression, and anxiety make up 60% (35/58) of the set of studies in this review (e.g. [3, 6, 72, 73]). There were 13 studies targeting more specific mental health conditions, such as academic stress (e.g. [14, 53]) or isolation (e.g. [17, 31]). The remaining ten papers did not clearly specify which mental health conditions or symptoms the studies were targeting [4, 20, 25, 29, 34, 36, 39, 83, 90, 91]. Concerning therapeutic tools, eight of the studies used different self-help methods, such as self-disclosure, self-reflection, self-awareness, and self-compassion as part of their chatbots [10, 23, 28, 36, 42, 43, 45, 89], and these focused on mitigating general mental health or well-being support, emotional regulation, anxiety and depression.

We note that stress-related mental health conditions in the studies generally targeted people between the ages of 16–19 (e.g. [14, 46, 48, 51]). Studies that target anxiety, depression, and general mental health conditions included participants aged 22–24 (e.g. [19, 23, 24, 71, 85]). In the studies that target isolation or loneliness, the spread of ages was wider, with most studies having a range or an average of between 14–20 years old (e.g. [17, 31, 32]).

In the studies where the therapeutic tools included self-help methods, the participants were slightly older, with most averaging between 21 and 24 years old (e.g. [23, 28, 36, 43]). This could be due to older teenagers and young adults generally understanding their own feelings better than younger teenagers do [15], and these kinds of methods are therefore more effective on them.

4.2 Study Types and Research Methods

We investigated the types of studies and the research methods used. A total of 26 studies were field studies, i.e., when the users used the chatbot in a real-world context (e.g. [3, 6, 14, 17, 24, 25, 27]). In 12 of these, general mental health conditions were covered [3, 6, 17, 31, 38, 43, 49, 71, 73, 80, 85, 89], and eight covered anxiety and/or depression specifically [24, 27, 35, 37, 60, 65, 71, 85]. The reason why field studies focused on general mental health and anxiety could be that field studies usually provide intervention over a longer period of time than in-lab studies, and those conditions generally need those longer periods of time to show any results.

Of all the studies, 27 were in-lab studies, under controlled experimental conditions, often with researchers present (e.g. [4, 32, 59, 69]). In 13 of them, general mental health conditions were covered [10, 18, 32, 40, 42, 43, 48, 59, 64, 69, 72, 79, 88], and eight covered anxiety and/or depression [16, 23, 32, 33, 48, 66, 81, 84]. As mentioned above, it is harder to see the effects of using a chatbot on user mental health when the user has only used it for a short period of time, which is often the case in in-lab studies. However, many of the in-lab studies that targeted general mental health conditions were conducted as a first study to test out their chatbot and understand if users preferred it to others, and what aspects of the chatbot were appealing to the users (e.g. [16, 28, 59, 64]).

Four studies were design studies, i.e., with a focus on the design of the chatbot [19, 26, 78, 87]. These studies all targeted either general mental health or anxiety and depression, and their aims were to evaluate aspects of the chatbot's design, which was done with the help of co-design in all four studies.

Most studies used a mix of research methods. Surveys were used in 69% of the studies (40/58) and in most cases together with other methods. Most often, it was used together with interviews and/or evaluations (evaluating whether the chatbot was good or if it helped with what they said it would), such as in [16, 17, 24, 32] and/or a comparative study (comparing the chatbot with different parameters, or comparing it to other tools for mental health support), such as in [27, 39, 40, 51].

The sample sizes in the studies varied, with 69% of them (40/58) having fewer than 100 participants (e.g. [10, 18, 36, 69]). Four studies had significantly larger sample sizes: 622 [90], 1054 [66], 10387 [80], and 17838 [25]. Three of the latter set of studies used online surveys as their data collection method [25, 66, 90]; therefore, the larger sample sizes are not surprising.

4.3 Different Relationships with Chatbots

We were interested in how the different relationships with chatbots were portrayed in the studies; a summary of these can be found in Table 1. We found that 41% (24/58) of the papers did not clearly mention the type of relationship to the chatbot (which could mean that the chatbot did not establish a relationship, or if it did, it was not reported). Nineteen papers reported on friendship of some kind (4 had friendship in focus, 15 had mentions of friendship), and 15 papers reported on therapeutic relationships of some kind (10 had therapeutic relationships in focus, 5 had mentions of therapeutic relationships). None of the papers included described the relationship to the chatbot as romantic.

In the four studies where friendship was in focus, the researchers described the chatbot as a friend, mentioning how it can act as a friend or that the users will think of it as a friend [29, 36, 46, 78].

In the five instances where therapeutic relationships were only mentioned, the papers refer in passing to a therapeutic alliance between the user and the chatbot, or note that it would act as a coach or a counselor [27, 28, 45, 65, 71]. For the ten studies where the therapeutic relationship was in focus, the studies targeted general mental health and well-being, anxiety and depression, or more specific cases such as academic stress due to the COVID-19 pandemic [20, 40, 44, 53, 55, 59, 60, 64].

Type of Relationship	Citation
Friendship Mentioned	[3, 10, 14, 23, 34, 38, 39, 42, 43, 66, 73, 80, 83, 85, 87]
Friendship in Focus	[29, 36, 46, 78]
Therapeutic Mentioned	[28, 45, 65, 71]
Therapeutic in Focus	[4, 19, 20, 40, 44, 53, 55, 59, 60, 64]
Not Mentioned	[6, 16–18, 24–26, 31–33, 35, 37, 48, 49, 51, 69, 72, 79, 81, 84, 88–91]

Table 1: Different types of relationships covered in the studies

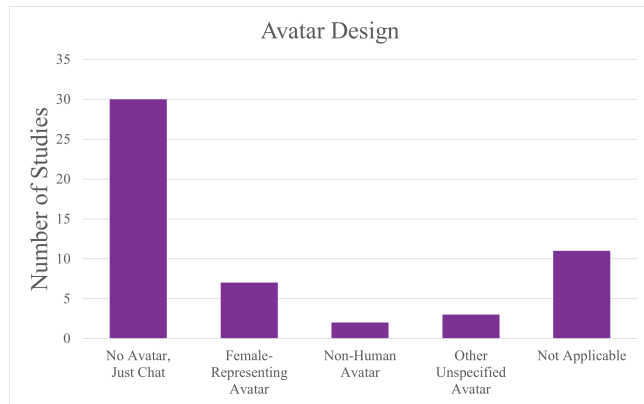


Figure 1: Design of the chatbot avatars

4.4 Chatbot Characteristics

In this section, we use the term **avatar**, meaning the *visual, embodied representation of a chatbot* [61]. In 11 papers, the interface of the chatbot was undefined [23, 25, 27, 38, 53, 66, 69, 71, 79, 89, 90]. In 47% (27/58) of the studies, the chatbot had a text-based chat interface as the only interaction modality (e.g. [48, 60, 65]). In some of the studies, the chat was complemented by other modalities. In six studies, speech (where the user speaks to the chatbot live or records audio messages) was used either in addition to text [28, 39, 79] or on its own [19, 20, 88]. Some also had other communication modalities alongside text, such as exercises [85], games [49], or mental health resources in the form of information text [49].

We were especially interested in the chatbots that used avatars, as we think that avatars can be impactful parts of the user experience of using a chatbot, and especially important for any social connections with it. In 16 studies an avatar was included [4, 19, 20, 26, 28, 36, 39, 49, 59, 64, 78, 83, 83–85, 88, 91]. See Fig. 1 for the spread of the design of the chatbot’s avatar.

5 Discussion

We will now discuss the most interesting findings, organized by the research questions, and show how they lead to our recommendations for chatbot research and development.

5.1 Chatbots for Mental Health

RQ1. What aspects of user mental health have been addressed in the research on chatbots, and what therapeutic tools are used for them? The results showed a wide variety of mental health conditions that are being targeted with various therapeutic tools. In many cases, the chatbot itself could be seen as a tool, supporting users with different self-help methods [10, 23, 28, 36, 42, 43, 45, 89], or with specific tasks such as journaling [35]. In the cases where the chatbot acted as a therapist rather than just a tool, we argue that this calls for caution in future research. For example, can a chatbot be a replacement for seeing a therapist, or should it be a supplement? This connects back to the research mentioned in related work; on one hand, therapeutic chatbots can mitigate the lack of sufficient therapeutic resources, but on the other hand, care seekers seem to have a low level of engagement with these chatbots, and there are

concerns regarding the efficacy of these interventions as well as potential ethical issues. Therapist-guided chatbots, acting as a tool under human control, might be a viable middle ground. However, as noted in related work, this requires extensive changes to the integration of therapeutic chatbots into existing healthcare, both technically and organizationally.

5.2 Study Designs

RQ2. What study designs have been used, and what are the most common research methods used in the studies? There was a wide spread of research methods, something that could be a result of the interdisciplinary body of research and the wide variety of studies included here. It is beneficial that many different research methods are used and reported, and as highlighted in Section 2 Related Work, this has been emphasized in earlier literature reviews as well [2, 41, 75, 76]. We however note the lack of design studies, and hope to see more of these types of studies conducted on chatbots and the mental health in young people. Insights and needs from the young people themselves should be taken into account, and research-by-design methods that can involve end users - such as participatory design and design fiction - could be especially useful.

5.3 Relationships with the Chatbot

RQ3. What are the most common relationship types with chatbots in the studies? We note that the companionship aspect is missing in many of the studies. We were interested in whether the chatbot had memory and initiative, two traits that are important for social interaction and social bonding [9, 11]. Three studies had chatbots with memory [24, 39, 83], but only one of these also had initiative [24]. In 13 studies, the chatbot took initiative [4, 6, 18, 23, 24, 37, 49, 55, 59, 78, 80, 81, 85]. So, these two crucial aspects of companionship are missing in the majority of studies. Also, a large number of studies either do not utilize companionship at all or have not defined or explained the kind of relationship between the user and the chatbot. We argue for caution when it comes to therapeutic chatbots and companionship, especially since previous work highlighted the potential ethical issues [1, 2, 41, 47, 70, 75, 76]. But also considering the huge problem with loneliness and social anxiety [86], companionship with chatbots might be an advantageous intervention in some mental health situations. Indeed, there are already studies that have investigated the effect of relationships with chatbots [12, 63, 74], but they do not specifically look at the mental health implications of this use, and they do not specifically look at a young user base, which is what we were anticipating to find (but did not find to any large extent) in this literature review.

5.4 Chatbot Characteristics and Interactions

RQ4. What characteristics and interaction modalities of chatbots can be found in the studies? Most of the chatbots in this review had chat interfaces and no other interaction or communication modalities. While chatting or texting are common ways of interacting with other people online, we wonder why so few studies include speech (such as audio recordings, video calls with the AI, and similar). One reason could be that it is more difficult to implement such a feature technically, and in many of the studies, the researchers themselves

were developing their chatbot, so they may not have had enough time or resources to add such features.

The design of an avatar impacts the user interaction [20, 67], and in particular, the gender representation of the avatar affects trust and rapport [21, 22]. Of the 16 papers that reported on an avatar design, seven of them had a female-representing avatar [4, 28, 39, 49, 59, 85, 91] (see Fig. 1), while none were male-representing and the remaining avatars had a variety of non-human designs. In one paper [28], the argument was made that previous research has shown that people generally prefer female-representing avatars when they meet them in a medical context due to female-representing avatars being perceived as having a more caring communication style, thus leading users to quicker self-disclosure and touching on more emotional subjects [52]. We wonder why it is not more common to use an avatar to represent the chatbot, and we also note that it is often not properly rationalized why the avatar looks the way it does. This might invite bias and reinforce gender stereotypes, which can have negative effects on the users, in particular young users who interact with them. We therefore encourage designers and researchers to motivate why they design their avatars in a certain way and to be aware of whether they are reinforcing gender or other stereotypes.

5.5 Recommendations for Chatbot Research and Development

We present recommendations that designers, developers, and researchers of chatbots for youth mental health should take into consideration, based on the findings from existing literature and the previous discussion points made.

- Gender issues in relation to chatbots should be explored more in depth to reduce bias and avoid reinforcing stereotypes, especially considering our findings regarding female-representing avatars.
- More research on whether companionship with a chatbot can mitigate loneliness and how that will affect young people should be carried out. However, when considering companionship with a chatbot, one should also be aware of the ethical concerns regarding the blurring of the lines between a companion and a therapist, and the need for integration and interoperability with existing healthcare solutions.
- Interdisciplinary research should be continued in this field, with an emphasis on establishing common therapeutic frameworks and widespread establishment of common evaluation methods. To complement the already interdisciplinary field, we recommend more design research on the area of chatbots and mental health to include new perspectives. Especially participatory design and design fiction enable a deeper focus on the users, the young people, and their insights. Design research can complement scientific experimentation by generating situated knowledge that explores how things could be in real contexts, such as using iterative, hands-on prototyping and reflective methods.

6 Conclusions and Future Work

In this systematic literature review, we have summarized and drawn conclusions based on 58 papers that cover chatbots, mental health,

and young people over the last decade of research. By including literature from both HCI and Psychology, we have gained a more holistic view of how chatbots have been studied in relation to young users and their mental health.

Our findings have shown that general mental health is the most common condition addressed in the literature, followed by anxiety and depression, and loneliness or isolation. A wide variety of therapeutic tools implemented into the chatbots have been used to target these conditions, some of the more common ones being different kinds of self-help methods.

Given the interdisciplinary field, many different research methods have been used in the studies, which we see as positive. Most commonly, surveys, interviews, and evaluations were used. We found an almost even spread between field studies and in-lab studies, and noted a lack of design-focused studies.

The relationship to the chatbot was not always described in the papers. When it was described, there was a mix of therapeutic relationships and friendship.

A majority of the chatbots in the studies had chat interfaces, and few reported on an avatar design. We offer a reflection on why it is important for researchers to thoroughly consider the design of their avatars so as to avoid biases and stereotypes.

Limitations: One of the main limitations of this study is that we might have missed relevant literature due to the databases. Had we searched on multiple or larger medical/psychology databases in particular, our pool of publications could have been larger. However, we believe that the set of papers we found from Web of Science gave a sufficient overview. We also omitted searching more technical databases, which could also have given us a wider set of papers. We chose not to look for technical papers, however, as they would be outside of the scope of this study. We also acknowledge that the coding scheme omitted detailed participant characteristics, but these were not always mentioned in the papers, so any data gathered would have been inconsistent.

Future Work: As mentioned in our recommendations, future work should carry on the interdisciplinary work regarding chatbots and young people's mental health, focusing on how different design methods, such as participatory design and design fiction, could bring more insights into young people's uses of chatbots and their effects on mental health. Future work should also study the impact of avatar design on the interactions between young people and chatbots, and identify how the use of both therapeutic and friendship relationships affects young people's mental health. We see opportunities to research the benefits and drawbacks of this technology while it is still being developed, allowing researchers to play a significant and much-needed role in shaping the further development of chatbots for mental health.

Acknowledgments

Mohammad Obaid is partially funded by the Swedish Research Council (VR), Project ID: 2020-04918.

References

- [1] Jennifer Apolinário-Hagen, Jessica Kemper, and Carolina Stürmer. 2017. Public Acceptability of E-Mental Health Treatment Services for Psychological Problems: A Scoping Review. *JMIR Ment Health* 4, 2 (2017), e10.
- [2] Filip K. Arnberg, Steven J. Linton, Monica Hulterantz, Emelie Heintz, and Ulf Jonsson. 2014. Internet-Delivered Psychological Treatments for Mood and Anxiety

- Disorders: A Systematic Review of Their Efficacy, Safety, and Cost-Effectiveness. *PLOS ONE* 9, 5 (2014), e98118.
- [3] Petter Bae Bae Brandtzæg, Marita Skjuve, Kim Kristoffer Kristoffer Dysthe, and Asbjørn Følstad. 2021. When the social becomes non-human: young people's perception of social support in chatbots. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–13.
- [4] Mukesh Barange, Sandratra Rasendrasoa, Maël Bouabdelli, Julien Saunier, and Alexandre Pauchet. 2022. Impact of adaptive multimodal empathic behavior on the user interaction. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [5] Sandhya Bhatt. 2025. Digital mental health: Role of artificial intelligence in psychotherapy. *Annals of Neurosciences* 32, 2 (2025), 117–127.
- [6] Ananya Bhattacharjee, Joseph Jay Williams, Jonah Meyerhoff, Harsh Kumar, Alex Mariakakis, and Rachel Kornfield. 2023. Investigating the role of context in the delivery of text messages for supporting psychological wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [7] Annabel Blake, Marcus Carter, and Eduardo Velloso. 2025. *Rapid Analysis - Character AI & Children*. Report.
- [8] Sarah-Jayne Blakemore. 2019. Adolescence and mental health. *The lancet* 393, 10185 (2019), 2030–2031.
- [9] Susan Bluck. 2003. Autobiographical memory: Exploring its functions in everyday life. *Memory* 11, 2 (2003), 113–123.
- [10] Wanling Cai, Yucheng Jin, Xianglin Zhao, and Li Chen. 2023. “Listen to Music, Listen to Yourself”: Design of a Conversational Agent to Support Self-Awareness While Listening to Music. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [11] Samuel Rhys Cox, Yi-Chieh Lee, and Wei Tsang Ooi. 2023. Comparing How a Chatbot References User Utterances from Previous Chatting Sessions: An Investigation of Users' Privacy Concerns and Perceptions. In *Proceedings of the 12th International Conference on Human-Agent Interaction*. 105–114.
- [12] Emmelyn AJ Croes and Marjolijn L Antheunis. 2021. Can we be friends with Mitsu? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships* 38, 1 (2021), 279–300.
- [13] Guoqiang Cui, Barbara Lockee, and Cuiqing Meng. 2013. Building modern online social presence: A review of social presence theory and its instructional design implications for future trends. *Education and information technologies* 18, 4 (2013), 661–685.
- [14] Johan Oswin De Nieva, Jose Andres Joaquin, Chaste Bernard Tan, Ruzel Khyvin Marc Te, and Ethel Ong. 2020. Investigating students' use of a mental health chatbot to alleviate academic stress. In *6th International ACM In-Cooperation HCI and UX Conference*. 1–10.
- [15] Andreas Demetriou and Karin Bakrachevic. 2009. Reasoning and self-awareness from adolescence to middle age: Organization and development as a function of education. *Learning and Individual Differences* 19, 2 (2009), 181–194.
- [16] Gilly Dosovitsky and Eduardo Bunge. 2023. Development of a chatbot for depression: adolescent perceptions and recommendations. *Child and adolescent mental health* 28, 1 (2023), 124–127.
- [17] César G Escobar-Viera, Giovanna Porta, Robert WS Coulter, Jamie Martina, Jeremy Goldbach, and Bruce L Rollman. 2023. A chatbot-delivered intervention for optimizing social media use and reducing perceived isolation among rural-living LGBTQ+ youth: development, acceptability, usability, satisfaction, and utility. *Internet Interventions* 34 (2023), 100668.
- [18] Katrin E Fabian, Katherine T Foster, Lydia Chwastiak, Morgan Turner, and Bradley H Wagenaar. 2023. Adapting a transdiagnostic digital mental health intervention for use among immigrant and refugee youth in Seattle: a human-centered design approach. *Translational Behavioral Medicine* 13, 11 (2023), 867–875.
- [19] Pedro Guillermo Feijóo-García, Chase Wrenn, Alexandre Gomes de Siqueira, Rashi Ghosh, Jacob Stuart, Heng Yao, and Benjamin Lok. 2024. Exploring the Effects of User-Agent and User-Designer Similarity in Virtual Human Design to Promote Mental Health Intentions for College Students. *ACM Transactions on Applied Perception* 22, 1 (2024), 1–41.
- [20] Pedro Guillermo Feijóo-García, Chase Wrenn, Jacob Stuart, Alexandre Gomes de Siqueira, and Benjamin Lok. 2023. Participatory design of virtual humans for mental health support among North American computer science students: Voice, appearance, and the similarity-attraction effect. *ACM Transactions on Applied Perception* 20, 3 (2023), 1–27.
- [21] Pedro Guillermo Feijóo-García, Chase Wrenn, Sterling Kalogeras, Chandler Payne, Benjamin Lok, and Olufisayo Omojokun. 2024. Effects of Gender Synchrony in User-Agent Interactions: Integrating the Designer as a Product Cue in Virtual Human Design for Mental Health Support. In *Proceedings of the 12th International Conference on Human-Agent Interaction*. 123–131.
- [22] Pedro Guillermo Feijóo-García, Mohan Zalake, Alexandre Gomes de Siqueira, Benjamin Lok, and Felix Hamza-Lup. 2021. Effects of Virtual Humans' Gender and Spoken Accent on Users' Perceptions of Expertise in Mental Wellness Conversations. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 68–75.
- [23] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health* 4, 2 (2017), e7785.
- [24] Russell Fulmer, Angela Joerin, Breanna Gentile, Lysanne Lakerink, Michiel Rauws, et al. 2018. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health* 5, 4 (2018), e9782.
- [25] Elton Fayiah Gbolli, Jason Bantjes, Lucy Jarvis, Sonja Swandevelder, Jean Du Plessis, Richard Shadwell, Charl Davids, Rone Gerber, Nuhaa Holland, and Xanthe Hunt. 2023. Intention to use digital mental health solutions: A cross-sectional survey of university students attitudes and perceptions toward online therapy, mental health apps, and chatbots. *Digital Health* 9 (2023), 20552076231216559.
- [26] Christine Grové. 2021. Co-developing a mental health and wellbeing chatbot with and for young people. *Frontiers in psychiatry* 11 (2021), 606041.
- [27] Yuhao He, Li Yang, Xiaokun Zhu, Bin Wu, Shuo Zhang, Chunlian Qian, and Tian Tian. 2022. Mental health chatbot for young adults with depressive symptoms during the COVID-19 pandemic: single-blind, three-arm randomized controlled trial. *Journal of medical Internet research* 24, 11 (2022), e40719.
- [28] Katherine Hopman, Deborah Richards, and Melissa M Norberg. 2023. A digital coach to promote emotion regulation skills. *Multimodal Technologies and Interaction* 7, 6 (2023), 57.
- [29] Hui-Wen Huang and Jessica Chang. 2024. Human-AI Interactions in Teacher Education: Examining Social Presence and Friendship. In *Proceedings of the 2024 International Conference on Artificial Intelligence and Teacher Education*. 64–69.
- [30] Yiting Huang and Hanyun Huang. 2025. Exploring the effect of attachment on technology addiction to generative AI chatbots: A structural equation modeling analysis. *International Journal of Human-Computer Interaction* 41, 15 (2025), 9440–9449.
- [31] Annie Kang, Sarah Hetrick, Tania Cargo, Sarah Hopkins, Nicola Ludin, Sarah Bodmer, Kiani Stevenson, Chester Holt-Quick, and Karolina Stasiak. 2023. Exploring Young Adults' Views About Aroha, a Chatbot for Stress Associated With the COVID-19 Pandemic: Interview Study Among Students. *JMIR Formative Research* 7 (2023), e44556.
- [32] Boyoung Kang and Munpyo Hong. 2025. Development and Evaluation of a Mental Health Chatbot Using ChatGPT 4.0: Mixed Methods User Experience Study With Korean Users. *JMIR Medical Informatics* 13 (2025), e63538.
- [33] Stanisław Karkosz, Robert Szymański, Katarzyna Sanna, Jarosław Michałowski, et al. 2024. Effectiveness of a web-based and mobile therapy chatbot on anxiety and depressive symptoms in subclinical young adults: randomized controlled trial. *JMIR formative research* 8, 1 (2024), e47960.
- [34] Liam Kettle and Yi-Ching Lee. 2024. User experiences of well-being chatbots. *Human Factors* 66, 6 (2024), 1703–1723.
- [35] Taewon Kim, Seolyeong Bae, Hyun Ah Kim, Su-woo Lee, Hwajung Hong, Chanmo Yang, and Young-Ho Kim. 2024. MindfulDiary: Harnessing Large Language Model to Support Psychiatric Patients' Journaling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–20.
- [36] Taewon Kim, Mintra Ruensuk, and Hwajung Hong. 2020. In helping a vulnerable bot, you help yourself: Designing a social bot as a care-receiver to promote mental health and reduce stigma. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
- [37] Maria Carolina Klos, Milagros Escoredó, Angela Joerin, Viviana Noemi Lemos, Michiel Rauws, and Eduardo L Bunge. 2021. Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial. *JMIR formative research* 5, 8 (2021), e20678.
- [38] Theodora Koulouri, Robert D Macredie, and David Olakitan. 2022. Chatbots to support young adults' mental health: an exploratory study of acceptability. *ACM Transactions on Interactive Intelligent Systems (TiIS)* 12, 2 (2022), 1–39.
- [39] Nikola Kovacevic, Christian Holz, Markus Gross, and Rafael Wampfler. 2024. On Multimodal Emotion Recognition for Human-Chatbot Interaction in the Wild. In *Proceedings of the 26th International Conference on Multimodal Interaction*. 12–21.
- [40] Mohammad Amin Kuhail, Nazik Alturki, Justin Thomas, and Amal K Alkhalifa. 2024. Human vs. AI counseling: College students' perspectives. *Computers in Human Behavior Reports* 16 (2024), 100534.
- [41] Liliana Laranjo, Adam G. Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie Y. S. Lau, and Enrico Coiera. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 25, 9 (2018), 1248–1258.
- [42] Minha Lee, Sander Ackermans, Nena Van As, Hanwen Chang, Enzo Lucas, and Wijnand IJsselstein. 2019. Caring for Vincent: a chatbot for self-compassion. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [43] Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2020. Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–27.
- [44] Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2021. Exploring the effects of incorporating human experts to deliver journaling guidance through a chatbot.

- Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.
- [45] Yi-Chieh Lee, Naomi Yamashita, Yun Huang, and Wai Fu. 2020. "I hear you, I feel you": encouraging deep self-disclosure through a chatbot. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–12.
 - [46] Carlos Miguel Legaspi Jr, Tristan Raphael Pacana, Kyle Loja, Christina Sing, and Ethel Ong. 2022. User perception of Wysa as a mental well-being support tool during the COVID-19 pandemic. In *Proceedings of the Asian HCI Symposium 2022*. 52–57.
 - [47] Han Li, Renwen Zhang, Yi-Chieh Lee, Robert E. Kraut, and David C. Mohr. 2023. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *npj Digital Medicine* 6, 1 (2023), 236.
 - [48] Irene Lopatovska, Olivia Turpin, Jessica Davis, Ellen Connell, Chris Denney, Hilda Fournier, Archana Ravi, Ji Hee Yoon, and Eesha Parasnis. 2022. Capturing teens' voice in designing supportive agents. In *Proceedings of the 4th Conference on Conversational User Interfaces*. 1–12.
 - [49] Nicola Ludin, Chester Holt-Quick, Sarah Hopkins, Karolina Stasiak, Sarah Hetrick, Jim Warren, and Tania Cargo. 2022. A chatbot to support young people during the COVID-19 pandemic in New Zealand: evaluation of the real-world rollout of an open trial. *Journal of Medical Internet Research* 24, 11 (2022), e38743.
 - [50] Bethanie Maples, Merve Cerit, Aditya Vishwanath, and Roy Pea. 2024. Loneliness and suicide mitigation for students using GPT3-enabled chatbots. *npj Mental Health Research* 3, 1 (2024), 4.
 - [51] Audrey Mariamo, Caroline Elizabeth Temcheff, Pierre-Majorique Léger, Sylvain Senecal, Marianne Alexandra Lau, et al. 2021. Emotional reactions and likelihood of response to questions designed for a mental health chatbot among adolescents: experimental study. *JMIR human factors* 8, 1 (2021), e24343.
 - [52] Marianne Schmid Mast, Judith A Hall, and Debra L Roter. 2007. Disentangling physician sex and physician communication style: their effects on patient satisfaction in a virtual medical visit. *Patient education and counseling* 68, 1 (2007), 16–22.
 - [53] Jingbo Meng, Minjin Rheu, Yue Zhang, Yue Dai, and Wei Peng. 2023. Mediated social support for distress reduction: AI Chatbots vs. Human. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–25.
 - [54] Pekka Mertala, Sonsoles López-Pernas, Henriikka Vartiainen, Mohammed Saqr, and Matti Tedre. 2024. Digital natives in the scientific literature: A topic modeling approach. *Computers in Human Behavior* 152 (2024), 108076.
 - [55] Varun Mishra, Sarah Hong, and David Kotz. 2024. Exploring the Relationship Between Intrinsic Motivation and Receptivity to mHealth Interventions. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 437–443.
 - [56] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. 2010. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International journal of surgery* 8, 5 (2010), 336–341.
 - [57] Clifford Nass and Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of social issues* 56, 1 (2000), 81–103.
 - [58] Gina Neff and Peter Nagy. 2016. Talking to bots: symbiotic agency and the case of Tay. *International journal of communication (Online)* (2016), 4915.
 - [59] Shreeya Nelekar, Amal Abdulrahman, Manik Gupta, and Deborah Richards. 2022. Effectiveness of embodied conversational agents for managing academic stress at an Indian University (ARU) during COVID-19. *British Journal of Educational Technology* 53, 3 (2022), 491–511.
 - [60] Ginger Nicol, Ruoyun Wang, Sharon Graham, Sherry Dodd, Jane Garbutt, et al. 2022. Chatbot-delivered cognitive behavioral therapy in adolescents with depression and anxiety during the COVID-19 pandemic: feasibility and acceptability study. *JMIR Formative Research* 6, 11 (2022), e40242.
 - [61] André Nusselder. 2009. *Interface Fantasy - a Lacanian Cyborg Ontology*. MIT Press, London.
 - [62] Tina M Olsson and Therése Skoog. 2025. Differences in wellbeing patterns among late adolescent boys and girls with and without experience of out-of-home care. *Vulnerable Children and Youth Studies* (2025), 1–20.
 - [63] Shuyi Pan and Yi Mou. 2024. Constructing the meaning of human–AI romantic relationships from the perspectives of users dating the social chatbot Replika. *Personal Relationships* 31, 4 (2024), 1090–1112.
 - [64] Johanna Peltola, Kirsikka Kaipainen, Katariina Keinonen, Noona Kiuru, and Markku Turunen. 2023. Developing a conversational interface for an ACT-based online program: Understanding adolescents' expectations of conversational style. In *Proceedings of the 5th International Conference on Conversational User Interfaces*. 1–16.
 - [65] Yue Qi. 2025. Pilot quasi-experimental research on the effectiveness of the Woebot AI chatbot for reducing mild depression symptoms among athletes. *International Journal of Human-Computer Interaction* 41, 1 (2025), 452–459.
 - [66] Samira Rasouli, Moojan Ghafurian, Elizabeth S Nilsen, and Kerstin Dautenhahn. 2024. University students' opinions on using intelligent agents to cope with stress and anxiety in social situations. *Computers in Human Behavior* 153 (2024), 108072.
 - [67] Deborah Richards, Bayan Alsharbi, and Amal Abdulrahman. 2020. Can I help you? Preferences of young adults for the age, gender and ethnicity of a Virtual Support Person based on individual differences including personality and psychological state. In *Proceedings of the Australasian Computer Science Week Multiconference*. 1–10.
 - [68] Debra J Rickwood, Kelly R Mazzer, and Nic R Telford. 2015. Social influences on seeking help from mental health services, in-person and online, during adolescence and young adulthood. *BMC psychiatry* 15, 1 (2015), 40.
 - [69] Erin Roehrer, Panatchakorn Pokawinkoon, Paul Watters, James D Sauer, and Joel Scanlan. 2024. Adolescent-Centric Design of an Online Safety Chatbot. *Journal of Computer Information Systems* (2024), 1–14.
 - [70] Ayanda Rogge. 2023. Defining, Designing and Distinguishing Artificial Companions: A Systematic Literature Review. *International Journal of Social Robotics* 15, 9 (2023), 1557–1579.
 - [71] Alicia Salamanca-Sanabria, Ahmad Ishqi Jabir, Xiaowen Lin, Aishah Alattas, A Baki Kocaballi, Jimmy Lee, Tobias Kowatsch, and Lorainne Tudor Car. 2023. Exploring the perceptions of mHealth interventions for the prevention of common mental disorders in university students in Singapore: qualitative study. *Journal of medical internet research* 25 (2023), e44542.
 - [72] Anita Schick, Jasper Feine, Stefan Morana, Alexander Maedche, and Ulrich Reininghaus. 2022. Validity of chatbot use for mental health assessment: experimental study. *JMIR mHealth and uHealth* 10, 10 (2022), e28082.
 - [73] Dominic Ethan Sia, Marco Jalen Yu, Justine Leo Daliva, Jaycee Montenegro, and Ethel Ong. 2021. Investigating the acceptability and perceived effectiveness of a chatbot in helping students assess their well-being. In *Proceedings of the Asian CHI Symposium 2021*. 34–40.
 - [74] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg. 2022. A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies* 168 (2022), 102903.
 - [75] Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems. *ACM Trans. Comput.-Hum. Interact.* 27, 5 (2020).
 - [76] John Torous, Sandra Bucci, Imogen H. Bell, Lars V. Kessing, Maria Faurholt-Jepsen, Pauline Whelan, Andre F. Carvalho, Matcheri Keshavan, Jake Linardon, and Joseph Firth. 2021. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry* 20, 3 (2021), 318–335.
 - [77] Pablo Vicente Torres-Carrión, Carina Soledad González-González, Silvana Aciar, and Germania Rodríguez-Morales. 2018. Methodology for systematic literature review applied to engineering and education. In *2018 IEEE Global engineering education conference (EDUCON)*. IEEE, 1364–1373.
 - [78] Jordi Tost, Rahel Flechtner, Rahel Maué, and Frank Heidmann. 2024. Caring for a companion as a form of self-care. Exploring the design space for irritating companion technologies for mental health. In *Proceedings of the 13th Nordic Conference on Human-Computer Interaction*. 1–15.
 - [79] Amy JC Trappey, Aislyn PC Lin, Kevin YK Hsu, Charles V Trappey, and Kevin LK Tu. 2022. Development of an empathy-centric counseling chatbot system capable of sentimental dialogue analysis. *Processes* 10, 5 (2022), 930.
 - [80] Dana Vertsberger, Navot Naor, Mirène Winsberg, et al. 2022. Adolescents' well-being while using a mobile artificial intelligence–powered acceptance commitment therapy tool: evidence from a longitudinal study. *Jmir Ai* 1, 1 (2022), e38171.
 - [81] Anna Viduani, Victor Cosenza, Helen L Fisher, Claudia Buchweitz, Jader Piccin, Rivka Pereira, Brandon A Kohrt, Valeria Mondelli, Alastair van Heerden, Ricardo Matsumura Araújo, et al. 2023. Assessing mood with the identifying depression early in adolescence Chatbot (IDEABot): development and implementation study. *JMIR human factors* 10 (2023), e44388.
 - [82] Emily A Vogels, Risa Gelles-Watnick, and Navid Massarat. 2022. Teens, social media and technology 2022. *Pew Research Center* 10 (2022).
 - [83] Wout Vossen, Maxwell Szymanski, and Katrien Verbert. 2024. The effect of personalizing a psychotherapy conversational agent on therapeutic bond and usage intentions. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 761–771.
 - [84] Xueyang Wang, Runyan Tian, Qiuyi Zeng, Chenye Tu, Shuning Zhang, Xin Yi, Hewu Li, and Pei-Luen Patrick Rau. 2024. The Synergy of Dialogue and Art: Exploring the Potential of Multimodal AI Chatbots in Emotional Support. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing*. 147–153.
 - [85] Ruth Williams, Sarah Hopkins, Chris Frampton, Chester Holt-Quick, Sally Nicola Merry, and Karolina Stasiak. 2021. 21-day stress detox: open trial of a universal well-being chatbot for young adults. *Social Sciences* 10, 11 (2021), 416.
 - [86] World Health Organization. 2024. Mental health of adolescents. <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health>. Accessed: 2025-04-28.
 - [87] Aimee-Rose Wrightson-Hester, Georgia Anderson, Joel Dunstan, Peter M McEvoy, Christopher J Sutton, Bronwyn Myers, Sarah Egan, Sara Tai, Melanie Johnston-Hollitt, Wai Chen, et al. 2023. An artificial therapist (Manage Your Life Online) to support the mental health of youth: Co-design and case series. *JMIR Human Factors* 10 (2023), e46849.

- [88] Kenji Yokotani, Gen Takagi, and Kobun Wakashima. 2018. Advantages of virtual agents over clinical psychologists during comprehensive mental health interviews using a mixed methods design. *Computers in human behavior* 85 (2018), 135–145.
- [89] Christopher You, Rashi Ghosh, Andrew Maxim, Jacob Stuart, Eric Cooks, and Benjamin Lok. 2022. How does a virtual human earn your trust? guidelines to improve willingness to self-disclose to intelligent virtual agents. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [90] Jordyn Young, Laala M Jawara, Diep N Nguyen, Brian Daly, Jina Huh-Yoo, and Afsaneh Razi. 2024. The role of ai in peer support for young people: A study of preferences for human-and ai-generated responses. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [91] Mohan Zalake, Alexandre Gomes De Siqueira, Krishna Vaddiparti, Pavlo Antonenko, and Benjamin Lok. 2021. Towards understanding how virtual human's verbal persuasion strategies influence user intentions to perform health behavior. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 216–223.