Journal of Health and Quality of Life 11, Issue 1 (2026) 59-67

Journal of Health and Quality of Life

Journal homepage:
https://karyailham.com.my/index.php/jhqol/
ISSN: 3030-5101

Designing Adaptive Digital Mental Health Assessment and Self-Management for University Students: An Evidence-Informed Framework

Nur Izzah Khayrin Aziz[1], Ahgalya Subbiah[1,*]

[1] Faculty of Information Sciences and Engineering, Management and Science University, Section 13, 40100 Shah Alam, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 1 July 2026
Received in revised form 31 August 2026
Accepted 1 September 2026
Available online 10 September 2026

Keywords:
Digital mental health; University students; PHQ-8; GAD-7; Self-efficacy; Technology acceptance; Adaptive interface

ABSTRACT

University students experience substantial psychological distress, yet digital mental health applications often struggle with sustained engagement and can blur the boundary between screening and diagnosis. This paper develops an evidence-informed design and evaluation framework for a university mental health assessment and self-management application. A targeted narrative synthesis was conducted using validated clinical screening literature, the Health Belief Model, the Technology Acceptance Model, and peer-reviewed studies of mobile mental health engagement, adaptive interfaces, and the published Mental Health Evaluation and Lookout Program (mHELP). The evidence indicates that screening should rely on validated instruments such as PHQ-8 and GAD-7, with explicit safety and referral pathways; engagement should be evaluated separately from intention; and perceived health threat, usefulness, ease of use, and resistance to change are important design variables. Published mHELP evidence further shows that self-efficacy can increase during digital coaching and may function as an outcome of mastery experiences rather than only as a pre-use predictor. The framework therefore integrates clinical screening, behavioural beliefs, adaptive interface design, self-management modules, and longitudinal engagement metrics within a safety boundary that prevents automated diagnosis or crisis management. The resulting model offers a defensible basis for future prototype development and prospective evaluation without misrepresenting previously published trial data as new evidence.

1. Introduction

University students face a convergence of academic pressure, financial strain, social transition, and uncertainty that can increase vulnerability to anxiety, depression, and stress. Large-scale university mental health research has also documented a persistent treatment gap: many students who meet criteria for common mental health problems do not receive formal care [12]. Digital mental health interventions are therefore attractive because they can extend access, provide low-intensity support, and offer screening or self-management functions outside conventional clinic hours [10,11].

[*] *Corresponding author.*
E-mail address: ahgalya_subbiah@msu.edu.my

The main challenge is no longer whether a smartphone can deliver mental health content, but whether students will continue to engage with it long enough for useful self-management routines to form. Meta-analytic evidence shows substantial attrition in smartphone-delivered mental health interventions, while observational app-use research shows that real-world engagement is often far lower than download counts imply [7,8]. Systematic review evidence further indicates that engagement is shaped simultaneously by user characteristics, intervention content, technology design, and the implementation context [9].

A second challenge concerns clinical validity and safety. Instruments such as the Patient Health Questionnaire-8 (PHQ-8) and Generalized Anxiety Disorder-7 (GAD-7) are validated screening measures for depressive and anxiety symptom severity [5,6], but they are not stand-alone diagnostic systems. A digital application that presents these instruments must therefore distinguish screening from diagnosis, communicate score meaning conservatively, and provide clear referral or escalation routes. The omission of suicidal-ideation content in the PHQ-8 does not eliminate the need for a separate safety workflow when the application is used in populations at risk.

A third challenge is theoretical. Technology adoption models typically treat beliefs such as perceived usefulness, ease of use, and self-efficacy as antecedents of intention. However, a published university mHealth study using the Mental Health Evaluation and Lookout Program (mHELP) found that perceived health threat, perceived usefulness, and resistance to change predicted intention, while self-efficacy increased during the intervention and was associated with lower symptom scores [1]. This pattern suggests that self-efficacy may also emerge through repeated mastery experiences rather than functioning only as a static pre-use determinant. That interpretation aligns with Bandura’s account of self-efficacy as a belief strengthened through successful performance and feedback [4].

The source draft for this manuscript combined a small Malaysian prototype description with numerical findings from the published mHELP study in the United States. Because those datasets are not interchangeable, the present revision does not claim the published mHELP statistics as original results. Instead, it uses verified peer-reviewed evidence to construct a design and evaluation framework for a future university digital mental health application. The objective is to integrate clinically defensible screening, behavioral adoption theory, adaptive interface principles, and measurable engagement outcomes into one coherent model that can later be tested prospectively.

2. Methodology

2.1 Review Design and Source Integrity

A targeted narrative evidence synthesis was undertaken to revise the manuscript. The synthesis prioritized: (i) original validation studies for PHQ-8 and GAD-7; (ii) foundational behavioral theory represented by the Health Belief Model (HBM), Technology Acceptance Model (TAM), and self-efficacy theory; (iii) systematic reviews and meta-analyses of engagement and clinical outcomes in digital mental health; (iv) recent studies of adaptive mHealth interfaces and AI-supported mental health interaction; and (v) the peer-reviewed mHELP evaluation published by Zahed, Markert, and Sasangohar [1]. The approach was deliberately narrative rather than systematic because the objective was manuscript correction and framework construction, not estimation of a pooled intervention effect.

The integrity rule applied throughout the revision was that numerical findings were retained only when they could be attributed to their actual published source. Claims of common-method-bias testing, AVE, composite reliability, HTMT, mediation, moderation, or predictive relevance were removed because the source draft did not provide evidence that these analyses had been performed.

The previously copied acknowledgement and funding statement from the Texas A&M mHELP article were also excluded from the revised manuscript.

2.2 Clinical Screening Components

The PHQ-8 was selected as a validated measure of depressive symptom severity. It contains eight items scored from 0 to 3, producing a total score from 0 to 24 [5]. The GAD-7 contains seven items scored from 0 to 3 and is widely used to quantify anxiety symptom severity [6]. Within a digital system, both instruments should be implemented with the original item wording and scoring logic, accompanied by a statement that the result is a screening indication rather than a clinical diagnosis.

For safety, the application should not interpret low questionnaire burden as justification for removing crisis safeguards. A university deployment should include a clearly visible help pathway, institution-specific counselling contacts, emergency information, and a protocol for handling high-risk disclosures collected elsewhere in the application. Automated scoring may support triage or personalized education, but it should not autonomously determine psychiatric diagnosis or crisis disposition.

2.3 Behavioural and Technology-Acceptance Framework

The HBM contributes constructs that explain why a student may perceive a need to act, including perceived health threat, perceived barriers, and cues to action [2]. TAM contributes perceived usefulness and perceived ease of use as central beliefs about whether the technology is worth using and sufficiently effortless [3]. Resistance to change can be treated as an inhibiting belief that counterbalances the enabling effects of usefulness and health concern. Self-efficacy is included as both a possible baseline belief and a longitudinal outcome because digital coaching may provide mastery experiences that strengthen confidence in self-management [4].

This dual-role treatment of self-efficacy is important. In the published mHELP study, self-efficacy did not emerge as a significant direct predictor of intention or engagement, yet it increased across the intervention, and higher self-efficacy was associated with lower stress, anxiety, and depression scores [1]. Rather than labelling the null adoption path as a failure, the framework treats self-efficacy as potentially endogenous to sustained interaction.

2.4 Design Translation and Adaptive Interface Principles

The framework translates evidence into design requirements rather than if more features automatically improve engagement. A recent discrete choice experiment found that mHealth users prefer adaptation that preserves usability and controllability, occurs infrequently, and makes relatively small changes to familiar functions [14]. The application should therefore adapt presentation and support gradually for example, by changing the order or prominence of self-management modules without unexpectedly altering core navigation or removing user control.

Artificial intelligence, when included, should be bounded to low-risk functions such as content prioritization, conversational guidance, summarization of self-management progress, or detection of interaction patterns that can trigger a suggestion to review support resources. Evidence from a randomized study of socially enriched AI chatbots suggests that interaction design can influence adherence, satisfaction, and symptom outcomes [15], but such findings do not justify autonomous clinical decision-making. AI features should remain transparent, auditable, and subordinate to validated screening instruments and human-accessible referral pathways.

3. Results

3.1 Evidence-to-Design Matrix

Table 1 summarizes the evidence most directly relevant to the corrected framework. The evidence converges on four design priorities: clinical validity, retention, usability, and user agency. These priorities are related but should be measured separately because a clinically valid screener can still have poor engagement, while a highly engaging interface can still produce unsafe or misleading health communication.

Table 1
Evidence translated into design requirements

Study	Evidence base	Key finding	Design implication
Zahed et al. (2025) [1]	129 college students; mHELP; HBM/TAM	Health threat, usefulness, and resistance to change predicted intention; ease of use, usefulness, and health threat differentiated engagement.	Measure beliefs and actual use separately; model self-efficacy longitudinally.
Linardon & Fuller-Tyszkiewicz (2020) [7]	Meta-analysis of 70 smartphone RCTs	Mean attrition was 24.1% at short-term and 35.5% at longer-term follow-up.	Make retention and adherence primary implementation outcomes.
Borghouts et al. (2021) [9]	Systematic review; 208 studies	Engagement depends on user, intervention, technology, and implementation factors.	Evaluate multilevel barriers instead of attributing dropout only to motivation.
LaMontagne et al. (2024) [10]	Meta-analysis; 2,974 college students	Mindfulness apps produced small-to-moderate mental health benefits, with variable usage.	Do not infer engagement from efficacy; collect objective usage metrics.
Wang et al. (2026) [14]	Discrete choice experiment; 186 mHealth users	Users preferred controllable, infrequent, small-scale adaptations that preserve usability.	Use conservative adaptive UI rules and explicit user control.
Katsaliaki et al. (2025) [13]	HTA-oriented analysis of 172 mental mHealth apps	Evaluation should include patient, societal, usability, and market-facing indicators.	Assess quality beyond symptom change, including credibility, accessibility, and adoption.
Xu & Ma (2025) [15]	Randomized AI-chatbot trial; 84 college students	Higher social-cue design improved alliance, adherence, satisfaction, and symptoms.	AI interaction style matters, but requires safety boundaries and clinical humility.

3.2 Published mHELP Findings as External Benchmark Evidence

The published mHELP evaluation provides the closest empirical benchmark to the source draft, but it must be cited as external evidence. The study enrolled 129 participants, with 32 allocated to control and 97 to intervention; 82 intervention participants completed the pre- and post-belief questionnaires used in the principal belief analyses [1]. The reported model explained 53% of the variance in intention. Perceived health threat ($\beta = 0.35$, $p < .001$) and perceived usefulness ($\beta = 0.25$,

$p = .033$) were positive predictors, whereas resistance to change was negative ($\beta = -0.29$, $p = .027$) [1].

Engagement analyses in the same publication showed significant group differences for perceived ease of use, perceived usefulness, and perceived health threat. The intervention group also showed significant longitudinal changes in self-efficacy, cues to action, and perceived ease of use [1]. These results support the framework's separation of intention, engagement, and self-efficacy development. They do not, however, establish that any new MSU application will reproduce the same effects; that requires a new prospective study with its own ethics approval, sample, and data.

3.3 Engagement, Attrition, and Adaptive Presentation

The literature indicates that engagement is a distinct outcome rather than a proxy for clinical effectiveness. Linardon and Fuller-Tyszkiewicz reported substantial attrition across smartphone-delivered mental health trials [7], while Baumel et al. showed that real-world use of popular mental health apps is often limited after initial adoption [8]. Borghouts et al. further showed that engagement barriers can arise from user factors, content, technology, or the implementation environment [9]. These findings support a design strategy that measures sessions, module completion, return frequency, and sustained use rather than relying only on self-reported intention.

Adaptive interfaces may reduce cognitive load when they preserve user control. Wang et al. found that users favored small and infrequent adaptations and disliked changes that disrupted commonly used functions [14]. Therefore, any AI-supported adaptation should be incremental, reversible, and explainable. The application should allow users to retain a stable core navigation structure while adapting the prominence of psychoeducation, journaling, breathing, or referral content based on recent interaction patterns and stated preferences.

3.4 Clinical Screening and Outcome Interpretation

The PHQ-8 and GAD-7 provide validated numerical symptom measures [5,6], but digital feedback must avoid diagnostic overreach. A score should be framed as an indication of symptom severity that can guide self-reflection or referral, not as a determination of a mental disorder. This distinction is especially important when AI-generated language is used to explain results, because fluent text may be perceived as more clinically authoritative than it is.

Recent digital intervention evidence supports cautious optimism. Meta-analytic findings indicate that mobile interventions can reduce distress and depressive symptoms in appropriate settings [10,11,16], while gamified approaches may influence transdiagnostic mechanisms such as rumination [17]. Nevertheless, effect sizes, adherence, and follow-up durability vary across populations. The proposed framework therefore treats symptom change, self-efficacy, and engagement as separate outcome domains rather than collapsing them into a single "success" metric.

4. Proposed Application and Evaluation Framework

4.1 Integrated Architecture

Figure 1 presents the proposed architecture. Validated screening generates a conservative risk and symptom profile, which informs safe feedback and the prioritization of self-management modules. HBM and TAM beliefs influence intention and early adoption, while actual engagement creates mastery experiences that may strengthen self-efficacy. An adaptive interface loop modifies presentation gradually while preserving user control. A safety boundary applies across the system and prevents automated screening outputs from being presented as diagnosis or crisis care.

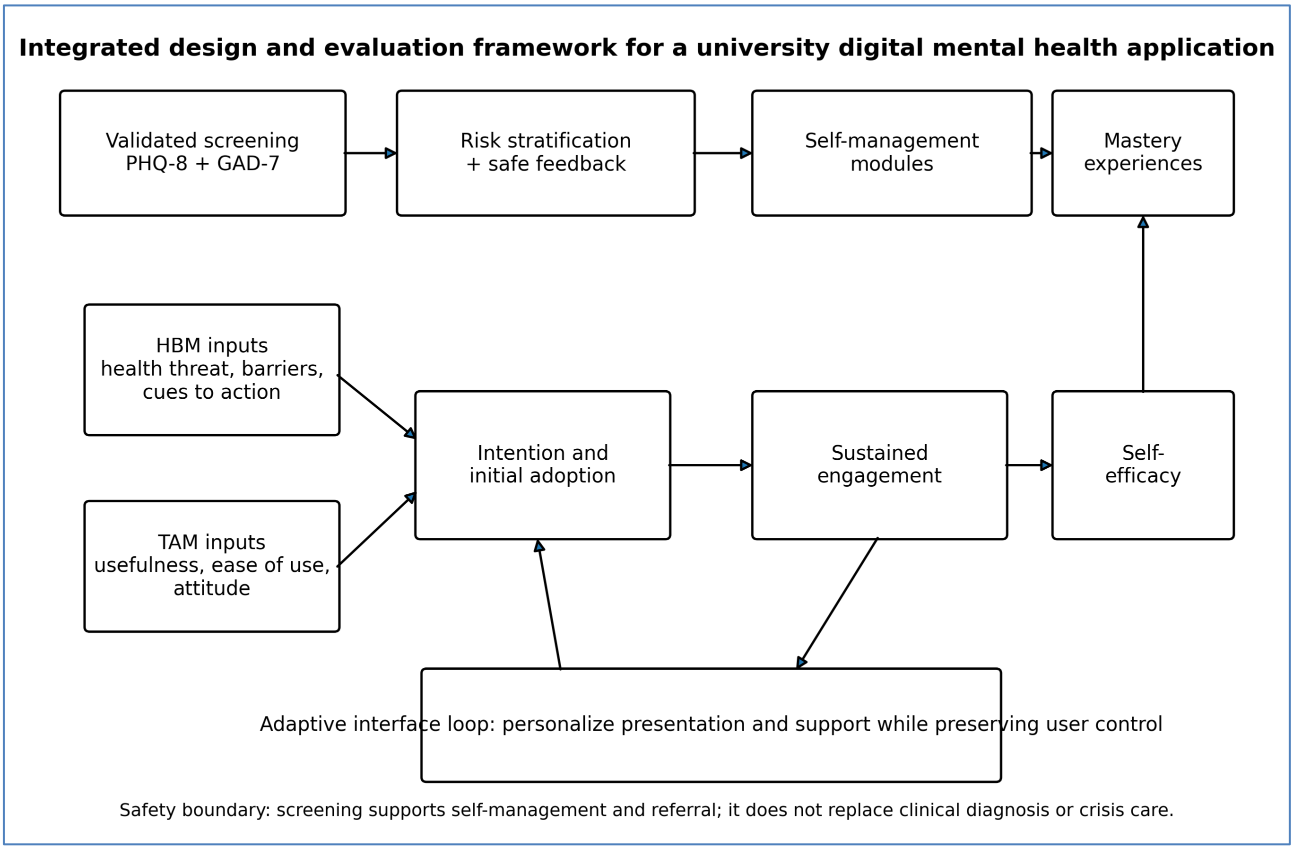

Fig. 1. Integrated design and evaluation framework for a university digital mental health application

4.2 Minimum Functional Modules

A future prototype should contain five minimum modules: (1) validated PHQ-8 and GAD-7 screening; (2) score explanation and psychoeducation written in non-diagnostic language; (3) low-intensity self-management activities such as breathing, journaling, structured reflection, or brief educational content; (4) engagement tracking based on objective interaction events; and (5) support and referral information appropriate to the university context. AI-enabled functions should be optional extensions rather than prerequisites for valid screening.

4.3 Prospective Evaluation Outcomes

A future empirical evaluation should predefine clinical, behavioural, usability, and safety outcomes. Table 2 specifies a minimum dataset. The design should distinguish between intention to use, actual engagement, and longitudinal self-efficacy. This distinction addresses a core weakness in the original compiled draft, where constructs from different studies were combined as though they belonged to one dataset.

Table 2
Minimum outcomes for a prospective evaluation

Domain	Measure/example	Purpose
Clinical screening	PHQ-8; GAD-7	Baseline and repeated symptom severity; no diagnostic claim
Technology beliefs	Perceived usefulness; perceived ease of use; resistance to change	Pre, mid, post change and association with intention
Health beliefs	Perceived health threat; barriers; cues to action	Association with intention and engagement
Self-efficacy	Validated general or mental-health-specific self-efficacy scale	Change over time; association with engagement and symptoms
Engagement	Sessions, active days, module completions, return interval	Objective sustained-use profile
Usability	SUS or equivalent plus task-based usability measures	Ease, errors, completion time, perceived burden
Safety	Referral views, escalation events, adverse events, false reassurance reports	Safety monitoring and governance
Qualitative experience	Semi-structured interview or open-ended feedback	Acceptability, burden, trust, reasons for discontinuation

4.4 Recommended Statistical Analysis

The statistical plan should match the study design and available data rather than importing tests from unrelated articles. For a repeated-measures intervention, descriptive statistics and mixed-effects or generalized estimating models are preferable when participant-level longitudinal observations are available. For ordinal or non-normal measures, justified non-parametric alternatives may be used. Reliability should be reported only for scales actually administered. If a structural model is tested, the manuscript should specify whether the analysis is PLS regression, PLS-SEM, or covariance-based SEM and report the corresponding measurement-model and structural-model diagnostics. Common-method-bias, mediation, moderation, Q^2, AVE, composite reliability, and HTMT should never be reported unless they were actually calculated from the study data.

4.5 Ethical and Safety Requirements

Any prospective student study requires institutional ethics approval, informed consent, data-minimization procedures, and a clear plan for protecting sensitive mental health information. The interface should state the limits of the tool, provide routes to professional support, and avoid persuasive design that discourages help-seeking. If AI-generated text is used, the system should identify it as automated assistance, constrain generation to approved content domains, log model-mediated recommendations for audit, and maintain a human-accessible escalation route.

5. Discussion

The corrected manuscript changes the central scientific claim. The available materials do not support a new randomized trial conducted at Management and Science University, nor do they support attributing the published U.S. mHELP dataset to new authors. What the evidence does support is a coherent framework for designing and evaluating a university mental health application without repeating common weaknesses in digital mental health research: overclaiming diagnosis, confusing intention with engagement, treating self-efficacy as fixed, and adding adaptive or AI features without governance.

Theoretically, the framework integrates HBM and TAM while allowing self-efficacy to operate dynamically. The published mHELP results are particularly informative because baseline self-efficacy did not directly predict adoption, but self-efficacy increased during the intervention and covaried with better mental health scores [1]. This pattern is consistent with Bandura's mastery-experience mechanism [4]. Future studies should therefore test reciprocal or longitudinal relationships rather than assuming a single cross-sectional path from self-efficacy to intention.

Practically, the framework suggests that a successful university application should be deliberately modest in its clinical claims and deliberate in its engagement design. Validated screening provides structure, but students also need credible explanations, low-friction self-management options, and visible pathways to human support. Adaptive interfaces should reduce burden rather than maximize novelty. The most useful AI functions are likely to be those that quietly improve relevance and accessibility while keeping the student in control.

For university administrators, the evaluation target should extend beyond symptom change. Health technology assessment work emphasizes credibility, usability, accessibility, and societal considerations [13]. Accordingly, deployment decisions should consider adoption, sustained use, safety events, equity of access, student trust, and the burden imposed on counseling services, alongside clinical symptom trajectories.

5.1 Limitations

This paper is an evidence-informed framework and narrative synthesis, not a systematic review or new clinical trial. It does not provide pooled effect estimates, and it does not validate a new AI model. The framework also draws partly on studies from contexts outside Malaysian higher education, including the published U.S. mHELP program. Cultural, institutional, linguistic, and service-delivery differences may alter both technology beliefs and help-seeking behavior. These limitations reinforce the need for a locally approved prospective study before claims of effectiveness are made.

6. Conclusions

Digital mental health applications for university students should be evaluated as sociotechnical health systems rather than as collections of screening forms and AI features. The revised framework combines validated PHQ-8 and GAD-7 screening, HBM and TAM adoption beliefs, objective engagement, adaptive interface principles, and longitudinal self-efficacy within an explicit safety boundary. Evidence from the published mHELP study is useful as a benchmark, but it must remain attributed to its original authors and dataset. The next defensible step is a prospective university study with a clearly defined sample, ethics approval, pre-specified outcomes, objective engagement measures, and analyses matched to the actual data collected. Such a design would allow future authors to make original claims about usability, adoption, self-efficacy, and mental health outcomes without compromising statistical or research integrity.

Acknowledgement
The authors acknowledge Management and Science University (MSU), Shah Alam, Malaysia, for institutional support.

References

[1] Zahed, Karim, Carl Markert, and Farzan Sasangohar. 2025. "Understanding the Role of Beliefs on Intentions and Actual Usage of a Tool for Self-Management of Mental Health among College Students." *Applied Ergonomics* 126: 104485. DOI: 10.1016/j.apergo.2025.104485.

[2] Janz, Nancy K., and Marshall H. Becker. 1984. "The Health Belief Model: A Decade Later." *Health Education Quarterly* 11 (1): 1–47. DOI: 10.1177/109019818401100101.

[3] Davis, Fred D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly* 13 (3): 319–340. DOI: 10.2307/249008.

[4] Bandura, Albert. 1997. *Self-Efficacy: The Exercise of Control.* New York: W. H. Freeman.

[5] Kroenke, Kurt, Tara W. Strine, Robert L. Spitzer, Janet B. W. Williams, Joyce T. Berry, and Ali H. Mokdad. 2009. "The PHQ-8 as a Measure of Current Depression in the General Population." *Journal of Affective Disorders* 114 (1–3): 163–173. DOI: 10.1016/j.jad.2008.06.026.

[6] Spitzer, Robert L., Kurt Kroenke, Janet B. W. Williams, and Bernd Löwe. 2006. "A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7." *Archives of Internal Medicine* 166 (10): 1092–1097. DOI: 10.1001/archinte.166.10.1092.

[7] Linardon, Jake, and Matthew Fuller-Tyszkiewicz. 2020. "Attrition and Adherence in Smartphone-Delivered Interventions for Mental Health Problems: A Systematic and Meta-Analytic Review." *Journal of Consulting and Clinical Psychology* 88 (1): 1–13. DOI: 10.1037/ccp0000459.

[8] Baumel, Amit, Frederick Muench, Stav Edan, and John M. Kane. 2019. "Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis." *Journal of Medical Internet Research* 21 (9): e14567. DOI: 10.2196/14567.

[9] Borghouts, Judith, Elizabeth Eikey, Gloria Mark, Cinthia De Leon, Stephen M. Schueller, Margaret Schneider, Nicole Stadnick, Kai Zheng, Dana Mukamel, and Dara H. Sorkin. 2021. "Barriers to and Facilitators of User Engagement With Digital Mental Health Interventions: Systematic Review." *Journal of Medical Internet Research* 23 (3): e24387. DOI: 10.2196/24387.

[10] LaMontagne, Liva G., Jennifer L. Doty, David C. Diehl, Tyler S. Nesbit, Nicholas A. Gage, Nayha Kumbkarni, and Sophia P. Leon. 2024. "Acceptability, Usage, and Efficacy of Mindfulness Apps for College Student Mental Health: A Systematic Review and Meta-Analysis of RCTs." *Journal of Affective Disorders* 367: 951–971. DOI: 10.1016/j.jad.2024.09.014.

[11] Riboldi, Ilaria, Daniele Cavaleri, Angela Calabrese, Chiara Alessandra Capogrosso, Susanna Piacenti, Francesco Bartoli, Cristina Crocamo, and Giuseppe Carrà. 2023. "Digital Mental Health Interventions for Anxiety and Depressive Symptoms in University Students during the COVID-19 Pandemic: A Systematic Review of Randomized Controlled Trials." *Revista de Psiquiatría y Salud Mental* 16: 47–58. DOI: 10.1016/j.rpsm.2022.04.005.

[12] Fitzsimmons-Craft, Ellen E., C. Barr Taylor, Michelle G. Newman, Nur Hani Zainal, Elsa E. Rojas-Ashe, Sarah Ketchen Lipson, Marie-Laure Firebaugh, et al. 2021. "Harnessing Mobile Technology to Reduce Mental Health Disorders in College Populations: A Randomized Controlled Trial Study Protocol." *Contemporary Clinical Trials* 103: 106320. DOI: 10.1016/j.cct.2021.106320.

[13] Katsaliaki, Korina, Sameer Kumar, and Panagiota Galetsi. 2025. "Patient and Societal Indicators for mHealth Apps' Evaluation Using Health Technology Assessment Framework." *Technovation* 140: 103143. DOI: 10.1016/j.technovation.2024.103143.

[14] Wang, Wei, Hourieh Khalajzadeh, John Grundy, Anuradha Madugalla, and Humphrey O. Obie. 2026. "User-Centric Requirements Prioritization in mHealth Applications: Insights from a Discrete Choice Experiment." *Information and Software Technology* 192: 108014. DOI: 10.1016/j.infsof.2026.108014.

[15] Xu, Shuo, and Tiancong Ma. 2025. "Depression Intervention Using AI Chatbots with Social Cues: A Randomized Trial of Effectiveness." *Journal of Affective Disorders* 389: 119760. DOI: 10.1016/j.jad.2025.119760.

[16] González-Robles, Alberto, Verónica Romero-Ferreiro, Pablo Roca, Tíscar Rodríguez-Jiménez, and Amanda Díaz-García. 2026. "Mobile App-Based Psychological Interventions for Depression and Anxiety in Primary Care: A Systematic Review and Meta-Analysis." *General Hospital Psychiatry* 100: 62–71. DOI: 10.1016/j.genhosppsych.2026.03.008.

[17] Pedrelli, Paola, Boaz Sofer, Shir-Lee Yatziv, and Moshe Bar. 2026. "Facilitating Thought Progression via a Gamified Mobile Application for Depression: Possible Mediators of Outcomes." *Behavior Therapy*. Online ahead of print. DOI: 10.1016/j.beth.2026.06.003.

[18] Tsai, Zoe, Adam Kiss, Shaheer Nadeem, Karen Sidhom, Sawayra Owais, Mateusz Faltyn, and Ryan J. Van Lieshout. 2022. "Evaluating the Effectiveness and Quality of Mobile Applications for Perinatal Depression and Anxiety: A Systematic Review and Meta-Analysis." *Journal of Affective Disorders* 296: 443–453. DOI: 10.1016/j.jad.2021.09.106.