

Artificial Intelligence-Based Mobile Phone Applications for Child Mental Health: A Comprehensive Review and Content Analysis

Fan Yang, Jianan Wei, Xuejun Zhao, Ruopeng An

Submitted to: JMIR mHealth and uHealth
on: March 28, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 24

0..... 24

0..... 24

Figures 25

Figure 1..... 26

Figure 2..... 27

Artificial Intelligence-Based Mobile Phone Applications for Child Mental Health: A Comprehensive Review and Content Analysis

Fan Yang¹ phd, MSW; Jianan Wei²; Xuejun Zhao³; Ruopeng An⁴

¹ School of Social Work University of Illinois Urbana Champaign Urbana US

² School of Public Administration Dongbei University of Finance and Economics Dalian CN

³ School of Computer Science and Engineering Macau University of Science and Technology Macau CN

⁴ New York University Silver School of Social Work Constance and Martin Silver Center on Data Science and Social Equity, New York, NY US

Corresponding Author:

Ruopeng An

New York University Silver School of Social Work
Constance and Martin Silver Center on Data Science and Social Equity,
726 Broadway
New York, NY
US

Abstract

Background: Mobile phone apps powered by Artificial Intelligence (AI) have emerged as powerful tools to address mental health challenges faced by children. This study aimed to comprehensively review AI-driven apps for child mental health, assessing their availability, quality, readability, characteristics, and functions.

Objective: This study aimed to comprehensively review AI-driven apps for child mental health, assessing their availability, quality, readability, characteristics, and functions.

Methods: Utilizing a systematic review approach, we initially identified 369 apps, which, after screening eligibility, resulted in 27 apps being included in this study. Quality was evaluated using the Mobile Application Rating Scale (MARS). A readability calculator was implemented to assess readability by utilizing the count of words, syllables, and sentences to generate a score indicative of the reading difficulty level. Content analysis was conducted to evaluate the apps' availability, characteristics, and functionality.

Results: Evaluation of the apps revealed three functional categories: chatbot (15 apps), journal logging (9 apps), and psychotherapeutic treatment (3 apps). A majority (74.1%) employed natural language processing technology. The average MARS score of 3.45 out of 10 demonstrated a need for quality improvement. Low readability (averaged 6.04 for the content and 9.44 for the introduction in the app store) and monotonous user interface implied inadequate child-friendly design. While 74.1% of apps used clinically validated technologies, rigorous clinical tests of the apps were often missing, with only one app undergoing a clinical trial. Concerns over response accuracy were also identified. High costs (74.1% required payment with a mean of \$20.16/month) could also limit accessibility.

Conclusions: This study highlighted an urgent need for improved quality, child-friendly design, rigorous clinical testing, and affordable pricing to maximize the benefits of these AI-based mental health apps for children.

(JMIR Preprints 28/03/2024:58597)

DOI: <https://doi.org/10.2196/preprints.58597>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to the public.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>, I will be able to make my accepted manuscript PDF available to anyone.

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Artificial Intelligence-Based Mobile Phone Applications for Child Mental Health: A **Comprehensive** Review and Content Analysis

Preprint
JMIR Publications

Abstract

Background: Mobile phone apps powered by Artificial Intelligence (AI) have emerged as powerful tools to address mental health challenges faced by children. This study aimed to comprehensively review AI-driven apps for child mental health, focusing on their availability, quality, readability, characteristics, and functions.

Methods: This study systematically analyzed AI-based mobile apps for child mental health. Quality was evaluated using the Mobile Application Rating Scale (MARS), which assessed various dimensions of app quality, including subjective quality, engagement, functionality, aesthetics, and information. An Automatic Readability Index calculator was implemented to assess readability by utilizing the count of words, syllables, and sentences to generate a score indicative of the reading difficulty level. Content analysis was conducted to examine the apps' availability, characteristics, and functionality.

Results: Out of 369 apps initially identified, 27 met the eligibility criteria for inclusion. The quality of the apps was assessed using MARS, with an average score of 3.47 out of 5, indicating a need for quality improvement. The readability analysis revealed suboptimal scores, with an average grade level of 6.62 for in-app content and 9.93 for app store descriptions. These results, combined with a monotonous user interface, suggest that many apps lack a child-friendly design, potentially hindering their usability and engagement for young users. Content analysis categorized the apps into three functional groups: chatbot-based apps (15 apps), journal logging apps (9 apps), and psychotherapeutic treatment apps (3 apps). While 20 out of 27 apps (74.1%) used clinically validated technologies, rigorous clinical tests of the apps were often missing, with only two apps undergoing clinical trials. Of the 27 apps analyzed, only 7 (25.9%) were free to use, while the majority, 20 apps, required a subscription or one-time payment. Among the paid apps, the average cost was \$20.16 per month, which may pose a financial barrier and limit accessibility for some users, particularly those from lower-income households.

Conclusions: AI-based mental health apps hold significant potential to address the unique challenges of child mental health but face critical limitations in design, accessibility, and validation. To fully realize their benefits, future research and development should focus on integrating child-centric design principles, ensuring affordability, and prioritizing rigorous clinical testing. These efforts are essential to harness the power of AI technologies in creating equitable, effective, and engaging solutions for improving child mental health outcomes.

Keywords: Mental Health; Children; Artificial Intelligence; Mobile Health; AI-driven Mobile Applications

Introduction

Mental health, as defined by the World Health Organization, is a state of well-being where individuals effectively manage stress, recognize their abilities, work productively, and contribute to their community. It includes conditions such as mental disorders, psychosocial disabilities, and states involving significant distress or functional impairment [1]. Concerns regarding mental health, particularly among children—defined as individuals under the age of 18 years [2]—have garnered increasing attention in recent years. Around the world, 1 in 7 children experience mental health conditions, yet many more remain largely unrecognized and untreated [1]. Children with mental health challenges face various risks, including social exclusion, discrimination, stigma, educational difficulties, risk-taking behaviors, physical health symptoms, and human rights violations [3]. To safeguard their lifelong health and well-being, it is crucial to shield children from adversity, promote psychological wellness, and ensure their accessibility to mental health care [4,5].

Many children grapple with mental health issues, and an overwhelming fraction encounter barriers to accessing effective, affordable, and high-quality mental health services [6]. Several factors contribute to this limited access, including lack of health insurance, inadequate financial support, cultural constraints, and extended waiting times due to a shortage of professional services [4]. Beyond these micro-level challenges, public health crises, notably the COVID-19 pandemic, have exacerbated disparities between the escalating mental health needs of children and the availability of quality care. During the pandemic, national surveys across various countries revealed a heightened mental health strain on children, exacerbated by lockdowns that limited access to services and reduced help-seeking out of fear of contracting virus [1, 7, 8].

The swift advancement of technology, particularly in communication, has paved the way for novel mental health support mechanisms. The widespread availability of mobile phones and associated apps makes it possible for individuals to easily tap into essential resources, irrespective of where they are [9]. This democratization of access has been particularly impactful for mental health services, which are now available at the fingertips of those in need, often at a fraction of the traditional cost [10]. The integration of AI in these apps has further revolutionized the landscape. As noted by [11], the infusion of AI technology not only enhances the precision of these platforms but also imbues them with a more empathetic and human touch. This allows for a tailored intervention, optimizing the support offered based on an objective analysis of user data.

Growing up in the digital era, children are inherently familiar with digital devices such as smartphones and tablets [12, 13]. These tools offer an intuitive and user-friendly interface, eliminating complexities or abstract concepts and reducing potential barriers for younger users [13]. Previous research highlights a strong preference among children and adolescents for AI-integrated tools [13]. Studies suggest that AI-informed mobile health apps hold significant potential in addressing mental health issues in children [14]. These apps are generally well-received, with children showing positive attitudes toward their integration. Experts view such tools as complementary to traditional in-person interventions, offering low-threshold access via smartphones and extending their reach to children in high-risk situations or underserved communities [14]. This trend underscores the readiness of these age groups to embrace and engage with emerging technologies.

While AI-powered mental health apps are increasingly prevalent in the market, concerns have been raised about their effectiveness and quality[15,16]. These concerns are heightened by the rapid pace of AI advancement and evolving mental health care guidelines. It is important to note that most of the current scientific literature concentrates on adult mental health, often overlooking the unique needs of children. Recent comprehensive reviews by Li et al. [17] and Hua et al. [18] have made significant contributions to the field by providing valuable insights into AI-based conversational agents and large language models, respectively. Our study builds upon this foundation by encompassing a diverse array of AI-driven mobile phone apps, both those exclusively designed for children and those that include children as part of their broader target audience. This includes not only conversational agents, but also AI techniques used for emotion and journal tracking, as well as psychotherapeutic treatment within mental health apps. Furthermore, while these previous works have addressed AI applications in a general population context, our research uniquely targets the mental health of children, a demographic with specific needs and challenges that are often underrepresented in the literature. By systematically reviewing AI-based mental health apps accessible to children, our study aims to highlight the current gaps and challenges in the development of child-specific resources within this emerging field. We evaluated five core areas: availability, quality, readability, characteristics, and functionality. Our insights aim to lay the groundwork for future research and guide the creation and development of AI-driven mental health apps that better cater to the needs of children, ultimately fostering the long-term psychological well-being of our future generations.

Method

This study aimed to systematically review AI-driven mobile apps designed for children's mental health. The review process involved multiple steps, including a thorough literature search, screening, and evaluating selected apps.

App Search and Screening Strategy

An extensive search was conducted in the two most widely used marketplaces for mobile phone apps: Android's Google Play Store and Apple's IOS App Store. Search terms included "AI mental health apps for child/ children" and "AI mental health apps for teenagers." These apps were subsequently screened based on the following criteria: (1) available in English; (2) can be downloaded for current use from Google Play or the Apple App Store; (3) the primary function is improving mental health; (4) the app must explicitly state that it is AI-based, with the AI component being central to its functionality in addressing mental health issues; (5) the minimum age of the user shall be less than 18 years old.

The minimum age for app users in app stores is determined by a combination of factors, including developer-provided information, content ratings, and compliance with legal regulations. Developers are required to disclose the intended audience and the highest age-rated content available in their apps, which is then aligned with standardized rating systems like those developed by the International Age Rating Coalition. These systems ensure that apps are categorized appropriately based on their content, features, and potential risks. Legal frameworks such as the Children's Online Privacy Protection Act in the U.S. and the General Data Protection Regulation in the European Union further influence these age ratings by setting strict guidelines for data collection from children, thereby shaping the minimum age requirements [19]. Using the app stores' minimum age ratings as a criterion for this study ensured that the reviewed apps include both those explicitly designed for children and those that, while not exclusively targeting children, include children as part

of their broader user population. This dual inclusion provides a comprehensive analysis of AI-based mental health support tools available to children under 18 years old.

Apps that did not meet these criteria were filtered out. Two research investigators independently conducted this screening process and reached a consensus on the final selection.

App Evaluation

Two investigators independently evaluated the apps. The Mobile Application Rating Scale (MARS) was primarily used to assess the quality of the apps, while a readability calculator was employed to evaluate the readability. Content analysis was conducted to determine the descriptive characteristics and functionality. Descriptive characteristics codes included: (a) price (free versus for-purchase); (b) confidentiality, concerning whether the app communicated and managed policies related to user privacy and data security; and (c) AI techniques adopted, referring to specific AI techniques employed by the apps. Mobile app functionality components were also evaluated, including (a) clinically-proven psychotherapeutic treatment, detailing whether the app's treatments or the app itself have been clinically validated; (b) target mental health problems about specific mental health issues the app was designed to address; (c) functions. Any disagreements about an app were resolved through discussion until a consensus was reached.

Readability of Mobile Apps

The readability of the mobile apps was assessed by two independent researchers using an Automatic Readability Index calculator. Considering the prevalence of children's apps being downloaded by parents but primarily utilized by the children themselves [20], our analysis strategically separated the readability assessment into two distinct parts: examining the descriptions listed in the app store and the in-app text. For the in-app text, all visible text on the application pages during typical use was included in the analysis. If a page contained minimal text insufficient for readability evaluation, further interaction with the app was conducted to generate additional text output, such as engaging with chat features. Unrelated content, such as advertisements, was excluded to focus on text that directly impacts user interaction and comprehension. The Flesch-Kincaid Grade Level, an integral component within the readability calculator, was selected as our standard for readability, given its widespread recognition within the field [21-23]. The readability score was calculated using the formula below. The index corresponds to the U.S. elementary school grade levels associated with the evaluated text, offering a universally comprehensible measure of text readability. A higher value on the readability index denotes a higher grade level, implying that more advanced reading proficiency is required.

$$\text{Readability} = 0.39 * \left(\frac{\text{total number of words}}{\text{total number of sentences}} \right) + \frac{11.8 * \text{total number of syllables}}{\text{total number of words}} - 15.59$$

Mobile App Rating Scale Assessment

The quality of the apps was evaluated using MARS, a widely recognized standardized quality measurement tool for health-related mobile apps [24–27]. MARS was characterized by two principal dimensions: the subjective quality scale, which imparted valuable insights into user experiences and formed a crucial part of the comprehensive evaluation index, and the objective scale, which provided a rigorous, structured evaluation of the apps' various technical and functional attributes. The objective scale comprised 19 items rated on a 5-point scale (1= inadequate, 2= poor, 3= acceptable, 4= good, 5= excellent). The items within the objective scale were further divided into four thematic categories: *Engagement*, *Functionality*, *Aesthetics*, and *Information*. Each category assessed diverse elements, from user interactivity and design appeal to the accuracy and credibility of the information provided. The final MARS score was the mean of the four sub-scores in these dimensions. This structured approach accomplished a comprehensive and detailed evaluation of the app's quality.

Content Analysis

In addition to evaluating the apps by the readability index calculator and the MARS scale, we also used content analysis to assess the included apps. Content analysis encompassed three main areas: availability, the descriptive characteristics of the apps, and functionality. Availability evaluated the presence of the apps across various app stores. Descriptive characteristics codes included: (a) price (free versus for-purchase); (b) confidentiality, concerning whether the app communicated and managed policies related to user privacy and data security; (c) AI techniques adopted, referring to specific AI techniques employed by the apps; and (d) children- Specificity. Mobile app functionality components were also recorded, including (a) clinically-proven psychotherapeutic treatment, detailing whether the app's treatments or the app itself have been clinically validated; (b) target mental health problems about specific mental health issues the apps were designed to address; and (c) functions.

Two coders (JW and XZ) underwent training to attain reliability across several rounds of sample mental health apps designed for children. They were required to achieve an interrater reliability criterion of $K = 0.80$ on ten independently coded apps before commencing coding on the included apps in this study [28]. The 27 apps were downloaded from their respective app stores, and each was assessed and coded by these two coders. An overall interrater reliability of 0.94 was achieved, indicating strong reliability [29].

Results

As of December 2024, this search yielded 369 apps. After removing 42 duplicated apps, 327 apps remained. The 327 apps were evaluated based on the inclusion criteria. This selection process resulted in 27 apps being included in this review study, with 4 apps specifically designed for children and 23 apps included children as part of their broad user base. The detailed selection process was demonstrated in Figure 1.

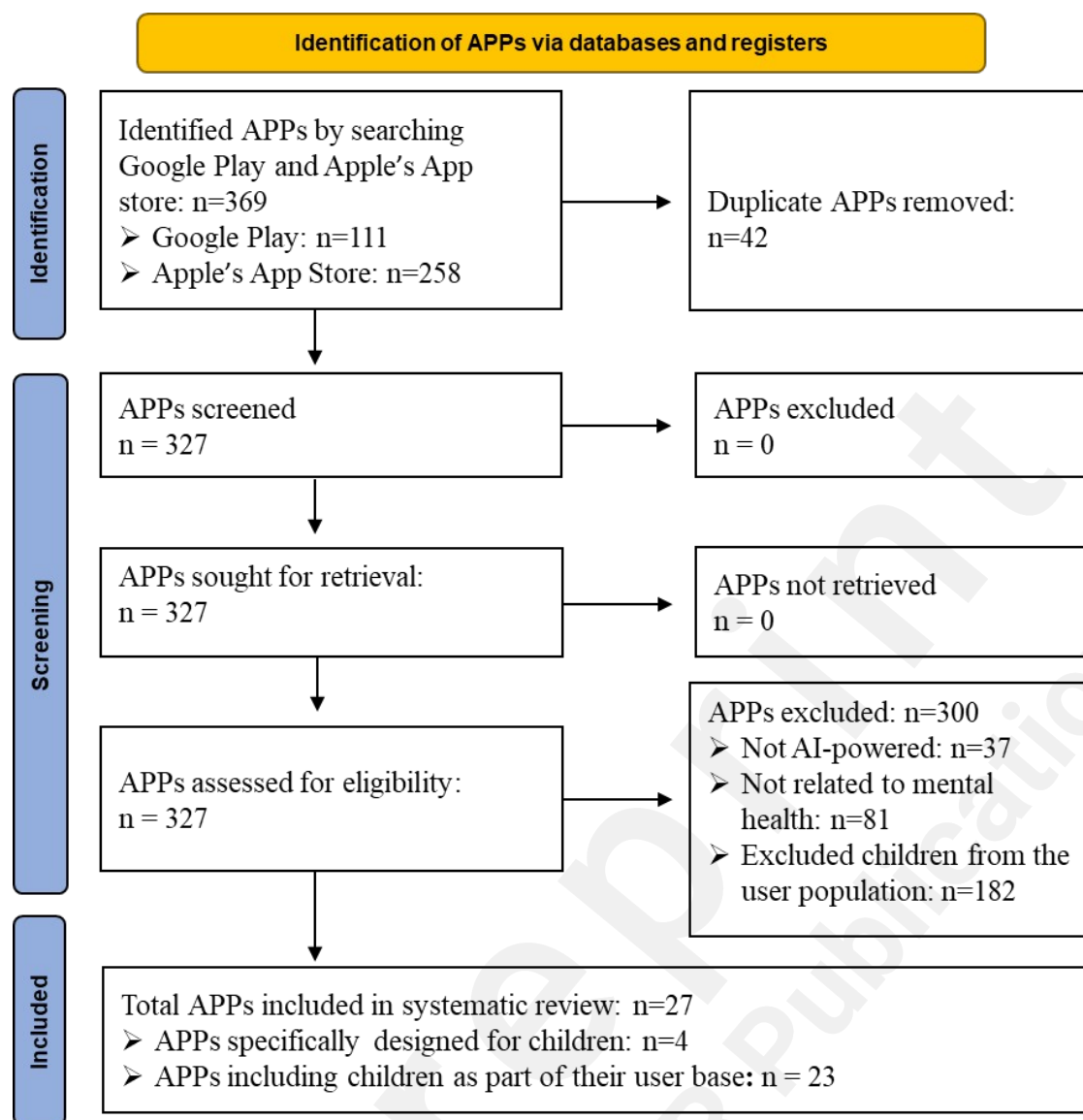


Figure 1. PRISMA of APP Search and Screening Strategy.

Readability

Our study utilized the Flesch-Kincaid Grade Level index in a readability calculator (<https://datayze.com/readability-analyzer>) to evaluate the readability of the app descriptions in the app store and in-app contents. The test revealed that the content readability within the apps spanned from grades 3 to 11, with an average readability of grade 6.62. This suggested a certain threshold of readability for younger children. The readability of the app description in the app store had a mean grade level of 9.93, which was notably higher than the content's readability. This could be due to the anticipation of parental involvement during the app download and installation processes, as parents are generally expected to have higher reading proficiency and are thus more likely to comprehend such introductions.

Quality Evaluation

Our findings revealed a suboptimal overall quality among existing apps, with MARS scores varying from 2.58 to 4.49, with an average of 3.47 on a scale of 1 to 5. Figure 2 employed a stacked bar graph to delineate the sub-scores and the composite MARS scores across these four dimensions. Of

particular concern was the *Information* quality, which had the lowest mean score of 3.19 and a minimum of 1.71. In contrast, *Engagement* scored the highest, with a mean of 3.56 and a score of 2.6-4.8.

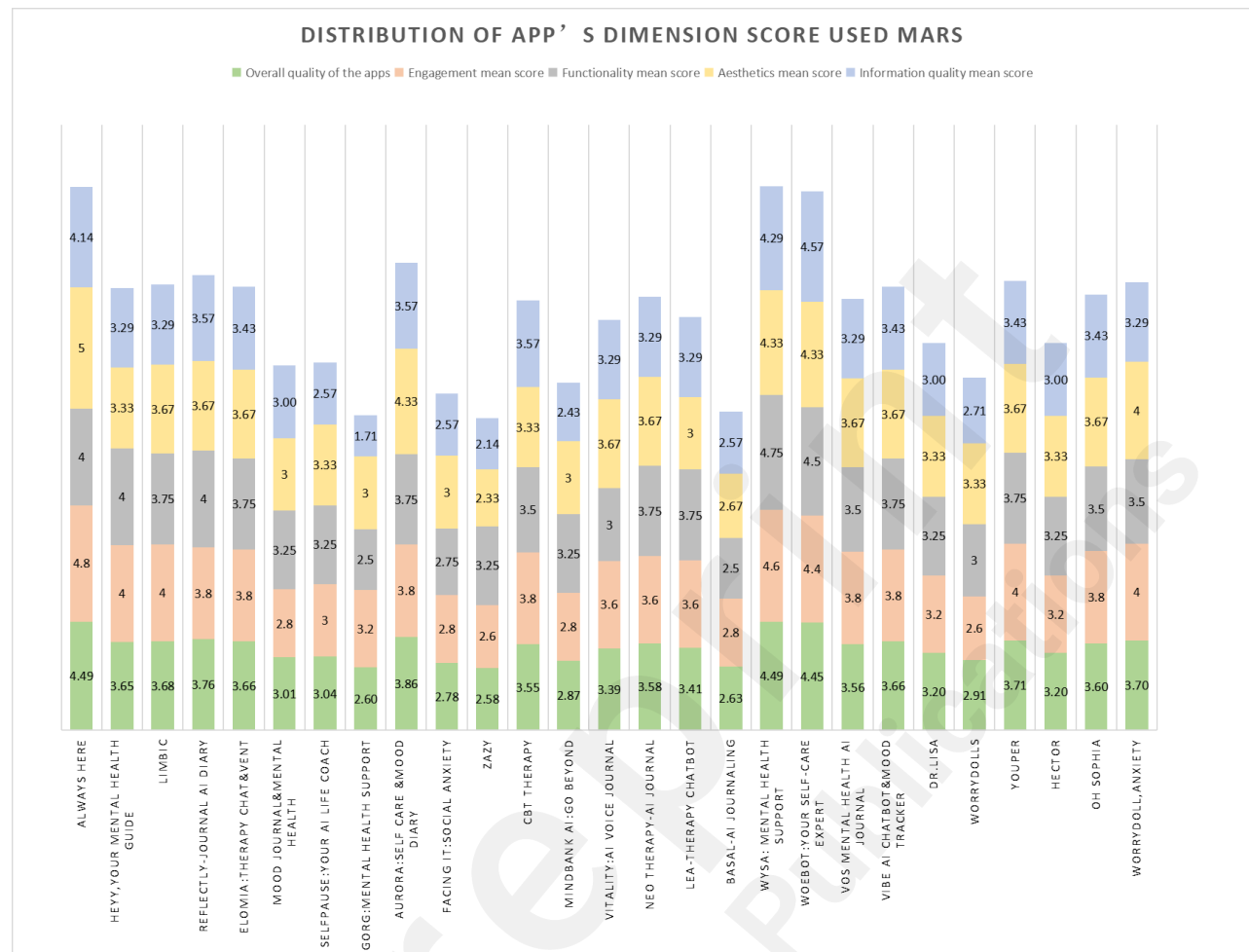


Figure 2. Total MARS Scores and Sub-scores across Four Dimensions.

Engagement subscale emerged as the highest-scoring category in our study. This subscale was assessed through five perspectives: *customization*, *interactivity*, *entertainment*, *interest*, and *target group*. Delving deeper into the results, 44.4% of the apps incorporated a baseline log-in assessment, emphasizing user *customization*, while 55.6% did not include this feature. The apps displayed varied engagement strategies: 55.6% primarily utilized dialogue interactions, and 33.3% focused on diary log inputs. However, many apps needed more rigorous tracking and oversight mechanisms, suggesting a gap in ensuring sustained user interactivity. Additionally, the apps exhibited diverse levels of *entertainment* and *interest*, with scores ranging from 1 to 5. These variances, especially in entertainment and interest, could affect the target group's appropriate level.

Four primary characteristics, *performance*, *ease of use*, *navigation*, and *gestural design*, were utilized to assess *Functionality*. In terms of *performance*, seven (26%) apps demonstrated low accuracy or speed in their features and components, such as buttons and menus, adversely affecting the user's experience. *Ease of use* was apparent in 24 apps (88.9%), as they facilitated effortless use for users with distinct menu labels and instructions. When evaluating *navigation*, these apps generally exhibited a coherent and uninterrupted transition between screens, with essential links in

place. The scale score of this component ranged from 3 to 5, with an average score of 3.41, indicating that the apps' navigation is relatively well-integrated and logically structured. In the domain of *gestural design*, the apps consistently exhibited intuitive user interactions, such as tapping, swiping, pinching, and scrolling, across various components and screens. The evaluation scores for this aspect ranged from 2 to 5, with an average score of 3.48.

In evaluating the *Aesthetic* appeal of the apps, particularly given their primary target audience of children, our findings revealed some limitations in their overall design. Based on the *Aesthetic* subscale's criteria – which included *layout*, *graphics*, and *visual appeal* – only 11 (40.7%) of the reviewed apps showcased vibrant and engaging designs that effectively captivated the intended population. In stark contrast, 55.6% of the apps exhibited monotonous designs that, while functional, lacked the visual allure necessary to maintain the interest of a younger audience. This aspect was particularly evident in the *layout* and *graphics*, where the arrangement and quality of buttons, icons, menus, and content were often not *visually appealing*. Especially in the *visual appeal* part, the score range was 2-5, and the average score was only 3.26, indicating that the overall appeal of the app design was not strong. Moreover, 14.8% of these apps predominantly featured dark color schemes, which could potentially create negative psychological impressions and detract from their attractiveness to children. This mismatch between the apps' aesthetic elements and their intended users' preferences highlighted a significant area for improvement in app design.

In assessing the apps from an *Information* perspective, our findings demonstrated high *accuracy* in describing features, aligning well with the first characteristic of the *Information* subscale. Nearly all apps accurately reflected their related features in app stores, avoiding exaggeration or misinformation. Regarding the *goals*, the majority of the apps were explicit in stating specific, measurable, and achievable objectives. Regarding the quality of information, 68% of the apps, primarily developed by research labs and scientific institutions, provided correct, well-written content relevant to their goals. This underlined their credibility as sources of information. Regarding *quantity*, most apps offer comprehensive yet concise information within the scope of their intended use. Additionally, the *visual information* in these apps generally presented concepts clearly and logically through various means like charts, graphs, images, and videos.

Content Analysis

Content analysis systematically evaluated app availability, descriptive characteristics, and functionality. In terms of app availability, all apps were sourced from the Apple App Store and Google Play. There were 22 apps available in both app stores. Four apps were exclusively available in the Apple App Store, and one was solely offered in Google Play.

We have identified critical, descriptive characteristics of these apps in terms of cost, confidentiality, AI techniques, and children's specificity. There were 20 apps requiring payment to access additional content, with prices ranging from \$2.99 to \$120 per month (mean=\$20.16), while only seven apps offer basic functionality for free. All apps ensured privacy protection by including privacy policies in the App Store and Google Play. AI technologies include Natural Language Processing (NLP), AI recommendation systems, Automatic Speech Recognition, and data analysis and tracking. NLP was used in 74.1% of the apps, mainly for chatbot interactions, while Automatic Speech Recognition was adopted in one voice input app. Nine emotion and journal tracking apps employed data analysis to monitor and track users' mental health. Although all the apps have age limits that make them accessible to children, most (85.2%) were not explicitly designed for a young audience, often lacking content and interfaces tailored to be child-friendly. Only a small proportion (14.8%, or four apps) were explicitly tailored for children, featuring age-appropriate content and user interfaces.

Functionality components were evaluated from the perspectives of clinically proven psychotherapeutic treatment, targeting mental health problems, functions, and challenges. Despite the utilization of clinically validated psychotherapeutic treatments like Cognitive Behavioral Therapy and Dialectical Behavior Therapy in 74.1% of the apps, the overall effectiveness of the apps largely remained untested, with two apps, Woebot and Youper, having undergone clinical trials. Woebot underwent a clinical trial involving 36,070 participants, which revealed that users of the app experienced a significant reduction in depression and anxiety symptoms after 3-5 days of use [30]. Similarly, Youper has shown clinical evidence of benefit in a study involving its paying users. This study found that Youper was effective in reducing symptoms of anxiety and depression, with significant improvements observed [31]. Although this provides promising support for both Woebot and Youper, further rigorous testing, particularly through additional RCTs, is needed to establish their clinical effectiveness more definitively. In terms of targeted mental health problems, 22 apps focused on general mental health problems, while five apps were explicitly designed to treat anxiety and depression. Included apps can be divided into three major categories based on functions: dialogue interaction, journal log tracking, and psychotherapeutic interventions, which accounted for 55.6%, 33.3%, and 11.1%, respectively.

The utilization of these apps revealed several challenges, which can be categorized into three primary areas. Initially, a significant concern was the reliance on static large language models, which do not update their parameters after deployment and therefore may not align with rapid advancements in the field. This led to 12 apps needing help with complex or nuanced user inputs, resulting in ineffective or delayed processing. The responses of seven apps tended to be repetitive and needed help to deliver meaningful conversation after multiple trials of conversation or when topics extended beyond standardized themes. Secondly, nine apps exhibited a critical deficiency in user stickiness and feedback mechanisms. User stickiness was evaluated by examining factors including the app's ability to re-engage users after periods of inactivity, and the presence of features that incentivize continuous interaction, such as personalized content, reminders, and progress tracking. The lack of robust tracking and feedback systems may precipitate user attrition, diminishing the apps' potential advantages. The third and equally significant challenge was observed in four apps, particularly concerning their approach to providing clear guidance. One app showed inconsistent dialogue transitions, where choices made by users were not appropriately reflected in the following responses, leading to confusion.

Additionally, the game elements within one app needed explicit instructions, creating usability challenges for young users. Explanations or instructions were absent from the results displayed on the card surface, making it difficult for children to comprehend. Collectively, these challenges resulted in a less intuitive and supportive user experience, highlighting the need for more dynamic and user-centric design in these apps.

Table 1: Content Analysis for Availability, Descriptive Characteristics, and Functionality

Availability	Subcategory	Value
Availability	Google Play	1 (3.7%)
Availability	Apple's App Store	4 (14.8%)
Availability	Both Google and Apple	22 (81.5%)
Descriptive Characteristics	Subcategory	Value
Cost	Free	7 (25.9%)

Cost	Purchased	20(74.1%)
Confidentiality	Y	27(100%)
Confidentiality	N	0(0%)
AI Techniques	Natural Language Processing	20 (74.1%)
AI Techniques	AI-based Recommendation System	2(7.4%)
AI Techniques	Automatic Speech Recognition	1(3.7%)
AI Techniques	Data Analysis and Tracking	7(25.9%)
Children Specificity	Y	4(14.8%)
Children Specificity	N	23(85.1%)
Functionality	Subcategory	Value
Clinical-proven	Y	25(92.6%)
Clinical-proven	N	2(7.4%)
Psychotherapeutic treatment	None	7(25.9%)
Psychotherapeutic treatment	Cognitive Behavioral Therapy	19(70.4%)
Psychotherapeutic treatment	Dialectical Behavior Therapy	4(14.8%)
Psychotherapeutic treatment	Acceptance and Commitment Therapy	1(3.7%)
Psychotherapeutic treatment	Exposure Therapy	1(3.7%)
Target Problems	General	21(77.8%)
Target Problems	Anxiety, Depression, and Stress	6 (22.2%)
Functions	Dialogue Interaction	15(55.6%)
Functions	Emotion and Journal Tracking	9(33.3%)
Functions	Psychotherapeutic Treatment	3(11.1%)

Discussion

Summary of findings

In an era characterized by the rapid progression of AI, an increasing interest in leveraging technological advancements to address population-level mental healthcare needs has emerged [28]. In the context of growing societal attention to child mental health challenges [4], our research assessed the quality, readability, characteristics, and limitations of AI-driven mental health apps oriented to children. We identified three main functional categories: chatbots (15 apps), journal logging (9 apps), and psychotherapeutic treatment (3 apps), with 74.1% employing natural language processing technology. The average MARS score of 3.47 out of 5 indicated a need for quality improvement. Readability was not sufficiently child-friendly. While most apps (92.6%) used clinically validated technologies, rigorous clinical testing was lacking, with only two apps having undergone clinical trials. Additionally, concerns about response accuracy and the high costs associated with these apps—74.1% required payment, averaging \$20.16 per month—could limit their accessibility and effectiveness for children.

Underrepresentation of Child-Specific Design

Empirical research has underscored a positive inclination among children and adolescents toward AI and its integration with interventions, suggesting a solid willingness within these age groups to explore these technologies [13]. However, the evident lack of child-centric design may undermine these potential advantages, as identified in our investigation.

In the context of readability, these apps primarily required advanced reading skills. Previous studies have shown that parents are often the gatekeepers of the app download and installation process. At the same time, children are the users [32]; we evaluated the readability of app descriptions in the app store and content separately. Our findings correspond with the broader readability challenges of

mobile apps designed for vulnerable populations [24]. The readability of user interface content—consistently 2-4 grade levels lower than the app descriptions—still presented a considerable readability threshold for younger users, raising the accessibility barrier for app usage and potentially hindering independent use by children [4].

The content incorporated in these apps was not explicitly tailored to children. For example, our findings revealed that certain apps primarily focused on alleviating occupational stress and addressing professional interpersonal conflict, thereby falling short of addressing the unique mental health requirements of children. This mismatch can be attributed to two key factors. First, due to the limited availability of apps exclusively designed for children's mental health, our study included apps that, while not specifically targeted at children, were accessible to them within the broader age range intended by the developers. As a result, some apps that children could access contained content not relevant to their specific needs. Second, this mismatch also reflects the developers' limited understanding of children's mental health needs, leading to the inclusion of content that does not align with children's requirements. App developers were predominantly technology firms and research laboratories without professional child mental health experts to inform app design. Hence, based on the recommendations made by prior studies [13], we advocated for a multidisciplinary team, comprised of AI and app technicians alongside child mental health professionals, to guide the future development of AI-driven apps designed to address children's mental health concerns. This approach should involve collaboration between AI and app developers, child mental health professionals, and endorsement by peak mental health organizations. The involvement of these key stakeholders ensures that the apps are clinically effective, relevant to the target population, and endorsed by authoritative bodies, thereby increasing their credibility and acceptance in clinical practice. Furthermore, the co-creation of these apps with children and clinicians can enhance their usability and ensure that they meet the specific needs of the target population.

The intervention delivery method within these apps also demonstrated a need for more child-friendly techniques. Diary-logging apps and chatbot dialogues were the dominating intervention strategies employed by these apps. However, diary logs and chatbot interactions needed substantial text input and the ability for users to clearly articulate their mental health concerns, an approach that has been evidenced as being more challenging for child users [4,28]. The expressive capabilities of children may need to be sufficiently developed to leverage these apps, thereby impacting their effectiveness. Incorporating child-friendly tools such as art therapy [4], an area made achievable by recent developments in text-to-image models like stable diffusion models and DALL·E, could offer potential solutions.

In addition to readability, app content, and intervention delivery method, the existing apps' visual aesthetics and user interface design also demonstrated a need for child-centric considerations. The color palette of many current apps could have been more varied, simplistic, and potentially unengaging, lacking vibrant imagery that could capture children's attention [33]. On the contrary, some apps adopted a dark aesthetic, which could appear somber and oppressive, potentially intensifying feelings of discomfort or even burden, particularly in users already experiencing low mood or mental health issues[34]. Hence, it was advisable for app designs to integrate richer, brighter colors and more engaging images. Including popular animation characters could enhance children's interest and engagement [4].

Barriers to Optimal Mental Health Outcome Concerns

To effectively embed the findings of this study into clinical practice, it is crucial to integrate AI-driven mental health apps into existing care frameworks. This integration can enhance the accessibility and quality of mental health care for children, particularly in underserved populations.

However, for these tools to be successfully adopted in clinical settings, their effectiveness in producing measurable behavior change and improving mental health outcomes must first be validated. While our review highlights the potential of AI-driven mobile phone applications in addressing child mental health, it is important to note that the effectiveness of these interventions in producing measurable behavior change remains largely untested. Therefore, future research should focus on evaluating the impact of these apps on users' mental health outcomes and behavioral modifications. Establishing this evidence is essential to ensure that these apps can be confidently and effectively incorporated into routine care, ultimately delivering the intended benefits to the target population.

The effectiveness of an app was fundamentally linked to the therapeutic technology it employed [28]. Most reviewed mobile apps utilized clinically validated technologies, such as CBT and DBT Therapy [35]. The effectiveness of an app was fundamentally linked to the diagnostic and therapeutic technology it employed [32]. However, these apps' effectiveness in enhancing child mental health largely remained clinically unproven, a conclusion drawn from our review of 27 such apps. Even though these apps primarily serve adjunctive support tools and cannot fully replace professional psychological diagnosis and treatment [11], we strongly recommend that these apps undergo rigorous clinical validation to ensure their effectiveness, underpinned by solid evidence.

User engagement and stickiness are critical factors that can influence the effectiveness of mental health interventions. The journey toward mental health improvement is cumulative, requiring consistent commitment over time [11]. This underscores the importance of regular app usage and the fostering of robust user engagement for optimal intervention functionality [36]. However, a significant shortcoming in the design of most current apps is the absence of effective user tracking mechanisms. User tracking mechanisms were examined by examining the presence of features such as usage analytics, personalized reminders, and progress tracking, which are essential for promoting regular interaction. These features are known to encourage regular use and engagement, which are important for the success of any intervention [37–39]. Several apps have implemented systems requiring notification activation to promote more regular user interaction [11], contributing to a more effective mental health intervention experience.

Furthermore, the interplay between AI-powered mental health apps and user privacy presented unique strengths and challenges that can indirectly affect mental health outcomes by altering the users' decisions on whether to engage with these apps. The stigma attached to mental health issues in various cultural and societal contexts heightens privacy concerns [32], especially among underrepresented populations with limited access to formal mental health care [40]. AI-powered mobile phone apps can positively enhance the sense of privacy credited to no human involvement [41]. However, chatbot dialogues and journal logging may involve significant exposure of personal and traumatic experiences, raising critical privacy concerns about handling such sensitive data [10]. This potential exposure may deter many privacy-sensitive individuals from using these apps. To alleviate these concerns, many apps have adopted privacy safeguards such as presenting privacy policies during initial installation, anonymizing data, and post-session deletion of user histories. Transparency in implementing and communicating these privacy-related strategies remained crucial.

The association between low socioeconomic status and adverse child mental health outcomes has been widely described [8,42–44]. On the flip side, cost often presents a significant obstacle in accessing mental health services, with affordability issues preventing nearly half of the individuals requiring mental health treatment in the United States from receiving it [45]. In this context, the presumed high accessibility of AI-powered mental health apps, both in academia and public perception, was considered a significant advantage, fostering wider user reach and coverage [33].

However, most apps we evaluated incorporate fee-based features, raising the barrier to use. This increased cost compromises the perceived accessibility of the AI and mobile app approach, potentially curtailing its effectiveness.

Challenges and Considerations in AI Techniques for Child-Oriented Mental Health Apps

The current landscape of AI in child-oriented mental health apps reveals a significant reliance on NLP to assess and intervene in users' mental health states. While NLP is a powerful tool for processing and analyzing language, this predominant focus on text-based interactions may not fully address the needs of children, who often struggle with limited reading, verbal, and writing skills, particularly when expressing complex mental health concerns. The lack of diverse interaction methods could hinder children's ability to engage with and benefit from these mental health interventions. In light of this, we suggest future research and apps incorporating child-friendly AI techniques, such as generative AI for image creation using natural language prompts, to complement NLP-based approaches. Image-based tools can help children express their emotions and thoughts more effectively, broadening the accessibility and impact of mental health apps for this demographic. By offering multiple modes of interaction, these apps could better accommodate the diverse needs and abilities of children, ultimately making them more effective tools for promoting mental health among younger users [3].

Consistent with prior research [10,11,41,46], our findings revealed that the quality of the AI techniques currently in use raised considerable concerns. The performance of NLP, in particular, demonstrated notable inconsistencies and disparities among different mobile phone apps. For instance, some apps needed help processing simple questions, often responding with repetitive and uninspiring answers after prolonged periods. This issue was especially evident when voice input—an interaction mode that naturally resonates with and is easily used by children [5]—was utilized. These apps' inability to handle such a fundamental form of interaction effectively raised significant questions about their usability and overall quality [47]. Given these limitations, users might be deterred from relying on these tools for consistent support, limiting their potential benefits.

In light of these quality issues, we suggest that the application of implementation science principles is vital for the practical deployment of AI-driven mental health apps. One approach is for developers to prioritize incorporating state-of-the-art language models. One of the key challenges in this context is preventing AI hallucinations—where AI generates inaccurate or misleading information. To mitigate this risk, we propose linking AI systems within these apps to established clinical guidelines using advanced information retrieval methods. This approach ensures that AI provides evidence-based recommendations, thereby reducing the likelihood of errors and ensuring that the advice given is both accurate and appropriate for children. Another approach involves the use of rule-based models, as seen in apps like Woebot, which prioritize safety by avoiding generative AI for user-facing text. Even with rule-based models, there is room for improvement by updating and refining the model parameters. This approach can enhance the user experience while maintaining the safety and reliability of the app. By integrating these strategies, AI techniques in AI-driven mental health apps for children can be better positioned to meet clinical standards and effectively support children's mental health needs.

Limitation

It is important to acknowledge the limitations of this study. The first limitation lies in the time-sensitive nature of the app search process. AI-based mental health apps are being developed and launched at an unprecedented rate, reflecting the rapid advancements in AI technology. To mitigate

this limitation, we conducted two app searches in October 2023 and December 2023. However, this rapid pace of development means that some newly launched apps may not have been captured in our analysis. Future studies could address this limitation by employing longitudinal designs with repeated app searches over extended periods to better capture the evolving landscape of AI-based mental health apps. The second limitation concerns the scope of the apps included. Currently, there are very few AI-based mental health apps specifically designed for children available on the market. As a result, we included all AI-based mental health mobile phone apps including both those explicitly designed for children and those that, while not exclusively targeting children, still consider them part of their intended user base. As AI-based mental health interventions specifically designed for children continue to grow, future research should focus on systematically reviewing apps developed exclusively for children to provide a more targeted analysis. Thirdly, we only included apps available in English, which may have excluded apps offered in other languages. This limitation highlights the need for future studies to explore apps in diverse languages to provide a more inclusive understanding of the global availability and accessibility of AI-based mental health apps for children.

Conclusion and Implication

AI-powered mobile apps were positioned at the frontier addressing children's mental health issues, uniquely combining accessibility and acceptance among the younger demographics. Our study, however, emphasized that the future of these AI-based tools was intrinsically tied to their capacity for significant evolution and enhancement. This included enhanced child-friendly design, clinically impactful outcomes, and superior quality of AI techniques, all of which should be considered non-negotiable facets of their development. At this juncture, our choices will undeniably carve out the path for AI-powered mental health support for children. It is our collective responsibility as a profession to ensure these digital solutions extend beyond technological sophistication, embodying empathy, effectiveness, and genuine child-centricity. By doing so, we can unlock the full potential of AI, creating a future where every child, regardless of their location or circumstances, has access to robust, reliable, and adequate mental health support.

References

- [1] World Health Organization. Improving the mental and brain health of children and adolescents. Available at: <https://www.who.int/activities/improving-the-mental-and-brain-health-of-children-and-adolescents>. Accessed September 3, 2023.
- [2] Lansdown G, Vaghri Z. Article 1: Definition of a Child. In: Vaghri Z, Zermatten J, Lansdown G, Ruggiero R, editors. *Monitoring state compliance with the UN Convention on the Rights of the Child. Children's Well-Being: Indicators and Research*. Vol. 25. Cham: Springer; 2022. Available from: https://doi.org/10.1007/978-3-030-84647-3_40.
- [3] World Health Organization. Mental health of adolescents. Available at: <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health>. Accessed September 3, 2023.
- [4] Zhou X, Edirippulige S, Bai X, et al. Are online mental health interventions for youth effective? A systematic review. *J Telemed Telecare* 2021;27:638–66.
- [5] Guo Y, Hopson LM, Yang F. Socio-ecological Factors Associated with Adolescents' Psychological Well-being: A multilevel analysis. *Int J Sch Soc Work* 2018;3.
- [6] Patrick SW, Henkhaus LE, Zickafoose JS, et al. Well-being of Parents and Children During the COVID-19 Pandemic: A National Survey. *Pediatrics* 2020;146:e2020016824.
- [7] Jones Harden B, Slopen N. Inequitable experiences and outcomes in young children: Addressing racial and social-economic disparities in physical and mental health. *Annu Rev Dev Psychol* 2022;4:133–59.
- [8] Gadermann AC, Thomson KC, Richardson CG, et al. Examining the impacts of the COVID-19 pandemic on family mental health in Canada: findings from a national cross-sectional study. *BMJ Open* 2021;11:e042871.
- [9] Welch V, Wy TJ, Ligezka A, et al. Use of mobile and wearable artificial intelligence in child and adolescent psychiatry: scoping review. *J Med Internet Res* 2022;24:e33560.
- [10] Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry* 2021;20:318–35.
- [11] Boucher EM, Harake NR, Ward HE, et al. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Rev Med Devices* 2021;18:37–49.
- [12] Linder LA, Newman AR, Stegenga K, et al. Feasibility and acceptability of a game-based symptom-reporting app for children with cancer: perspectives of children and parents. *Support Care Cancer* 2021;29:301–10.
- [13] Tan K, Chen XS. Keeping Up with Technology: Socioemotional and Equity Challenges with Children and Schools. *Child Sch* 2023;45:127–30.
- [14] Götzl C, Hiller S, Rauschenberg C, et al. Artificial intelligence-informed mobile mental health apps for young people: a mixed-methods approach on users' and stakeholders' perspectives. *Child Adolesc Psychiatry Ment Health* 2022;16:86. Available from: <https://doi.org/10.1186/s13034-022-00522-6>.
- [15] Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. *Heal Expect* 2021;24:1072–124.
- [16] Hamdoun S, Monteleone R, Bookman T, et al. AI-based and digital mental health apps: Balancing need and risk. *IEEE Technol Soc Mag* 2023;42:25–36.
- [17] Li H, Zhang R, Lee Y-C, et al. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digit Med* 2023;6:236.
- [18] Hua Y, Liu F, Yang K, et al. Large language models in mental health care: a scoping review. *ArXiv Prepr ArXiv240102984* 2024.
- [19] App Store Connect Help. Manage app information: Set an app age rating. [Accessed 2024-12-

- 18]. Available from: <https://developer.apple.com/help/app-store-connect/manage-app-information/set-an-app-age-rating/>
- [20] Pearson HA, Montazami A, Dubé AK. Why this app: Can a video-based intervention help parents identify quality educational apps? *Br J Educ Technol* 2023;54:712–33.
- [21] Fitzsimmons PR, Michael BD, Hulley JL, et al. A readability assessment of online Parkinson's disease information. *J R Coll Physicians Edinb* 2010;40:292–6.
- [22] Munsour EE, Awaisu A, Hassali MAA, et al. Readability and comprehensibility of patient information leaflets for antidiabetic medications in Qatar. *J Pharm Technol* 2017;33:128–36.
- [23] Guo Y, Yang F, Hu F, et al. Existing mobile phone apps for self-care management of people with Alzheimer disease and related dementias: Systematic analysis. *JMIR Aging* 2020;3:1–10.
- [24] Guo Y, Yang F, Hu F, et al. Existing Mobile Phone Apps for Self-Care Management of People With Alzheimer Disease and Related Dementias: Systematic Analysis. *JMIR Aging* 2020;3(1):E15290 <https://aging.jmir.org/2020/1/E15290> 2020;3:e15290.
- [25] Hensley CP, Witte MSM, Cai J, et al. Assessment of mobile health applications for management of knee and/or hip osteoarthritis using the mobile application rating scale. *JCR J Clin Rheumatol* 2023;29:245–53.
- [26] Musgrave LM, Kizirian N V, Homer CSE, et al. Mobile phone apps in Australia for improving pregnancy outcomes: systematic search on app stores. *JMIR MHealth UHealth* 2020;8:e22340.
- [27] Stoyanov SR, Hides L, Kavanagh DJ, et al. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. *JMIR MHealth UHealth* 2015;3:e27.
- [28] Bry LJ, Chou T, Miguel E, et al. Consumer smartphone apps marketed for child and adolescent anxiety: a systematic review and content analysis. *Behav Ther* 2018;49:249–61.
- [29] Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378.
- [30] 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:e27868.
- [31] Mehta A, Niles AN, Vargas JH, et al. Acceptability and effectiveness of artificial intelligence therapy for anxiety and depression (Youper): longitudinal observational study. *J Med Internet Res* 2021;23:e26771.
- [32] Barua PD, Vicnesh J, Gururajan R, et al. Artificial intelligence enabled personalised assistive tools to enhance education of children with neurodevelopmental disorders—a review. *Int J Environ Res Public Health* 2022;19:1192.
- [33] Lipschitz JM, Connolly SL, Miller CJ, et al. Patient interest in mental health mobile app interventions: demographic and symptom-level differences. *J Affect Disord* 2020;263:216–20.
- [34] Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. *J Med Internet Res* 2017;19:e176.
- [35] Denecke K, Schmid N, Nüssli S. Implementation of cognitive behavioral therapy in e-mental health apps: literature review. *J Med Internet Res* 2022;24:e27791.
- [36] Schachner T, Keller R, v Wangenheim F. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. *J Med Internet Res* 2020;22:e20701.
- [37] Kankanhalli A, Xia Q, Ai P, et al. Understanding personalization for health behavior change applications: A review and future directions. *AIS Trans Human-Computer Interact* 2021;13:316–49.
- [38] Alshurafa N, Jain J, Alharbi R, et al. Is more always better? Discovering incentivized mHealth intervention engagement related to health behavior trends. *Proc ACM Interactive, Mobile, Wearable Ubiquitous Technol* 2018;2:1–26.
- [39] Oakley-Girvan I, Yunis R, Longmire M, et al. What works best to engage participants in mobile app interventions and e-health: a scoping review. *Telemed e-Health* 2022;28:768–80.
- [40] Timmons AC, Duong JB, Simo Fiallo N, et al. A call to action on assessing and mitigating

bias in artificial intelligence applications for mental health. *Perspect Psychol Sci* 2023;18:1062–96.

[41] Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res* 2019;21:e13216.

[42] Ravens-Sieberer U, Kaman A, Otto C, et al. Mental health and psychological burden of children and adolescents during the first wave of the COVID-19 pandemic—results of the COPSYS study. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* 2021:1–10.

[43] Lampert T, Hoebel J, Kuntz B, et al. Gesundheitliche ungleichheiten bei kindern und jugendlichen in Deutschland—zeitliche entwicklung und trends der KiGGS-studie 2019.

[44] Reiss F, Meyrose A-K, Otto C, et al. Socioeconomic status, stressful life situations and mental health problems in children and adolescents: Results of the German BELLA cohort-study. *PLoS One* 2019;14:e0213700.

[45] Skorbura JA, Yam J. Is there an App for that?: Ethical issues in the digital mental health response to COVID-19. *AJOB Neurosci* 2022;13:177–90.

[46] Velez JA, Keene JR, Corwin M, et al. A Visual Interactive Narrative Intervention (VINI) for aphasia education: Can digital applications administer augmented input to educate stroke survivors with aphasia? *Patient Educ Couns* 2021;104:2536–43.

[47] Pham KT, Nabizadeh A, Selek S. Artificial intelligence and chatbots in psychiatry. *Psychiatr Q* 2022;93:249–53.

Supplementary Files

Untitled.

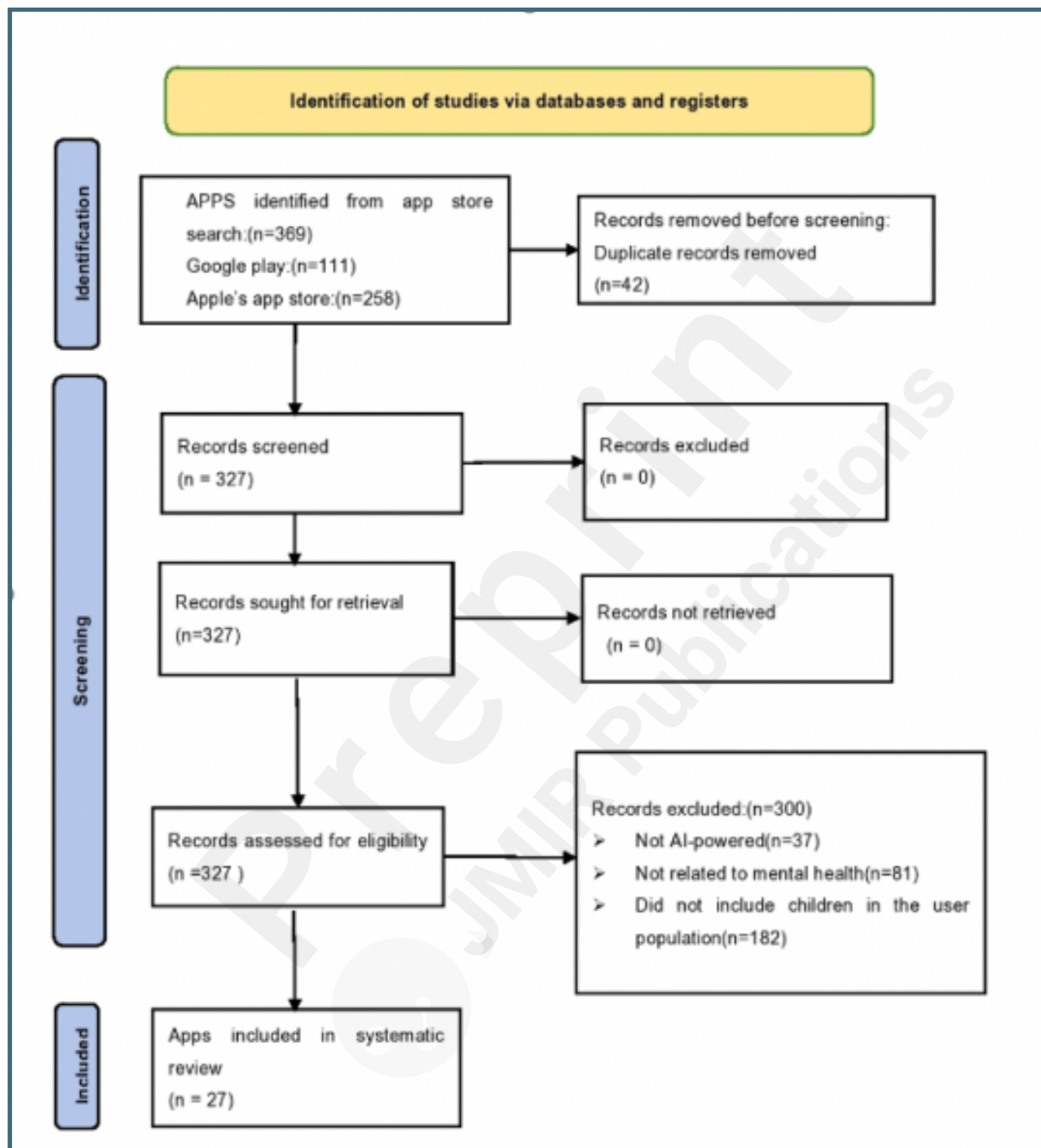
URL: <http://asset.jmir.pub/assets/aa72a1ff248ba274d422d1be91f727f8.docx>

Untitled.

URL: <http://asset.jmir.pub/assets/304883fa9b4d5e22005f394a63a767d0.docx>

Figures

PRISMA of APP Search and Screening Strategy.



Total MARS Scores and Sub-scores across Four Dimensions.

