Review

Digital Gamified Interventions Targeting Internalizing Mental Health Outcomes in Children and Adolescents: Systematic Review and Meta-Analysis

Julia Franziska Baschab[1], MSc; Eva Möhler[1,2], Prof Dr Med; Justine Hussong[2], PhD

[1]Child and Adolescent Psychiatry, Saarland University, Homburg, Saarland, Germany

[2]Child and Adolescent Psychiatry, Universitätsklinikum des Saarlandes, Homburg, Saarland, Germany

Corresponding Author:
Julia Franziska Baschab, MSc
Child and Adolescent Psychiatry
Saarland University
julia.baschab@uni-saarland.de
Homburg, Saarland
Germany
Phone: 49 684114238
Email: julia.baschab@uni-saarland.de

Abstract

Background: Gamified digital mental health interventions have been proposed as a scalable and engaging approach to addressing the substantial treatment gap in child and adolescent mental health, including limited access to evidence-based care and difficulties with treatment engagement. However, the evidence remains fragmented, and it is unclear which specific gamification elements contribute to clinical effectiveness and sustained engagement.

Objective: This review aimed to (1) characterize the gamification elements implemented in digital mental health interventions for children and adolescents; (2) evaluate their effects on internalizing mental health outcomes; (3) assess engagement, adherence, and user experience outcomes; and (4) descriptively explore whether specific gamification elements are associated with differences in clinical or engagement outcomes.

Methods: A systematic review and meta-analysis was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. Six electronic databases were searched for studies published between January 2018 and March 2026. Eligible studies examined gamified digital interventions targeting internalizing symptoms in children and adolescents. Gamification elements were systematically coded across 10 categories using a predefined coding framework. A 3-level random-effects meta-analysis was conducted for controlled studies reporting validated mental health outcomes.

Results: A total of 27 studies (N=28,777; mean age range 5.05-18.78 years) were included. Feedback, challenges, and progress tracking were the most frequently implemented gamification elements; social interaction and competition were rare. The meta-analysis of 5 eligible controlled studies yielded a small effect in favor of gamified interventions (g=–0.28, 95% CI –0.54 to –0.02; P=.04). However, the 95% prediction interval (–0.79 to 0.24) included zero, indicating substantial uncertainty regarding the effect expected in future comparable settings. Narrative synthesis suggested potentially more favorable effects for depression and anxiety outcomes in some clinical or at-risk samples and in interventions incorporating structured therapeutic content, although these patterns could not be formally examined as moderators. Engagement and adherence limitations were frequently reported across studies.

Conclusions: Gamified digital interventions may be associated with small improvements in internalizing outcomes in children and adolescents, although the conditions under which these interventions are most effective remain unclear. The limited evidence base and uncertainty across settings preclude firm conclusions regarding efficacy. Future research should determine which gamification elements enhance clinical outcomes and engagement, for whom they are most effective, and under which conditions.

(J Med Internet Res 2026;28:e106933) doi: 10.2196/106933

KEYWORDS

adolescents; anxiety; children; depression; digital mental health; gamification; meta-analysis; serious games; systematic review

XSL•FO
RenderX

Introduction

Mental health disorders are among the most prevalent and burdensome conditions affecting children and adolescents worldwide. Approximately 1 in 5 young people meet diagnostic criteria for a mental health disorder, with anxiety and depressive disorders being particularly common [1,2].

More recent evidence further indicates that internalizing symptoms increased substantially during the COVID-19 pandemic, with meta-analytic findings showing marked increases in depressive symptoms and smaller but significant increases in anxiety symptoms among children and adolescents compared with prepandemic levels [3,4]. Despite this substantial burden, most affected youth do not receive adequate care, with treatment gaps especially pronounced in low- and middle-income countries (LMICs) [5]. Early intervention is critical for preventing the chronification of mental health difficulties and reducing long-term impairment [6,7].

Digital mental health interventions have been proposed as a scalable and accessible approach to addressing these gaps. Delivered via mobile applications, web-based platforms, or computer-based programs, these interventions can complement traditional face-to-face care, serve as a bridge between sessions, and provide support before or after formal treatment [8,9]. This flexibility is particularly relevant for children and adolescents, who are highly familiar with digital environments and may perceive digital tools as a low-threshold and less stigmatizing way of accessing support [10].

An increasing number of interventions already target a range of outcomes, including anxiety, depression, stress, and general well-being [11-13]. However, despite their potential, digital interventions often experience low engagement and poor adherence, which can substantially limit their effectiveness in real-world settings [14]. Gamification refers to the application of game design elements in nongame contexts and has been proposed as a strategy to enhance engagement, particularly among younger populations [15]. Common elements include points, rewards, levels, avatars, narratives, and challenges. These features are assumed to enhance motivation and sustained use by targeting both extrinsic incentives (eg, rewards) and intrinsic motivational processes, such as autonomy, competence, and relatedness, as described in self-determination theory [16]. Within self-determination theory, autonomy refers to experiencing choice and volition, competence to experiencing mastery and effectiveness, and relatedness to experiencing social connection and belonging. Gamification mechanics may support these needs in different ways; for example, customizable avatars and meaningful choices may foster autonomy, levels, progress indicators, and challenges may foster competence, whereas cooperative features, peer interaction, or socially oriented game mechanics may strengthen relatedness.

Gamification approaches range from the integration of isolated game elements into otherwise conventional interventions to fully developed serious games designed for therapeutic or educational purposes [17]. A growing number of systematic reviews and meta-analyses have examined the effectiveness of gamified digital mental health interventions for children and adolescents, with mixed findings. For example, David et al [18] reported small and highly heterogeneous effects of serious games on mental health outcomes, with no sustained effects at follow-up. Bryant et al [19] found modest but significant effects for attention-deficit/hyperactivity disorder (ADHD) and depression (both g=0.28), while effects on anxiety were not significant. At the broader eHealth level, specific gamification elements such as rewards have been identified as potentially beneficial components of effective interventions alongside cognitive-behavioral techniques and self-monitoring [13]. Nevertheless, the theoretical grounding of gamification remains inconsistent, with many interventions relying primarily on extrinsic motivational strategies and lacking a coherent conceptual framework [11,12].

Taken together, these reviews offer a promising but inconsistent evidence base, and several critical gaps remain. To our knowledge, no review to date has systematically mapped which specific gamification elements are associated with clinical or engagement outcomes. This represents a critical limitation, as evidence that gamification “works” in general provides limited guidance for intervention design without identifying the underlying mechanisms and components driving these effects. Furthermore, user experience outcomes have rarely been examined alongside clinical outcomes, despite their central role in real-world adoption.

To address these gaps, this systematic review and meta-analysis synthesizes evidence on the effectiveness, engagement, and user experience of gamified digital mental health interventions for children and adolescents, with a primary focus on internalizing outcomes, including anxiety, depression, stress, emotional distress, and psychological well-being. In addition, engagement, adherence, satisfaction, and usability outcomes are synthesized where available.

This review covers studies published between January 1, 2018, and March 18, 2026. The lower boundary was chosen to extend beyond the search horizon of David et al [18] and to capture the rapid expansion of gamified digital mental health interventions in recent years, which are more representative of current app- and web-based technologies than earlier, less interactive systems.

The objectives of this review were to (1) characterize gamification elements implemented in digital mental health interventions for children and adolescents; (2) evaluate the effects of these interventions on internalizing mental health outcomes; (3) assess engagement and adherence outcomes, as well as user experience indicators such as satisfaction and usability, where reported; and (4) descriptively explore whether specific gamification elements are associated with differences in clinical or engagement outcomes.

By integrating evidence on intervention characteristics, clinical outcomes, and user-related outcomes, this review aims to identify current methodological gaps and inform the design and evaluation of future gamified digital mental health interventions for youth.

XSL•FO
RenderX

Methods

Overview

This systematic review and meta-analysis was conducted and reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [20]. The review protocol was registered before data extraction on PROSPERO (CRD420261343227). The PRISMA checklist is provided in Multimedia Appendix 1.

Ethical Considerations

No ethics approval or informed consent was required because this systematic review and meta-analysis exclusively synthesized data from previously published studies and involved no collection of new individual-level participant data.

Eligibility Criteria

Studies were included if they met the following criteria: (1) the preregistered target population comprised children and adolescents aged 4-18 years. Mixed-age samples were eligible when the study explicitly targeted adolescent or youth populations and included participants within the predefined age range. Studies focusing exclusively on adults were excluded. (2) The study evaluated a digital mental health intervention incorporating at least one gamification element (eg, points, rewards, levels, avatars, narrative/story elements, challenges, or feedback systems). Both fully gamified interventions (eg, serious games) and nongame digital interventions with embedded gamification features were eligible. Eligible delivery formats included mobile applications, web-based programs, and computer-based interventions. (3) The study reported at least one outcome related to mental health symptoms or psychosocial functioning (eg, anxiety, depression, stress, emotional distress, well-being, resilience, social-emotional functioning, bullying-related outcomes, or self-regulation) and/or implementation-related outcomes such as engagement, adherence, usability, user satisfaction, or therapeutic/intervention alliance. (4) Empirical studies using randomized controlled, quasi-experimental, or pilot/feasibility designs and published in peer-reviewed journals between January 1, 2018, and March 18, 2026, were eligible.

Studies were excluded if they evaluated nondigital interventions, digital interventions without identifiable gamification elements, or interventions not primarily targeting mental health or psychosocial outcomes. Nonempirical publications (eg, reviews, commentaries, protocols, and case reports) were also excluded. Interventions focusing primarily on cognitive or memory training, digital literacy or academic skill development, general lifestyle or health behavior, autism-specific social communication or neurodevelopmental skills, or externalizing behaviors were considered outside the scope of this review.

Specific decision rules were applied to potentially overlapping intervention classes. Virtual reality–based exposure interventions were excluded when immersive exposure constituted the primary therapeutic modality, and exergames were excluded when physical activity constituted the primary therapeutic mechanism; movement-related components remained eligible when embedded within a broader mental health intervention. ADHD-specific cognitive or attention training was excluded when ADHD-related cognitive functioning was the primary intervention target, whereas studies involving participants with ADHD remained eligible when the intervention primarily targeted internalizing mental health outcomes. Interventions developed for pediatric populations with a primary physical or somatic diagnosis (eg, cancer or chronic illness) were excluded when mental health outcomes were secondary to disease-related psychosocial support. These intervention classes were considered conceptually distinct and beyond the predefined scope of this review.

Search Strategy and Article Selection

A systematic literature search was conducted across 6 electronic databases: PubMed, Web of Science, ScienceDirect, PubPsych, EBSCOhost, and IEEE Xplore. Searches were conducted between March 17 and 18, 2026. The search strategy combined terms relating to four conceptual domains: (1) gamification (eg, gamification, serious game* and game-based); (2) digital health delivery (eg, mHealth, mobile app*, and web-based intervention*); (3) the target population (eg, child*, adolescen*, and youth); and (4) mental health outcomes (eg, anxiety, depression, stress, and well-being).

A standardized search strategy was applied across databases where possible, with syntax adapted to database-specific functionality and search limitations. In particular, a simplified combination of the core concepts was used in ScienceDirect, whereas the IEEE Xplore search was adapted by replacing truncation-based terms with explicit synonyms. The complete database-specific search strategies, including exact search strings, search dates, filters, and numbers of retrieved records, are provided in Multimedia Appendix 2. To enhance comprehensiveness, backward and forward citation searching was conducted for all included studies.

Study Selection

Records were managed using Zotero and Rayyan. Duplicate records were identified using Zotero's duplicate-detection function and manually verified before removal. The first author screened all titles and abstracts against the predefined eligibility criteria and subsequently assessed potentially eligible articles at the full-text stage. A second reviewer was consulted when eligibility was uncertain during full-text assessment, with disagreements or uncertainties resolved through discussion and consensus. Because the second reviewer did not independently screen a predefined subset of records, interrater reliability was not calculated. This screening procedure is acknowledged as a methodological limitation of the review.

Quality Assessment

The methodological quality of all included studies was assessed using the Mixed Methods Appraisal Tool (MMAT) [21], which allows for the appraisal of qualitative, quantitative, and mixed methods study designs within a single coherent framework.

The first author conducted the initial appraisal of all included studies, which was subsequently cross-checked by a second reviewer. Any uncertainties or discrepancies were discussed until consensus was reached. Because the second reviewer did

XSL•FO
RenderX

not independently assess the studies before reviewing the first author's ratings, interrater reliability was not calculated. This procedure is acknowledged as a limitation of the quality assessment. The full quality assessment results are provided in Multimedia Appendix 3.

Data Synthesis

Data were extracted by one reviewer (JFB) into a structured Microsoft Excel spreadsheet developed for this review. Extracted study characteristics included bibliographic information, country, study aims and design, control condition, follow-up period, sample size and demographic characteristics, study setting, clinical status and diagnosis, inclusion criteria, and intervention characteristics. Intervention-level information included intervention type and platform, delivery mode and guidance, intervention duration, therapeutic components, and core intervention mechanics. Outcome data were extracted for mental health symptoms and related psychosocial outcomes, including the outcome measure, reporter, assessment time points, designation as a primary or secondary outcome, and reported results. Information on intervention use, adherence, and user experience was also extracted. For studies included in the quantitative synthesis, numerical data required for effect-size calculation, including group sample sizes, means and SDs, or reported between-group standardized effect estimates, were additionally extracted where available.

Gamification characteristics were coded within the same extraction framework. Ten recurring gamification elements were coded at the study level: rewards, feedback, narrative/story, levels or progression, challenges/tasks/quests, avatars, personalization, social interaction, competition/leaderboards, and progress tracking. For each element, its implementation was described and its presence or absence coded. For descriptive purposes, gamification intensity was operationalized as the number of identified elements and categorized as low (0-3 elements), moderate (4-6 elements), or high (7-10 elements). These cutoffs were developed pragmatically to facilitate comparison across interventions and do not represent validated or empirically established thresholds. Engagement strategy was additionally classified as intrinsic, extrinsic, or mixed. The coding framework was informed by gamification concepts described in the literature and iteratively refined during data extraction to accommodate recurring features across the included interventions. Accordingly, the element-level coding and higher-order classifications were used as an exploratory descriptive framework rather than as a validated taxonomy. Data extraction and gamification coding were conducted by one reviewer and were not independently duplicated. Uncertainties were resolved through reexamination of the full-text articles.

The narrative synthesis followed a structured descriptive approach, with studies grouped according to (1) gamification elements, (2) mental health outcomes, and (3) engagement, adherence, and user experience outcomes. Engagement and adherence were treated as related but distinct constructs. Adherence referred to the extent to which participants completed intervention components or the prescribed intervention dose. Engagement was conceptualized more broadly and included initiation (eg, intervention uptake or first use), sustained usage (eg, repeated use over time), dose (eg, number of sessions or modules accessed), completion (eg, completion of prescribed intervention content), and indicators of engagement quality (eg, enjoyment, satisfaction, or perceived involvement). Because the included studies varied substantially in how these constructs and other outcomes were operationalized and reported, findings not eligible for quantitative pooling were synthesized descriptively using the metrics and direction of effects reported in the original studies.

A total of 5 controlled studies reporting comparable internalizing mental health outcomes, including anxiety, depression, stress, psychological distress, emotional symptoms, and internalizing symptoms, were included in the quantitative synthesis. All meta-analyses were conducted in R version 4.6.0 [22] using the *metafor* package. The *dplyr* and *ggplot2* packages were used for data handling and visualization, respectively.

Effect sizes were expressed as standardized mean differences (Hedges *g*) using postintervention data or reported between-group standardized effect estimates, depending on the data available in the original studies. Because several studies contributed multiple outcomes, dependent effect sizes were modeled using a 3-level random-effects meta-analytic model with restricted maximum likelihood estimation. The model accounted for sampling variance (level 1), variance between outcomes within studies (level 2), and variance between studies (level 3). Given the small number of independent studies, t-distribution–based inference with containment-adjusted degrees of freedom was used. A 95% prediction interval was calculated to characterize the expected range of effects across future comparable studies. Statistical heterogeneity was assessed using the I^2 statistic, with values of approximately 25%, 50%, and 75% interpreted as indicating low, moderate, and high heterogeneity, respectively [23].

A conventional random-effects model with the Hartung-Knapp-Sidik-Jonkman adjustment was conducted as a sensitivity analysis. Because this model does not account for dependence among multiple outcomes from the same study, the 3-level model was retained as the primary analysis. Moderator analyses examining gamification elements, clinical status, and intervention type were not conducted because the number of independent studies was insufficient for reliable moderator analyses [24].

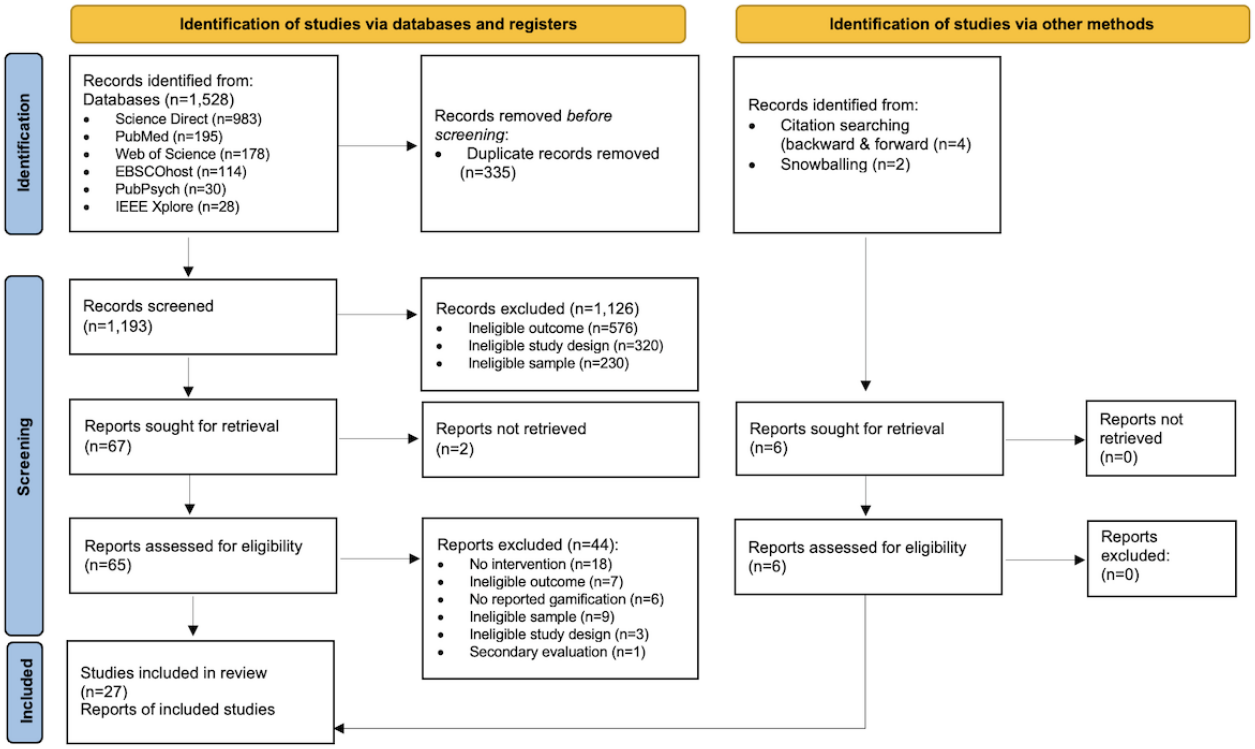
Results

Study Selection

The study selection process is illustrated in the PRISMA 2020 flow diagram (Figure 1 [20]). A total of 1528 records were identified through database searching. After removal of 335 duplicates, 1193 records were screened at title and abstract level. Of these, 1126 records were excluded based on the predefined eligibility criteria. The main reasons for exclusion were: ineligible outcome (n=576), ineligible study design (eg, reviews or commentaries; n=320), ineligible population (eg, adult samples; n=230), and no full-text availability (n=2).

XSL•FO
RenderX

Figure 1. Flow diagram of systematic search as per PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.



A total of 65 full-text articles were assessed for eligibility, of which 44 were excluded: no eligible intervention (n=18), ineligible sample (n=9), ineligible outcome (n=7), no identifiable gamification elements (n=6), ineligible study design (n=3), and secondary evaluation (n=1). Articles were classified as having no eligible intervention when they did not evaluate a digital mental health intervention meeting the predefined eligibility criteria, and as having no identifiable gamification elements when neither the intervention description nor available intervention materials contained gamification elements according to the coding framework applied in this review.

Following full-text screening, 21 studies met the inclusion criteria and were included in the review. To ensure comprehensive coverage, backward and forward citation searching, as well as snowballing, were conducted on all included studies, resulting in the identification of 6 additional eligible studies. This yielded a final sample of 27 studies included in the review.

Overview of Included Studies

The final sample comprised 27 studies published between 2018 and 2025, collectively reporting on 28,777 participants across 18 countries. Studies were predominantly conducted in high-income, English-speaking countries, with the United Kingdom (k=4) and the United States (k=3) most frequently represented. Several studies were conducted in LMICs, including South Africa [25], Uganda [26], and Mexico [27].

Sample sizes varied considerably, ranging from N=5 [28] to N=21,320 [29]. The largest sample [29] reflects a real-world retrospective implementation study and is not directly comparable to controlled trial samples. Excluding this study, sample sizes ranged from 5 to 3153. Mean ages ranged from 5.05 years [30] to 18.78 years [31]. Several studies included mixed-age samples extending to 19 years. Dietvorst et al [31] included 2 adolescent samples with mean ages of 16.19 and 18.78 years, respectively. These studies were retained because they explicitly targeted adolescent or youth populations and included participants within the preregistered target age range. The proportion of female participants ranged from 33.0% [32] to 91.7% [33]. Most studies targeted adolescents (approximately 12-17 years), with comparatively fewer studies focusing on younger children [30,34-37]. A complete overview of study characteristics is presented in Table 1.

XSL•FO
RenderX

Table 1. Study characteristics of included studies (k=27).

Author	Country	Study design	Participants, n	Age, mean (range)	Setting	Clinical status	Intervention	Type	Duration	Theoretical framework
Birrell et al [38]	Australia	Cluster RCT[a]	166	15.3 (14-16)	School	Nonclinical	Mind your Mate	App (blended)	12 months	Psychoeducation, peer support
Bohr et al [39]	Canada	Pilot cluster RCT	24	—[b] (13-18)	Community	At-risk	SPARX	Serious game	7 weeks	CBT[c]
Chen et al [40]	China	RCT	105	13.3 (13-15)	School	Nonclinical	GISCC	Serious game	4 weeks	Social learning theory, CBT
David et al [41][d]	Romania	3-arm RCT	142	13.0 (10-16)	School	Nonclinical	REThink	Serious game	4 weeks	REBT[e]
David et al [42][d]	Romania	Secondary analysis	48	13.0 (10-16)	School	Nonclinical	REThink	Serious game	4 weeks	REBT
Dietvorst et al [31]	Netherlands	Feasibility study	1720	17.8 (—)	Community (online)	Nonclinical	Grow It!	App (serious game)	3-6 weeks	CBT
Dietvorst et al [43]	Netherlands	Pre-post (proof-of concept)	3153	17.9 (—)	Community (online)	Nonclinical	Grow It!	App (serious game)	3-6 weeks	CBT
Fleming et al [29]	New Zealand	Retrospective cohort	21,320	14.9 (12-19)	Community (online)	Mixed (mild-moderate depression)	SPARX	Serious game	7 modules	CBT (iCBT[f])
Grist et al [33]	United Kingdom	Qualitative study (interviews)	44	15.98 (—)	Clinic (CAMHS[g])	Clinical	BlueIce	App	12 weeks	DBT[h], CBT, mindfulness
Hugh-Jones et al [44]	United Kingdom	Feasibility study	31	— (15-17)	School	At-risk	MindMate2U	App	6 weeks	CBT, ACT[i], mindfulness
Keinonen et al [45]	Finland	Cluster RCT	106	11.1 (10-12)	School	Nonclinical	Magis	Serious game	4 weeks	ACT
Lepe-Salazar et al [27]	Mexico	Pilot RCT	92	17.6 (16-19)	School	Nonclinical	Bernstein	Serious game (RPG[j])	Single session (~60 min)	CBT (MCS[k])
Lincke et al [46]	Germany	Feasibility study	36	15 (12-18)	Clinic (outpatient)	Clinical (MDD[l])	e-MICHI	App (blended)	6 weeks	CBT
Lockwood et al [34]	United Kingdom	Pre-post evaluation	120	9.8 (7-12)	School	Nonclinical/at-risk	Lumi Nova	App (serious game)	8 weeks	CBT (exposure)
McCashin et al [35]	Ireland	Pragmatic RCT	122	9.9 (8-12)	Primary care	Clinical	Pesky gNATs	Serious game	7 game levels	CBT
Miranda et al [47]	United States	Feasibility study	24	— (12-18)	Mixed (clinic and community)	Mixed (clinical: AN[m], nonclinical)	BALANCE	App	At least 30 min	DBT, mindfulness
Mittmann et al [48]	Austria	Mixed methods	91	11.4 (10-13)	School	Nonclinical	LINA	Serious game (AR[n])	Single session (~1 hr)	Contact theory, SEL[o]
Moffett et al [25]	South Africa	Pilot RCT	195	16.1 (15-19)	School	At-risk/clinical	Kuamsha	App (serious game)	11 weeks	Behavioral activation
Nicolaidou et al [36][p]	Cyprus	Usability study	74	8.4 (5-12)	Community	Nonclinical	Kids' Stress Relief	App (interactive storytelling)	Single session (~7.5 min)	Mindfulness, positive psychology

XSL•FO
RenderX

Author	Country	Study design	Participants, n	Age, mean (range)	Setting	Clinical status	Intervention	Type	Duration	Theoretical framework
Nicolaidou et al [30][p]	Cyprus	Qualitative study	20	5.1 (4-6)	Community	Nonclinical	Kids' Stress Relief	App (interactive storytelling)	Single session	Mindfulness, CBT-informed
Pera et al [28]	Italy	Usability study	5	— (9-12)	Clinic	At-risk/clinical (anxiety)	iCBT WebApp	App (iCBT)	Single session (pilot)	CBT
Peuters et al [49]	Belgium	Quasi-RCT (cluster)	279	13.6 (12-15)	School	Nonclinical	#LIFE-GOALS	App	12 weeks	HAPA[q], BCT[r], Self-regulation
Pine et al [32]	New Zealand	Feasibility study	45	12.5 (12-14)	School	Nonclinical	Match Emoji	Casual video game	2 weeks	Mindfulness, 5 ways to well-being
Pozuelo et al [26]	Uganda	Feasibility study	31	17.6 (15-19)	Community	Nonclinical	Kuamsha	App (gamified narrative)	11 weeks	Behavioral activation
Pramana et al [50]	United States	Open clinical trial	35	11.2 (9-14)	Clinic	Clinical (anxiety)	SmartCAT 2.0	App	8 sessions	CBT
Stallard et al [51]	United Kingdom	Mixed methods (nested in RCT)	60	15.6 (12-17)	Clinic (CAMHS)	Clinical	BlueIce	App	12 weeks	DBT-informed
Williams et al [52]	United States	Cluster RCT	699	11.6 (11-14)	School	Nonclinical	LST + Galaxia	Blended (game + e-learning)	~20 sessions	CBT, social skills training

[a]RCT: randomized controlled trial.

[b]Not available.

[c]CBT: cognitive behavioral therapy.

[d]David et al [42] is the secondary analysis of David et al [41].

[e]REBT: rational emotive behavior therapy.

[f]iCBT: internet-delivered cognitive behavioral therapy.

[g]CAMHS: child and adolescent mental health services.

[h]DBT: dialectical behavior therapy.

[i]ACT: acceptance and commitment therapy.

[j]RPG: role-playing games.

[k]MCS: multiple composite scenarios.

[l]MDD: major depressive disorder.

[m]AN: anorexia nervosa.

[n]AR: augmented reality.

[o]SEL: social-emotional learning.

[p]Nicolaidou et al [36] represents the preliminary work for Nicolaidou et al [30].

[q]HAPA: health action process approach.

[r]BCT: behavior change technique.

Quality Assessment

Methodological quality was assessed using the MMAT [21]. All included studies met the MMAT screening criteria and were assessed using the corresponding design-specific criteria. Randomized controlled trials (RCTs) generally reported appropriate randomization procedures and adequate baseline comparability. However, completeness of outcome data and adherence to assigned interventions varied across studies, with limitations particularly evident in trials characterized by substantial attrition or limited intervention uptake. Blinding of outcome assessors was frequently not reported or not feasible, representing a potential source of bias, particularly for self-reported outcomes. Among nonrandomized and observational studies, outcome measures were generally appropriate, although participant representativeness and control of potential confounding varied. Qualitative and mixed methods studies were assessed using the corresponding design-specific MMAT criteria. Detailed MMAT ratings for all included studies are provided in Multimedia Appendix 3.

Study Characteristics

Study Design

The included studies used a heterogeneous range of methodological designs, including RCTs, cluster RCTs,

XSL•FO
RenderX

quasi-experimental studies, single-arm pilot or feasibility studies, observational designs, and qualitative or mixed methods approaches.

Ten studies used randomized controlled designs, including cluster RCTs, pilot RCTs, quasi-RCTs, and standard RCTs [25,27,35,38-41,45,49,52]. Control conditions varied across studies and included waitlist controls, active comparator conditions, treatment-as-usual conditions, and alternative intervention conditions. One study included both waitlist and active face-to-face comparison conditions [42]; however, this study represented a secondary analysis of the RCT reported by David et al [41] and was therefore not counted as an independent randomized trial.

A total of 13 studies used quantitative nonrandomized or observational designs. These included feasibility and usability studies, single-arm pre-post evaluations, an open clinical trial, a retrospective cohort study, and a secondary analysis of previously collected RCT data. Several studies explicitly indicated that they were not powered to evaluate clinical efficacy [26,28,32,44,46,47]. Two studies used qualitative designs [30,33], and 2 used mixed methods designs [48,51], with the latter including 1 mixed methods study nested within an RCT [51].

Follow-up data were available in only a minority of studies. Where reported, follow-up periods ranged from 2 weeks [32] to 12 months [38]. Notable longer-term assessments included a 3-month follow-up [46] and a 24-week follow-up assessment [25]. Overall, long-term follow-up data were limited across the included studies, indicating a gap in evidence regarding sustained intervention effects.

Settings and Clinical Characteristics

Studies were conducted across 3 primary settings: school-based, clinical, and community or online settings. School-based studies (k=11) [25,27,34,38-42,44,45,52] primarily targeted nonclinical or at-risk populations within universal or selective prevention frameworks. Clinical studies (k=7) [28,33,35,46,47,50,51] were conducted in outpatient, primary care, inpatient, or community mental health service settings. Community or online-based studies (k=9) [26,29-32,36,43,48,49] were delivered outside formal educational or clinical institutions, often via digital platforms accessible in naturalistic contexts.

In terms of clinical status, studies included nonclinical general population samples (k=16), at-risk or subclinical samples (k=5), and clinical samples (k=9). Clinical samples included participants with diagnosed or clinically significant symptoms of major depressive disorder [46], anxiety disorders (including generalized anxiety disorder, social anxiety disorder, and separation anxiety disorder [28,35,50]), internalizing problems broadly defined [35], self-harm and comorbid depression and anxiety [33,51], and anorexia nervosa as the primary clinical diagnosis [47]. Several studies used dimensional inclusion criteria based on questionnaire cutoff scores rather than formal diagnoses [25,26]. Categories were not mutually exclusive because several studies included mixed clinical samples.

Intervention Characteristics

Interventions varied substantially in type, delivery format, and theoretical grounding. Most interventions could be classified as either serious games (purpose-built digital games embedding therapeutic content) or gamified applications (nongame digital interventions incorporating discrete game elements). A smaller number of studies used blended approaches combining digital intervention components with face-to-face or guided human support [25,26,46,52].

Delivery platforms included mobile applications delivered via smartphones or tablets, web-based or computer-based programs, and one augmented reality-based intervention. Delivery modes ranged from fully unguided self-directed use [29,32] to therapist-guided or teacher-facilitated delivery [35,52], with a subset using peer mentor support [25,26]. Intervention duration ranged from single-session interventions of approximately 7.5-90 minutes [27,30,36] to longitudinal programs lasting up to 12 months [38]. Most structured interventions comprised 4-12 modules or sessions delivered over 4-12 weeks.

Cognitive-behavioral and related therapeutic approaches were most common, including cognitive behavioral therapy (CBT) and CBT-informed interventions, rational emotive behavior therapy [41,42], dialectical behavior therapy (DBT) [33,47,51] and acceptance and commitment therapy (ACT) [44,45]. Other theoretical approaches included behavioral activation [25,26], mindfulness-based approaches [30,32,36], social learning theory [40], and contact theory [48]. A subset of studies did not explicitly report a theoretical framework or relied primarily on psychoeducational content without an identified therapeutic model [38,49].

Gamification Elements

The gamification elements present in each included study are summarized in Table 2. Across all 27 studies, 10 distinct gamification elements were identified and coded at study level: rewards, feedback, narrative or story, levels or progression, challenges or quests, avatars, personalization, social interaction, competition or leaderboards, and progress tracking.

XSL•FO
RenderX

Table 2. Gamification elements across included studies (k=27). Checkmarks indicate the presence of a gamification element, whereas empty cells indicate its absence.[a]

Study		Gamification elements										Intensity
Author	Intervention	Rewards	Feedback	Narrative	Levels	Challenges	Avatar	Personalization	Social interaction	Competition	Progress tracking	
Birrell et al [38]	Mind your Mate		✓			✓		✓	✓		✓	Moderate
Bohr et al [39]	SPARX	✓	✓	✓	✓	✓	✓				✓	High
Chen et al [40]	GISCC	✓	✓	✓	✓	✓	✓	✓			✓	High
David et al [41]	REThink	✓	✓	✓	✓	✓	✓				✓	High
David et al [42]	REThink	✓	✓	✓	✓	✓	✓				✓	High
Dietvorst et al [31]	Grow It!	✓	✓		✓	✓		✓	✓	✓	✓	High
Dietvorst et al [43]	Grow It!	✓	✓		✓	✓		✓	✓	✓	✓	High
Fleming et al [29]	SPARX	✓	✓	✓	✓	✓	✓				✓	High
Grist et al [33]	BlueIce		✓			✓		✓			✓	Moderate
Hugh-Jones et al [44]	MindMate2U		✓		✓	✓		✓			✓	Moderate
Keinonen et al [45]	Magis		✓	✓	✓	✓						Moderate
Lepe-Salazar et al [27]	Bernstein	✓	✓	✓	✓	✓	✓	✓			✓	High
Lincke et al [46]	e-MICHI		✓		✓	✓		✓			✓	Moderate
Lockwood et al [34]	Lumi Nova	✓	✓	✓	✓	✓	✓	✓			✓	High
McCashin et al [35]	Pesky gNATs		✓	✓	✓	✓	✓				✓	Moderate
Miranda et al [47]	BALANCE	✓	✓			✓	✓				✓	Moderate
Mittmann et al [48]	LINA		✓	✓	✓	✓	✓		✓			Moderate
Moffett et al [25]	Kuamsha	✓	✓	✓	✓	✓		✓			✓	High
Nicolaidou et al [36]	Kids' Stress Relief	✓	✓	✓	✓	✓					✓	Moderate
Nicolaidou et al [30]	Kids' Stress Relief	✓	✓	✓	✓	✓						Moderate
Pera et al [28]	iCBT WebApp	✓	✓		✓	✓	✓	✓			✓	High
Peuters et al [49]	#LIFEGOALS	✓	✓	✓		✓	✓	✓			✓	High
Pine et al [32]	Match Emoji		✓		✓	✓						Low
Pozuelo et al [26]	Kuamsha	✓	✓	✓	✓	✓		✓			✓	High
Pramana et al [50]	SmartCAT 2.0	✓	✓		✓	✓		✓			✓	Moderate
Stallard et al [51]	BlueIce		✓					✓			✓	Low
Williams et al [52]	LST + Galaxia	✓	✓	✓	✓	✓	✓	✓			✓	High
Total present, n	N/A[b]	18	27	16	22	26	13	16	4	2	23	N/A
Prevalence, %	N/A	67	100	59	81	96	48	59	15	7	85	N/A

[a]Rewards included points, badges, trophies, virtual currencies, or tangible prizes. Feedback comprised in-app responses, mood-tracking responses, or push notifications. Narrative = story-based or thematic framing elements. Levels = progression systems or staged advancement. Challenges = quests, missions, or task-based activities. Avatar = fixed or customizable characters representing or guiding the user. Personalization = individualized goals,

XSL•FO
RenderX

tailored content, adaptive tasks, or personalized monitoring. Social interaction = peer interaction, collaboration, or social sharing features. Competition = leaderboards or competitive mechanics. Progress tracking = dashboards, completion indicators, logs, or monitoring systems. Gamification intensity was classified according to the number and integration of coded elements: low intensity (0-3 elements), moderate intensity (4-6 elements), and high intensity (7-10 elements).

[b]N/A: not applicable.

Feedback was the only element in all 27 (100%) studies, encompassing in-app performance feedback, mood tracking responses, therapist or teacher notifications, and push notifications. Challenges or task-based activities were present in 26 (96%) studies, and progress tracking (including dashboards, module completion indicators, or in-game logs) in 23 (85%) studies, reflecting their near-universal integration as core mechanisms for structuring goal-directed activity and monitoring engagement across intervention types.

Levels or progression structures were used in 22 (81%) studies, enabling sequential skill acquisition and providing an architectural scaffold for therapeutic content delivery. Rewards (eg, points, badges, virtual coins, trophies, or tangible prizes) were present in 18 (67%) studies, and both narrative or story-based elements and personalization features were present in 16 (59%) studies each. Narrative elements ranged from fully developed fantasy-based story worlds (eg, SPARX, REThink, Lumi Nova, and Bernstein) to more loosely framed thematic or story-based structures. Personalization encompassed user-selected goals, culturally tailored content, individualized mood monitoring, and therapist-adapted exposure tasks. Avatar systems appeared in 13 (48%) studies, ranging from fixed guide characters (eg, RETMAN in REThink; David gNATtenborogh in Pesky gNATs) to fully customizable player avatars (eg, Lumi Nova).

In contrast, social and competitive elements were comparatively rare. Social interaction features (eg, peer collaboration or social sharing functions) were present in only 4 (15%) studies, and competition or leaderboard mechanics in only 2 (7%) studies. Overall, the coded interventions therefore relied more frequently on individually oriented than on socially comparative gamification features.

Gamification intensity was operationalized based on the number of coded elements per study, with predefined cutoffs used to facilitate descriptive grouping rather than as empirically validated thresholds. Two studies were classified as low intensity (2-3 elements): BlueIce [51] and Match Emoji [32]. Eleven studies were classified as moderate intensity (incorporating 4 to 6 elements), including Mind your Mate [38], BlueIce in the earlier feasibility study [33], MindMate2U [44], Magis [45], e-MICHI [46], BALANCE [47], LINA [48], Pesky gNATs [35], both Kids' Stress Relief studies [30,36], and SmartCAT 2.0 [50]. A total of 14 studies were rated as high intensity (7 or more of the 10 coded elements): GISCC [40], both REThink studies [41,42], both Grow It! studies [31,43], SPARX [29,39], Bernstein [27], Lumi Nova [34], both Kuamsha studies [25,26], internet-delivered cognitive behavioral therapy (iCBT) WebApp [28], #LIFEGOALS [49], and LST + Galaxia [52].

Gamification Elements and Outcomes: Descriptive Patterns Across Studies

A central aim of this review was to explore whether particular gamification features appeared to co-occur with favorable mental health outcomes. Given the co-occurrence of multiple gamification and therapeutic components, the heterogeneity of study designs and populations, and the absence of dismantling or factorial designs, these comparisons were descriptive and exploratory.

For this exploratory comparison, 5 independent studies, represented by 6 publications reporting statistically significant improvements on validated symptom or psychosocial outcomes [34,40-42,46,52], were contrasted with 4 studies reporting null or predominantly nonsignificant findings on comparable outcomes [25,38,45,49]. Secondary analyses of the same trial were not counted as independent studies, and studies reporting only feasibility, qualitative, proxy, or nonsymptom outcomes were not included in this descriptive comparison. Among the 5 independent studies reporting significant improvements, 4 used high-intensity gamification, 4 incorporated narrative or story-based elements, and 4 included personalization. However, these features were also present in several studies with null or predominantly nonsignificant findings, and no individual gamification element consistently distinguished interventions with favorable outcomes from those with null or mixed findings. Rewards were present in interventions reporting improvements in depression or anxiety [34,40,41,46], but were also used in studies with predominantly null findings [38,49].

Effects on Mental Health Outcomes

Depression

Depression was among the most frequently examined mental health outcomes across the included studies. As summarized in Table 3, findings were mixed, with significant improvements reported in some controlled and uncontrolled studies, whereas others found small or nonsignificant effects.

XSL•FO
RenderX

Table 3. Outcome summary: mental health, engagement, and user experience (k=27).[a]

Author	Primary outcomes	Key results	Direction	Adherence/user experience
Birrell et al [38]	Depression, anxiety, distress	Marginal ↓ depression (*d*=–0.31, but ns[b]); no effects on anxiety/distress	Null	49% registered; low use (mean 1.57 modules)
Bohr et al [39]	Depression, resilience, ER[c]	No effects on depression/resilience; ↓ maladapt. ER strategies	Mixed	Not reported
Chen et al [40]	Cyberbullying	Cyberbullying ↓ (*d*=0.29; *P*<.05); secondary ↓ depression/anxiety/stress (*d*=0.46-0.54)	Positive	Not reported
David et al [41]	Depression, emotional symptoms	Significant ↓ depressive mood and symptoms (*d*=0.66-0.84)	Positive	13.9% dropout at baseline; no further attrition
David et al [42]	Symptom change	Game performance predicted symptom improvement (*r*=0.29)	Positive	Not reported
Dietvorst et al [31]	Reflection, well-being	Increased reflection (66%-72%); no clinical effects tested	Exploratory	App rated 7.1-7.2/10; 72.6%-75.6% recommend
Dietvorst et al [43]	Depression, anxiety, well-being	Significant ↑ well-being (*d*=0.20-0.32); secondary small ↓ depression/anxiety (*d*=0.08-0.17)	Positive	App rated 7.14-7.17/10; challenge completion 25%-33%
Fleming et al [29]	Depression	Significant ↓ depression (*d*=0.38-0.51)	Positive	Very low completion (3.1% full; 7.4% ≥4 modules)
Grist et al [33]	Self-harm	↓ self-harm urges (qualitative evidence)	Exploratory	100% would recommend
Hugh-Jones et al [44]	Mental health, well-being	Nonsignificant trend toward improvement (not powered)	Trend	Mean topic rating 4.12/5; 94% motivated
Keinonen et al [45]	Psychological flexibility	No main effect overall; ↓ in high-inflexibility subgroup (*d*=0.35); ↓ loneliness (*P*=.02)	Mixed	Not reported
Lepe-Salazar et al [27]	Social skills	↑ assertiveness (*P*<.05)	Positive	Usability 3.63/5; 8% dropout
Lincke et al [46]	Depression	Significant ↓ depression; maintained at follow-up	Positive	80% good/excellent adherence; 26% dropout
Lockwood et al [34]	Anxiety	Significant ↓ anxiety (*d*=0.35); other primary measures ns	Positive	mean 11 sessions over mean 18 days; 75% “easy”
McCashin et al [35]	Internalizing problems	No group effect; within-group improvement	Mixed	88%-90% recommend; retention 66% (T2), 44% (T3)
Miranda et al [47]	Eating disorder distress	↓ distress reported (qualitative evidence)	Exploratory	100% user-friendly; 88% recommend
Mittmann et al [48]	Stigma, belonging	↑ understanding and belonging (qualitative)	Exploratory	Usability 33.64/45, enjoyment & social connectivity highest-rated
Moffett et al [25]	Depression	No significant effects (*d*=–0.07)	Null	80.2% adherence (≥4 modules)
Nicolaidou et al [36]	Stress, relaxation	High immediate relaxation response (94.4%, descriptive)	Exploratory	70.4% recommend; usability=grade B
Nicolaidou et al [30]	Emotion recognition	High comprehension and engagement (descriptive)	Exploratory	85% wanted app for home use
Pera et al [28]	Anxiety, usability	High usability; anxiety improved	Exploratory	No dropout (single session, n=5)
Peuters et al [49]	Health-related quality of life	No effect (in-school); HRQoL[d] declined more in intervention (remote subgroup; *P*<.01)	Negative	Low engagement: 18% never used; 30% stopped by week 2
Pine et al [32]	Anxiety, depression, mindfulness	Small ↓ anxiety/depression (*P*=.049); mindfulness ns	Mixed	High engagement (87% ≥60 min; 71% active at week 6)
Pozuelo et al [26]	Depression	No significant effects; trends only	Trend	Retention 94% (29/31); 45% continued post study
Pramana et al [50]	Engagement	↑ Usage frequency (median 68 vs 37; *P*<.01)	Positive	~2x/day use; highest early, tapering later
Stallard et al [51]	Self-harm prevention	77% reported prevention of ≥1 episode (qualitative)	Exploratory	4.09/5 stars; 95.0% used ≥once, 73.3% ≥6 times

XSL•FO
RenderX

Author	Primary outcomes	Key results	Direction	Adherence/user experience
Williams et al [52]	Bullying, life skills knowledge	Significant ↓ bullying/cyberbullying (P<.05); ↑ skills knowledge (P<.01)	Positive	Not reported

[a]Direction categories were standardized as follows: positive=statistically significant favorable findings on the reported outcomes without significant deterioration on other relevant outcomes; mixed=a combination of significant and nonsignificant findings across relevant outcomes; null=no statistically significant effects; negative=statistically significant deterioration in at least one mental health or psychosocial outcome without evidence of significant improvement; trend=nonsignificant change in the expected direction; exploratory=descriptive or qualitative favorable findings without inferential testing of the relevant outcome. Effect sizes are reported as presented in the original studies and may therefore not be directly comparable across studies.

[b]ns: nonsignificant.

[c]ER: emotion regulation.

[d]HRQoL: health-related quality of life.

Lincke et al [46] found significant reductions in depression severity on both the Beck Depression Inventory-II (mean change=–6.76; P=.01) and the clinician-rated Children's Depression Rating Scale-Revised (P<.001) in adolescents with major depressive disorder receiving e-MICHI as an adjunct to outpatient treatment, with improvements stable at 3-month follow-up. Fleming et al [29] observed moderate within-person reductions on the Patient Health Questionnaire adapted for adolescents (PHQ-A) across SPARX modules (d=0.38 for modules 1-4; d=0.51 for modules 1-7) in a large real-world implementation sample, noting, however, that only 3.1% of registered users completed all 7 modules. Chen et al [40] reported a significant between-group effect on the Depression Anxiety Stress Scale-21 (DASS-21) depression subscale (d=0.46; P=.02). David et al [41] found significant within-group reductions in depressive mood for the REThink arm (d=0.84; P<.01) that were superior to the waitlist condition at posttest (d=0.66; P=.01).

In contrast, several controlled trials did not detect significant between-group effects. Moffett et al [25] reported a negligible effect for Kuamsha vs an active control on the PHQ-A (d=–0.07; P=.65). Bohr et al [39] found no significant changes in Centre for Epidemiologic Depression Scale-Revised (CESD-R) scores in their pilot RCT of SPARX among Inuit youth, although improvements were observed in maladaptive cognitive emotion regulation strategies. Birrell et al [38] similarly found no significant between-group effect on the PHQ-A at 6- or 12-month follow-up (d=–0.31; P=.052).

Evidence from uncontrolled or naturalistic studies was similarly mixed. The Grow It! app [43] showed small but statistically significant reductions in Children's Depression Inventory Questionnaire-12 scores across 2 large cohorts (cohort 1: d=0.08; cohort 2: d=0.17), although the absence of a control group limits causal inference. Pine et al [32] reported small reductions in anxiety and depression following Match Emoji (P=.049), whereas Pozuelo et al [26] observed only a nonsignificant trend toward improvement in PHQ-A following Kuamsha use in Uganda (d=0.30; P=.16).

Anxiety

A total of 6 studies reported anxiety or closely related internalizing symptom outcomes. Lockwood et al [34] reported significant reductions in parent-reported anxiety symptoms on the Spence Child Anxiety Scale-Parent version 8 after 8 weeks of Lumi Nova use (d=0.35; P=.01). Chen et al [40] demonstrated significant between-group reductions in anxiety on the DASS-21 in the GISCC group compared to controls (d=0.53; P=.01). Pine et al [32] observed a small but statistically significant reduction in Revised Children's Anxiety and Depression Scale total scores over 2 weeks following use of the Match Emoji casual game (mean difference=–3.42; P=.049). Pera et al [28] found that all 5 participants in their single-session iCBT pilot moved from clinically elevated to subclinical scores on the Multidimensional Anxiety Scale for Children-Second Edition, though the very small sample precludes inferential conclusions. Birrell et al [38] found no significant between-group difference on the 7-item Generalized Anxiety Disorder 7 scale (d=–0.13; P=.41). McCashin et al [35] reported significant improvement in internalizing symptoms over time using the Child Behavior Checklist but no significant group × time interaction compared to a waitlist control condition (P=.71). In addition, a higher proportion of intervention participants moved from the clinical to the nonclinical range at postassessment compared to control (22.5% vs 0%; χ^2_1=7.53; P=.01), though this difference was not maintained at 3-month follow-up.

Stress and Emotional Distress

Studies targeting stress, emotional distress, or general psychological symptoms yielded heterogeneous results. Chen et al [40] found significant reductions in perceived stress on the DASS-21 stress subscale (d=0.54; P<.01). Nicolaidou et al [36] reported that 94.4% of children felt happy or relaxed immediately following a brief breathing exercise in the Kids' Stress Relief app, though this was assessed as an immediate postexercise self-report without a validated scale. Birrell et al [38] found no significant change in Kessler 6 scale psychological distress scores over 12 months.

Well-Being and Resilience

Several studies examined well-being, resilience, or positive mental health outcomes. The Grow It! app [43] demonstrated significant improvements in both affective and cognitive well-being across 2 large cohorts (ds=0.20-0.32), with higher-risk adolescents showing the greatest gains. Hugh-Jones et al [44] found a mean Warwick and Edinburgh Mental Well-being Scale increase of 7.06 points in the MindMate2U intervention group compared to 3.67 in the waitlist group, though the feasibility study was not powered for inferential testing. Keinonen et al [45] found no significant intervention effects on emotional and behavioral problems, health-related quality of life, self-rated mood, school satisfaction, or friendship satisfaction in their cluster RCT of Magis. However, a

XSL•FO
RenderX

significant group × time interaction was observed for loneliness ($F_{1,102}$=5.77; *P*=.02), which decreased in the intervention group while increasing in the control group. Bohr et al [39] found no significant intervention effect on resilience as measured by the shortened, 12-item version of the Child and Youth Resilience Measure according to the prespecified 2-tailed criterion; the original publication did not provide an unambiguous 2-tailed *P* value for this effect. Peuters et al [49] found no significant change in health-related quality of life (HRQoL) over time among adolescents receiving normal in-school education (χ^2_2=0.97; *P*=.61; T0-T2: *b*=0.87, 95% CI –0.88 to 2.63; *P*=.33). In contrast, pandemic-related education restrictions significantly moderated the intervention effect (χ^2_2=13.13; *P*=.001), with HRQoL declining more in the intervention than in the control group among adolescents receiving partly remote education (T0-T2: *b*=–4.10, 95% CI –6.28 to –1.93; *P*<.001).

Bullying-Related Distress and Social-Emotional Outcomes

Three studies examined outcomes related to bullying, cyberbullying, or social-emotional functioning. Williams et al [52] demonstrated significant reductions in both social and cyberbullying perpetration in the LST + Galaxia cluster RCT (social bullying perpetration: Wald χ^2_1=3.87; *P*<.049; cyberbullying perpetration: Wald χ^2_1=4.20; *P*<.04). Physical and verbal bullying perpetration did not differ significantly between groups (adjusted generalized linear model means: physical, 1.20 vs 1.32; verbal, 1.73 vs 1.69), and no significant group differences were reported for bullying victimization. Separate inferential statistics for these nonsignificant outcomes were not provided in the original publication. Chen et al [36] reported a significant reduction in cyberbullying following the GISCC intervention (*d*=0.29; *P*<.05), alongside an improvement in general self-efficacy (*d*=0.43; *P*<.01). Lepe-Salazar et al [27] reported a significant improvement in assertiveness scores (Rathus Assertiveness Schedule) using the Bernstein role-playing games in a school-based exploratory trial, conceptualized as a proxy for a bullying protective factor (treatment group: –2.78 to 1.69; *P*=.01; control group comparison via 1-way ANOVA; *P*=.01).

Self-Harm and Emotion Regulation

Two studies targeted self-harm in clinical samples. Grist et al [33] reported qualitative findings indicating that adolescents using BlueIce experienced the app as helpful in managing urges to self-harm, with several participants reporting a perceived cessation of self-harming behavior during the 12-week intervention period. Similarly, Stallard et al [51] found that 76.6% of those who used BlueIce when thinking about self-harm reported that the app had helped prevent at least one episode of self-harm, although 70.2% still reported engaging in self-harm at some point during the same period.

In addition, Bohr et al [39] found no significant reduction in depressive symptoms on the CESD-R but reported significant reductions in hopelessness (Hopelessness Scale for Children; *b*=–1.08; *P*=.02), self-blame (*b*=–0.86; *P*=.03), rumination (*b*=–0.77; *P*=.03), and catastrophizing (*b*=–0.90; *P*=.02). Other cognitive emotion regulation strategies assessed with the Cognitive Emotion Regulation Questionnaire-Short did not show significant intervention effects. These findings should be interpreted cautiously given the small pilot sample and the use of 1-tailed tests for several hypothesized effects.

Engagement, Adherence, User Experience

Engagement and adherence were reported using heterogeneous metrics across studies, including intervention uptake, usage frequency, dose, completion, compliance, and subjective indicators of engagement quality. Pramana et al [50] found significantly higher usage frequency for the gamified SmartCAT 2.0 compared with its nongamified predecessor (median 68 vs 37 uses; *P*<.01). Fleming et al [29] observed substantial attrition across successive levels of SPARX use: of 21,320 registered users, 63.6% (n=13,564) initiated the intervention; among those who started SPARX, only 7.4% (n=997) completed at least 4 modules and 3.1% (n=416) completed all 7 modules. Dietvorst et al [31] reported experience-sampling compliance rates of 16.7%-30.9%, while challenge completion rates in the longitudinal Grow It! study [43] ranged from 25% to 33%. Birrell et al [38] similarly reported limited uptake and dose, with 48.9% of participants downloading the app and a mean of 1.57 modules opened. In contrast, Moffett et al [25] reported relatively high intervention adherence for Kuamsha, with 80.2% of participants completing at least 4 of 6 app modules and 63.5% completing all 6 recommended modules. In addition, study retention at the 11-week primary end point was 100%, with all randomized participants completing the primary end point assessment; this reflects assessment completion rather than intervention adherence.

Indicators of engagement quality, acceptability, and usability were also heterogeneous. Dietvorst et al [31] reported overall app satisfaction ratings of 7.1-7.2 out of 10, whereas Birrell et al [38] reported app ratings of 3.38-3.48 out of 5. Moffett et al [25] reported significantly higher acceptability (Acceptability of Intervention Measure: mean 4.22, SD 0.71 vs mean 3.98, SD 0.75; *P*=.02) and appropriateness (Intervention Appropriateness Measure: mean 4.22, SD 0.64 vs mean 3.89, SD 0.80; *P*=.002) for Kuamsha compared with the active control app, whereas perceived feasibility did not differ significantly between groups (Feasibility of Intervention Measure: mean 4.09, SD 0.63 vs mean 3.96, SD 0.69; *P*=.16). Nicolaidou et al [36] reported a mean System Usability Scale score of 73.55 (grade B; 70th percentile) for the Kids' Stress Relief app. Miranda et al [47] reported favorable usability ratings for the BALANCE app among healthy controls, with most participants describing it as user-friendly and helpful. McCashin et al [35] found that 88% of children and 83% of parents considered Pesky gNATs potentially helpful for future use, while 90% of children and 88% of parents would recommend it. Mittmann et al [48] reported a mean Game User Experience Satisfaction Scale-18 usability score of 33.64 out of 45 for the LINA augmented reality game, with enjoyment and social connectivity among the more favorably rated aspects.

Qualitative findings provided additional information on the quality and context of engagement. In Moffett et al [25], participants identified in-app points and games, relatable stories

XSL•FO
RenderX

and characters, and peer-mentor support as features facilitating engagement, whereas barriers included excessive reading demands, technical difficulties, and challenges coordinating calls with peer mentors. More broadly, qualitative findings across studies highlighted perceived relevance, narrative engagement, and personalization as potential facilitators of intervention use, whereas reported barriers included poor app stability, technical difficulties, perceived mismatches between game aesthetics and the target age group, and, in resource-limited settings, constraints related to connectivity and device access [25,26]. Overall, the substantial variation in how engagement and adherence were operationalized and reported limited direct comparisons across interventions and precluded conclusions about whether particular engagement patterns were associated with mental health outcomes.

Follow-Up Outcomes

Longer-term effects were assessed in only a small subset of studies, with considerable variation in follow-up duration. Overall, evidence for sustained improvements in mental health outcomes was limited and mixed.

Lincke et al [46] reported significant reductions in depressive symptoms during the e-MICHI intervention, which remained stable at the 3-month follow-up across self-reported and clinician-rated measures. Birrell et al [38] assessed outcomes at 6 and 12 months and found a small reduction in depressive symptoms in the intervention group relative to the active control over the 12-month period (d=–0.31), although this effect did not reach statistical significance. No significant intervention effects were observed for anxiety or psychological distress. McCashin et al [35] found no significant between-group differences in internalizing symptoms or secondary mental health outcomes at the 3-month follow-up; moreover, the advantage observed for clinically meaningful change immediately after the intervention was no longer evident at follow-up. Moffett et al [25] similarly found no significant between-group difference in depressive symptoms at the 24-week follow-up (adjusted mean difference=–0.11, 95% CI –1.70 to 1.48; P=.89). For REThink, David et al [41] reported postintervention outcomes but did not report the planned 6-month follow-up results. The subsequent publication [42] was a secondary analysis of in-game performance rather than an analysis of longer-term intervention effects. Therefore, no 6-month REThink outcomes were available for extraction from the included publications.

Taken together, the available follow-up evidence provides limited indication that improvements may persist beyond the immediate intervention period. However, the small number of studies with longer-term assessments and the heterogeneity of follow-up periods preclude firm conclusions regarding the durability of intervention effects.

Meta-Analyses

Included Studies and Effect Sizes

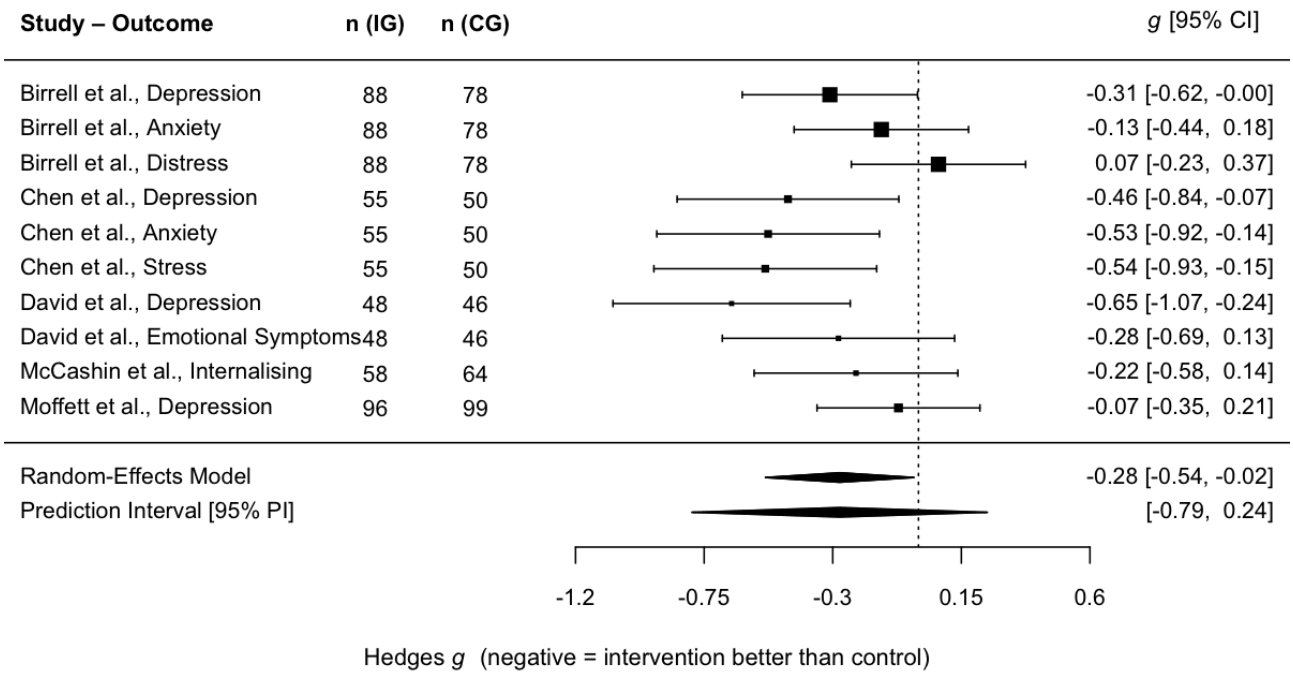
Of the 27 studies included in the systematic review, 17 were not eligible for the meta-analysis because they used uncontrolled, observational, qualitative, or other designs without an eligible comparator condition. Of the remaining 10 studies with controlled designs, 5 were included in the meta-analysis [25,35,38,40,41], contributing a total of 10 effect sizes across validated mental health outcomes. Five controlled studies were excluded from the quantitative synthesis for the following reasons: Bohr et al [39] did not report sufficient statistics to calculate effect sizes; Keinonen et al [45] did not assess a clinical mental health outcome as a primary measure; Lepe-Salazar et al [27] assessed assertiveness rather than mental health outcomes; Peuters et al [49] primarily targeted lifestyle behaviors; and Williams et al [52] assessed bullying behavior only, without mental health symptom outcomes. Sample sizes across included studies ranged from N=105 [40] to N=195 [25]. Three studies used waitlist or no-intervention control conditions [35,40,41], and 2 used active control conditions [25,38]. Follow-up periods ranged from 4 weeks to 12 months. Effect sizes at the outcome level ranged from g=–0.65 for depressive mood [41] to g=+0.07 for psychological distress [38].

Overall Effect

Across all 5 studies and 10 outcomes, the 3-level random-effects model yielded a statistically significant small overall effect in favor of gamified digital interventions, g=–0.28, 95% CI –0.54 to –0.02; P=.04. Individual study–outcome effect sizes and the pooled 3-level estimate are presented in Figure 2 [25,35,38,40,41]. Negative values indicate greater symptom reduction in the intervention group relative to the control group. Using conventional benchmarks for standardized mean differences, the pooled estimate corresponds to a small effect [53].

XSL•FO
RenderX

Figure 2. Forest plot of effect sizes for gamified digital mental health interventions across controlled studies [25,35,38,40,41]. The forest plot displays Hedges g effect sizes with 95% CIs for each study-outcome combination included in the meta-analysis (K=5 independent studies; k=10 outcomes). The pooled effect is based on the primary 3-level random-effects model accounting for multiple outcomes nested within studies (g=–0.28, 95% CI –0.54 to –0.02; *P*=.04). The 95% prediction interval ranged from –0.79 to 0.24. Negative values indicate greater symptom reduction in the intervention group (IG) relative to the control group (CG).



The 95% prediction interval ranged from –0.79 to 0.24 and included zero, indicating uncertainty regarding the effect that may be expected in a future comparable setting. Given the small number of independent studies (K=5), the pooled estimate should therefore be interpreted cautiously as an exploratory quantitative summary rather than as definitive evidence of efficacy.

The estimated between-study variance was σ^2=0.03, whereas the estimated between-outcome variance within studies was negligible (σ^2=0.00). The corresponding pseudo-I^2 estimates were 45.3% at the between-study level and 0.0% at the within-study level. The test for heterogeneity was Q_9=15.23 (*P*=.08). Given the limited number of independent studies, these variance-component estimates should be interpreted cautiously.

As a sensitivity analysis, a conventional random-effects model using the Hartung-Knapp adjustment yielded a comparable pooled estimate (*g*=–0.28, 95% CI –0.45 to –0.12; *P*=.004). Because this model does not account for the dependence of multiple outcomes within studies, the 3-level model was retained as the primary analysis.

Moderator and Subgroup Analyses

Moderator analyses were planned to examine gamification elements, clinical status, and intervention type as potential sources of heterogeneity. However, with only 5 independent studies available, formal moderator analyses were not conducted, as the recommended minimum number of studies for reliable moderator analyses (typically ≥10 studies per moderator) was not met [24]. Meta-regression was therefore considered inappropriate given the limited statistical power and risk of unstable or spurious estimates. Potential differences according to intervention characteristics could therefore only be explored descriptively and should be regarded as hypothesis-generating.

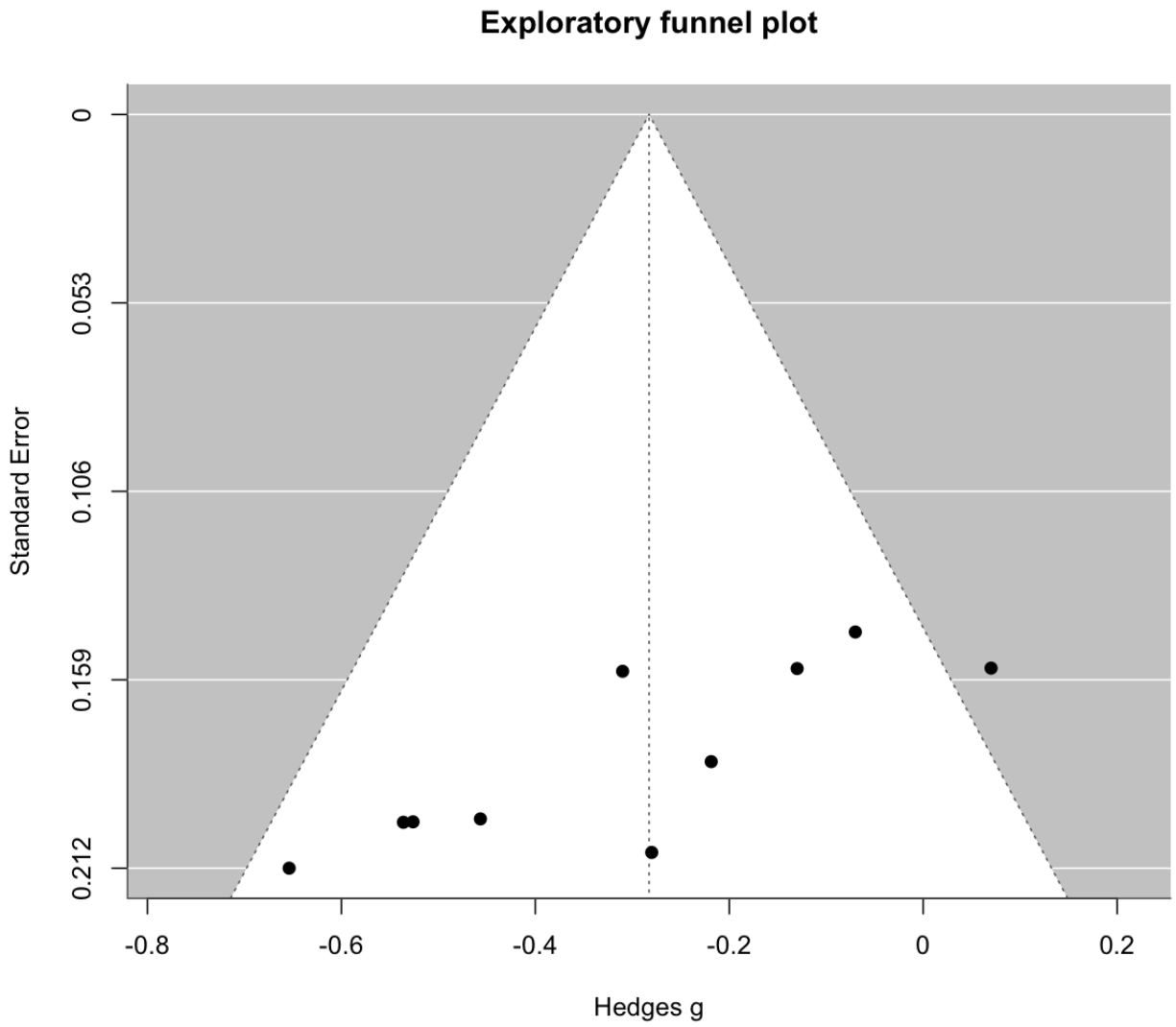
Publication Bias

Assessment of publication bias was substantially limited by the small number of independent studies (K=5) and by the inclusion of multiple dependent effect sizes within several studies. An exploratory Egger-type regression produced a statistically significant result (t_8=–4.00; *P*=.004), with a limit estimate of *b*=1.03 (95% CI 0.28-1.79). However, the standard Egger test assumes independent effect sizes, an assumption violated by the present dataset. Moreover, funnel-plot asymmetry tests are unreliable with so few independent studies. This result should therefore not be interpreted as evidence of publication bias [54].

Exploratory visual inspection of the funnel plot, see Figure 3, suggested some asymmetry. However, given the small number of independent studies, this pattern may reflect sampling variability or genuine between-study heterogeneity rather than selective publication. Publication bias could therefore neither be reliably confirmed nor excluded.

XSL•FO
RenderX

Figure 3. Exploratory funnel plot of study-outcome effect sizes. Each point represents one outcome-level effect size (Hedges g) plotted against its SE. Because only 5 independent studies contributed to the meta-analysis and several studies contributed multiple dependent effect sizes, the funnel plot and Egger-type regression should be interpreted as exploratory only and cannot reliably confirm or exclude publication bias.



Discussion

Principal Findings

This systematic review synthesized evidence from 27 studies examining gamification elements, internalizing mental health outcomes, engagement, adherence, and user experience in gamified digital mental health interventions for children and adolescents. Feedback, challenges, and progress tracking were the most frequently implemented gamification elements, whereas social interaction and competition were uncommon. The meta-analysis of 5 eligible controlled studies yielded a small effect in favor of gamified interventions on internalizing mental health outcomes (g=–0.28, 95% CI –0.54 to –0.02; P=.04). Engagement and adherence varied substantially across studies, and the available evidence did not allow conclusions regarding whether specific gamification elements were associated with more favorable clinical or engagement outcomes. This finding is consistent with prior reviews by David et al [18] and Bryant et al [19], which reported similarly modest effects for serious games and gamified digital health interventions in youth populations. However, the 95% prediction interval (–0.79 to 0.24) included zero, indicating substantial uncertainty regarding the effect that may be expected in future comparable settings. Given the small number of eligible studies and the heterogeneity in intervention formats, populations, and outcome measures, the pooled estimate should be interpreted as suggestive evidence of a small average benefit rather than as a precise or broadly generalizable estimate of efficacy.

The narrative synthesis similarly revealed mixed findings. Significant improvements in depression and anxiety were observed in several controlled and uncontrolled studies, including interventions delivered to clinical, at-risk, and nonclinical populations. At the same time, several controlled studies reported small or nonsignificant effects. Although some of the more favorable findings occurred in interventions grounded in cognitive-behavioral approaches or involving therapeutic guidance, the available evidence was insufficient to determine whether clinical status, therapeutic framework,

XSL•FO
RenderX

guidance, or other intervention characteristics systematically moderated outcomes. This uncertainty was also evident in the analysis of gamification characteristics. Interventions with favorable outcomes frequently incorporated multiple elements, including narrative, personalization, rewards, feedback, and progression structures. However, these features were also present in interventions with null or mixed findings. No individual gamification element consistently distinguished interventions with favorable outcomes from those with less favorable findings. Gamification should therefore not be regarded as an isolated active ingredient, and its potential contribution is likely intertwined with therapeutic content, delivery context, and patterns of intervention use.

Several findings nevertheless suggest that intervention exposure and engagement warrant closer investigation. David et al [42] found that higher in-game performance predicted subsequent symptom improvement, while Fleming et al [29] observed larger within-person reductions in depressive symptoms among participants completing more SPARX modules. These associations cannot establish a dose-response effect, as they may reflect self-selection, baseline differences, or other confounding factors. However, they highlight the importance of distinguishing the potential efficacy of an intervention from the extent to which users initiate, engage with, and complete it.

Taken together, the findings suggest that gamified digital mental health interventions may offer modest benefits for youth mental health, while substantial uncertainty remains regarding which intervention components, populations, and delivery contexts are most likely to benefit. The conceptual overlap between many gamification elements and established behavior change techniques may represent a useful avenue for future research. For example, progress tracking, feedback, rewards, and structured challenges overlap conceptually with techniques represented in established behavior change taxonomies, including self-monitoring, feedback on behavior, reinforcement, and graded tasks [55]. However, the gamification categories used in the present review were not retrospectively coded according to the Behavior Change Technique Taxonomy (v1), and direct equivalence between gamification elements and individual behavior change techniques should therefore not be assumed. Future reviews could systematically map gamification components onto established behavior change taxonomies to improve comparability across interventions and help disentangle the contribution of gamification from that of the underlying therapeutic content.

This review systematically characterized gamification elements across the included studies. Feedback systems, challenges, and progress tracking were most commonly implemented, whereas social interaction and competitive features were comparatively rare. Interventions varied considerably in both the number and combination of gamification elements used, and no individual element consistently distinguished interventions with favorable outcomes from those with null or mixed findings.

The frequent co-occurrence of gamification elements with therapeutic components makes it difficult to isolate their specific contribution. The descriptive patterns observed in this review cannot establish whether narrative integration, rewards, personalization, or other individual features contributed to intervention effectiveness. Instead, observed outcomes may reflect combinations of gamification features, therapeutic content, population characteristics, intervention intensity, guidance, and implementation factors. Moreover, several interventions lie on a continuum between gamified interventions and serious games, further blurring conceptual boundaries. Component-level analyses, dismantling studies, and factorial designs are needed to determine whether particular gamification features contribute independently or interactively to mental health outcomes.

Developmental considerations may also influence how gamification features are experienced and used. Different developmental stages are characterized by changing needs for structure, autonomy, identity exploration, and social interaction, suggesting that the relevance of particular game elements may vary across childhood and adolescence. Qualitative findings in the included studies indicated that users valued customizable features such as avatars, colors, and music, as well as mood and trigger tracking. These findings suggest that personalization and emotionally responsive features warrant further investigation, although the present evidence does not establish that they improve clinical outcomes.

From a developmental perspective, gamification may also be understood in relation to the broader role of play in childhood and adolescence. Play provides a developmentally familiar context for exploration, learning, social interaction, and emotion regulation, and game-based environments may therefore provide opportunities to practice therapeutic skills in engaging and comparatively low-stakes contexts. Features such as narrative, role-play, challenges, and feedback may support experiential learning and repeated practice. However, the mechanisms through which play-related processes contribute to clinical outcomes in digital mental health interventions remain insufficiently investigated, and the present review cannot determine whether these processes mediate intervention effects.

Progression systems and gradually increasing challenges may also be conceptually related to flow theory, according to which engagement is facilitated when perceived challenges are appropriately matched to an individual's skills. From this perspective, levels, adaptive difficulty, and structured progression may support sustained engagement by maintaining a balance between challenge and competence. However, the included studies rarely assessed flow directly, and this potential mechanism therefore remains theoretical rather than empirically established within the present evidence base.

The relative scarcity of social and competitive features may reflect concerns about social comparison or the appropriateness of competitive mechanics in vulnerable populations, although the reasons for these design choices were rarely examined directly. At the same time, carefully structured peer-based features may represent an underexplored avenue for adolescent interventions, particularly where social connectedness is therapeutically relevant. Overall, the current evidence does not indicate that a greater number of gamification elements necessarily produces better outcomes. Rather, future research should examine whether specific features are developmentally

XSL•FO
RenderX

appropriate, therapeutically congruent, and capable of supporting meaningful and sustained engagement over time.

Engagement and adherence emerged as a recurring and critical limiting factor across the evidence base. Mental health apps are generally poorly used in real-world settings, with high dropout rates observed even in nongamified interventions [56]. This review suggests that gamification alone may not reliably overcome these implementation challenges. For example, Fleming et al [29] reported that only 3.1% of the 13,564 adolescents who started SPARX completed all 7 modules—just 2.0% of the original 21,320 registrants.

The interpretation of dropout was further complicated by inconsistent reporting across studies. For the purposes of interpreting the present evidence, study attrition refers to noncompletion of scheduled study assessments, whereas intervention abandonment refers to discontinuation or noncompletion of the intervention itself, even when participants remained available for follow-up assessment. These distinct concepts were not consistently differentiated across the included studies, limiting direct comparisons of dropout rates. A further challenge is the substantial variability in usage when engagement is left largely to user discretion. Lockwood et al [34], for instance, reported a mean of 11 sessions over 15 days (SD 9.41), indicating pronounced interindividual differences in engagement that were not systematically controlled for in most studies. This variability complicates interpretation of outcome data, as it remains unclear whether nonresponse reflects insufficient exposure or limited intervention efficacy.

While gamification is primarily designed to enhance engagement and motivation, its potential risks in youth mental health interventions have received comparatively little empirical attention. A key concern relates to the potential for problematic or excessive use associated with certain gamification mechanics. Features such as variable reward schedules, streak systems, achievement badges, and progressive unlocking are deliberately designed to reinforce repeated engagement and may draw on mechanisms similar to those implicated in problematic use of video games and social media [57,58]. In children and adolescents with internalizing disorders, such mechanisms may blur the boundary between therapeutic engagement and compulsive use, particularly given the always-available nature of app-based interventions and the limited self-regulatory capacities in younger users. Notably, none of the included studies systematically assessed problematic or excessive use, and only a minority implemented explicit usage limits (eg, 34, 40 minutes per day). This gap is concerning and warrants attention in future research and clinical implementation guidelines.

A second and closely related issue is the question of optimal gamification "dose," that is, how much gamification is necessary to sustain therapeutic engagement, and at what point does it become counterproductive or even harmful? This question remains almost entirely unaddressed in the empirical literature. This review did not identify a consistent relationship between gamification intensity and more favorable outcomes, and no included study directly compared varying levels of gamification intensity within the same intervention framework. Theoretically, excessive gamification may undermine intrinsic motivation through overjustification effects [59,60], shift the user's focus from therapeutic content to game mechanics, or create unrealistic expectations about the nature of mental health recovery. Conversely, insufficient gamification may fail to differentiate an app from a conventional psychoeducational tool and provide little motivational added value. The optimal balance between motivational scaffolding and therapeutic focus has not been empirically established, and no included study directly compared varying levels of gamification intensity within the same intervention framework. These issues may be particularly relevant in clinical populations, where avoidance-driven usage patterns could shape engagement with gamification features.

Further, in the context of already elevated media consumption among children and adolescents [61,62], the use of gamified mental health apps may contribute to cumulative digital exposure, with potential implications for sleep, attention, and overall well-being [63,64]. These potential trade-offs have not been systematically examined in the included studies and remain an open question for future research.

These considerations may be particularly important for children and adolescents, whose ability to independently regulate digital media use and evaluate potential risks varies with developmental stage. Gamified mental health interventions for minors should therefore not be conceptualized solely as autonomous self-management tools [65,66]. Depending on age, symptom severity, and intervention content, appropriate caregiver involvement and professional guidance may provide important safeguards, particularly when interventions are used outside structured clinical settings. Emerging integration of generative AI and conversational agents into digital mental health tools additionally raises questions regarding inappropriate or inaccurate responses, privacy, and reliance on automated support. These issues warrant specific evaluation before such features are implemented in gamified interventions for young users.

Altogether, these considerations highlight the importance of examining which forms and intensities of gamification are most likely to support meaningful, sustainable, and clinically relevant engagement, rather than assuming that more gamification necessarily leads to better outcomes.

Limitations of This Review

This review has several limitations related to both the underlying evidence base and the review methodology.

First, the meta-analysis included only 5 controlled studies, limiting statistical power and precluding moderator analyses. The pooled effect should therefore be interpreted as an exploratory estimate rather than a definitive measure of efficacy. Second, substantial heterogeneity across study designs, populations, interventions, outcomes, and follow-up periods limited comparability and synthesis, reflecting the early and fragmented state of the field. Third, methodological quality of outcome assessment was variable. Most studies relied on self-report measures, often without clinician-rated confirmation, and several used nonvalidated or study-specific instruments. In addition, none of the included RCTs reported blinded outcome

XSL•FO
RenderX

assessment, increasing risk of bias. Fourth, control conditions were frequently weak or inconsistent. Many studies used waitlist controls, which may inflate effect estimates due to expectancy effects and natural symptom fluctuation, while active controls varied considerably in content and intensity. Fifth, sample and contextual factors further limit generalizability. Many studies included nonclinical populations with low baseline symptom severity, while clinical samples were often small or confounded by concurrent treatment. In addition, several studies were conducted during or shortly after the COVID-19 pandemic, which may have independently influenced mental health trajectories and outcomes. Peuters et al [49] explicitly reported that pandemic-related measures moderated intervention effects, with beneficial effects observed only for adolescents who continued in-school education, and negative effects for those in partly remote settings. Sixth, several studies, particularly single-session or very brief evaluations [28,30,36], had extremely limited exposure durations that are unlikely to produce lasting clinical change. The inclusion of such studies in a review of therapeutic effectiveness, while appropriate for capturing feasibility and usability evidence, underscores the need to distinguish between proof-of-concept findings and evidence of efficacy. Seventh, methodological reporting was inconsistent, with insufficient detail on intervention components, samples, and analytic procedures. Data protection and privacy issues were rarely addressed. Given that these interventions collect sensitive mental health data from minors, including mood logs, symptom scores, and behavioral patterns, the absence of systematic reporting on data security, password protection, and compliance with applicable regulations such as the General Data Protection Regulation represents a notable gap that future research and clinical implementation should address explicitly.

Important developmental and contextual evidence gaps also remain. Few studies included children under 10 years of age [30,34-36], limiting conclusions for younger children with distinct developmental needs regarding literacy, attention, and emotion regulation. Evidence from LMICs was similarly limited [25-27]. In these settings, barriers related to device access, connectivity, infrastructure, and cultural adaptation may constrain implementation [25,26]. The potential scalability of digital interventions should therefore be distinguished from their practical accessibility, and clinical efficacy should not be assumed to translate directly into population-level feasibility. The review methodology introduces additional limitations. Study selection was conducted primarily by one reviewer, with a second reviewer consulted when eligibility was uncertain, while quality assessments were initially conducted by one reviewer and subsequently cross-checked by a second. Data extraction and gamification coding were not independently duplicated. These procedures may have increased the risk of selection, extraction, assessment, and classification bias. Furthermore, the iteratively developed gamification coding framework should be regarded as descriptive rather than as a validated taxonomy. The restriction to English-language publications from 2018 onward, potential omission of recently published studies, and limited ability to assess publication bias represent additional limitations. Finally, the available evidence lacked sufficient clinical granularity to determine whether gamified interventions are differentially effective for specific symptom dimensions, such as cognitive vs somatic manifestations of anxiety or depression. Similarly, the evidence does not permit conclusions regarding appropriate levels of care, including outpatient, intensive outpatient, residential, or inpatient settings. These findings should therefore not be interpreted as supporting specific clinical placement recommendations. Future studies should examine differential effects across symptom profiles, clinical severity, and treatment settings.

Overall, current evidence should be interpreted cautiously. Future research should prioritize adequately powered RCTs with active control conditions, standardized outcome measures, improved reporting, and greater methodological transparency.

Despite the methodological limitations of the current evidence base, several potentially relevant patterns emerged from the narrative synthesis. Favorable outcomes were observed across interventions with different combinations of gamification features, therapeutic frameworks, and delivery formats, but no individual gamification element or level of gamification intensity consistently distinguished interventions with positive outcomes from those with null or mixed findings. These observations highlight the need to examine how gamification interacts with therapeutic content, population characteristics, guidance, and implementation context in adequately powered studies.

For intervention design, gamification should be integrated into a coherent evidence-based therapeutic framework rather than applied as a superficial layer of motivational features. Although CBT-based approaches currently dominate the evidence base, these findings do not establish that CBT or particular combinations of gamification elements are inherently more effective. Developers may benefit from incorporating built-in evaluation mechanisms that capture not only usage metrics but also exposure to and fidelity with therapeutic components, allowing for a more meaningful assessment of real-world intervention engagement.

For research, future studies should prioritize adequately powered RCTs with active control conditions, longitudinal follow-up assessments, validated clinical outcome measures, and preregistered analysis plans. Dismantling designs that isolate the effects of specific gamification elements would substantially advance mechanistic understanding [67,68]. Future research should also examine whether gamification can effectively support therapeutic approaches beyond CBT, including skills-based interventions such as DBT, as the present evidence is insufficient to determine whether particular therapeutic modalities are more suitable for gamification than others. Studies should additionally report concurrent treatments, assess intervention exposure and engagement systematically, and address data protection considerations explicitly. Particular attention should be paid to younger age groups and low-resource settings, where the evidence base remains limited.

The potential scalability of gamified digital mental health interventions should be interpreted cautiously in LMIC contexts. Clinical effectiveness and practical accessibility should be considered separately, as barriers related to devices, connectivity, infrastructure, technical support, and digital literacy may restrict implementation. Future research should

XSL•FO
RenderX

therefore evaluate context-specific implementation requirements and barriers before assuming scalability in low-resource settings.

For clinical practice, this evidence is insufficient to support the routine recommendation of specific gamified apps as stand-alone treatments or to determine their appropriate placement within specific levels of care. Gamified interventions may nevertheless have potential as adjunctive or low-intensity tools, but their role alongside established treatments requires further evaluation. Clinicians should be aware of the substantial variability in engagement with digital interventions and consider whether the intervention format, therapeutic content, and level of guidance are appropriate for the individual patient and clinical context. When evaluating gamified digital mental health interventions, clinicians and families may consider whether the intervention is grounded in an evidence-based therapeutic framework, provides developmentally appropriate content, offers transparent information on data protection and safety, and includes an appropriate level of human support or guidance. Services considering implementation should additionally account for technical support, staff responsibilities, data-security procedures, and pathways for responding to symptom deterioration or safety concerns. Particular caution is warranted for young people with severe or complex mental health conditions, for whom the effectiveness, safety, and appropriate level of care of gamified interventions remain insufficiently established.

Conclusion

This systematic review and meta-analysis provides suggestive evidence that gamified digital mental health interventions may be associated with small improvements in internalizing outcomes, particularly depression and anxiety, in children and adolescents. However, the limited and heterogeneous evidence base and the uncertainty of effects across settings preclude firm conclusions regarding efficacy. More broadly, the findings suggest that gamification should not be viewed as a therapeutic mechanism in itself, but rather as a design strategy whose value may depend on how specific game elements are integrated with evidence-based therapeutic content, developmental needs, appropriate guidance, and the broader implementation context. Gamification alone is unlikely to ensure sustained engagement or therapeutic effectiveness, and potential trade-offs, including problematic or excessive use and the still unclear optimal level of gamification, require greater attention.

Future development and evaluation should therefore move beyond asking whether gamified interventions are effective in general and instead determine which gamification elements support clinical outcomes and meaningful engagement, for whom they are appropriate, and under which conditions. Establishing this evidence will be essential for translating the potential of gamification into developmentally appropriate, clinically meaningful, and responsibly implemented digital mental health interventions for children and adolescents.

Funding

The authors declared no financial support was received for this work.

Data Availability

This study is a systematic review and meta-analysis based on previously published data. All data extracted and analyzed in this study are derived from published articles cited in the manuscript. No new primary datasets were generated. Extracted data and analysis materials are available from the corresponding author upon reasonable request.

Generative AI Use

Generative AI tools were used to support selected stages of manuscript preparation. ChatGPT (GPT-5.5; OpenAI) was used for language editing, grammar, structural refinement, improving the overall clarity and flow of the manuscript, and assisting with the optimization of statistical analysis code. Claude Sonnet 4.6 (Anthropic) was used to assist with the organization of extracted information during data extraction. All extracted information, statistical code, and AI-assisted outputs were manually checked and verified by the authors. AI tools were not used to make study eligibility decisions or independently determine the statistical analyses or interpretation of findings. All final coding, analytical, and interpretive decisions were made and verified by the authors, who retain full responsibility for the accuracy and content of the manuscript.

Authors' Contributions

JFB contributed to conceptualization, methodology, investigation, data curation, formal analysis, software, validation, visualization, project administration, and writing – original draft. JH contributed to supervision and writing – review and editing. EM contributed to conceptualization and writing – review and editing. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.
[DOCX File , 269 KB-Multimedia Appendix 1]

XSL•FO
RenderX

Multimedia Appendix 2

Search strings.
[DOCX File , 18 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Quality assessment.
[DOCX File , 60 KB-Multimedia Appendix 3]

References

1. Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, et al. Child and adolescent mental health worldwide: evidence for action. Lancet. 2011;378(9801):1515-1525. [doi: 10.1016/s0140-6736(11)60827-1]
2. Sappenfield O, Alberto C, Minnaert J, Donney J, Lebrun-Harris L, Ghandour R. National survey of children's health data briefs. NIH. URL: https://www.ncbi.nlm.nih.gov/books/NBK603024/ [accessed 2026-04-15]
3. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. JAMA Pediatr. 2021;175(11):1142-1150. [FREE Full text] [doi: 10.1001/jamapediatrics.2021.2482] [Medline: 34369987]
4. Madigan S, Racine N, Vaillancourt T, Korczak DJ, Hewitt JMA, Pador P, et al. Changes in depression and anxiety among children and adolescents from before to during the COVID-19 pandemic. JAMA Pediatr. 2023;177(6):567-581. [doi: 10.1001/jamapediatrics.2023.0846]
5. Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, Benjet C, Bruffaerts R, et al. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. Psychol Med. 2017;48(9):1560-1571. [doi: 10.1017/s0033291717003336]
6. McGorry PD, Mei C. Early intervention in youth mental health: progress and future directions. Evid Based Ment Health. 2018;21(4):182-184. [FREE Full text] [doi: 10.1136/ebmental-2018-300060] [Medline: 30352884]
7. Parry TS. The effectiveness of early intervention: a critical review. J Paediatr Child Health. 1992;28(5):343-346. [doi: 10.1111/j.1440-1754.1992.tb02688.x] [Medline: 1382490]
8. Löchner J, Carlbring P, Schuller B, Torous J, Sander L. Digital interventions in mental health: an overview and future perspectives. Internet Interv. 2025;40:100824. [FREE Full text] [doi: 10.1016/j.invent.2025.100824] [Medline: 40330743]
9. Torous J, Linardon J, Goldberg SB, Sun S, Bell I, Nicholas J, et al. The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. World Psychiatry. 2025;24(2):156-174. [FREE Full text] [doi: 10.1002/wps.21299] [Medline: 40371757]
10. Potts C, Kealy C, McNulty JM, Madrid-Cagigal A, Wilson T, Mulvenna MD, et al. Digital mental health interventions for young people aged 16-25 years: scoping review. J Med Internet Res. 2025;27:e72892. [FREE Full text] [doi: 10.2196/72892] [Medline: 40344661]
11. Bas-Sarmiento P, Julián-López C, Fernández-Gutiérrez M, Poza-Méndez M, Marín-Paz A. Gamified eHealth interventions for health promotion and disease prevention in children and adolescents: a scoping review. Humanit Soc Sci Commun. 2025;12(1):397. [doi: 10.1057/s41599-025-04670-w]
12. Damaševičius R, Maskeliūnas R, Blažauskas T. Serious games and gamification in healthcare: a meta-review. Information. 2023;14(2):105. [doi: 10.3390/info14020105]
13. Dietvorst E, Legerstee J, Vreeker A, Koval S, Mens M, Keijsers L. A serious game smartphone application (Grow it!) that promotes wellbeing and resilience in Dutch adolescents during the COVID-19 pandemic. Bipolar Disord. 2021;23:60-60. [doi: 10.2196/preprints.29832]
14. Smith KA, Ward T, Lambe S, Ostinelli EG, Blease C, Gant T, et al. Engagement and attrition in digital mental health: current challenges and potential solutions. NPJ Digit Med. 2025;8(1):398. [FREE Full text] [doi: 10.1038/s41746-025-01778-w] [Medline: 40604240]
15. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Association for Computing Machinery. 2011. Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; 2011 September 28-30; New York, NY, USA. [doi: 10.1145/2181037.2181040]
16. Luarn P, Chen C, Chiu Y. Enhancing intrinsic learning motivation through gamification: a self-determination theory perspective. Int J Inf Learn Technol. 2023;40(5):413-424. [doi: 10.1108/ijilt-07-2022-0145]
17. Lukka L, Palva JM. The development of game-based digital mental health interventions: bridging the paradigms of health care and entertainment. JMIR Serious Games. 2023;11:e42173. [FREE Full text] [doi: 10.2196/42173] [Medline: 37665624]
18. David OA, Costescu C, Cardos R, Mogoaşe C. How effective are serious games for promoting mental health and health behavioral change in children and adolescents? A systematic review and meta-analysis. Child Youth Care Forum. 2020;49(6):817-838. [doi: 10.1007/s10566-020-09566-1]

XSL•FO
RenderX

19. Bryant BR, Sisk MR, McGuire JF. Efficacy of gamified digital mental health interventions for pediatric mental health conditions: a systematic review and meta-analysis. JAMA Pediatr. 2024;178(11):1136-1146. [doi: 10.1001/jamapediatrics.2024.3139] [Medline: 39312259]
20. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. J Clin Epidemiol. 2021;134:178-189. [FREE Full text] [doi: 10.1016/j.jclinepi.2021.03.001] [Medline: 33789819]
21. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. Educ Inf. 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
22. The R project for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org/ [accessed 2026-11-03]
23. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-560. [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
24. Borenstein M, Hedges L, Higgins J, Rothstein H. Introduction to Meta-Analysis. Hoboken, New Jersey. John Wiley & Sons; 2009.
25. Moffett B, Pozuelo JR, Musenge E, Makhanya Z, O'Mahen HA, Craske MG, et al. A behavioral activation digital intervention incorporating gamification and peer support for adolescent depression in rural South Africa: a pilot randomized controlled trial (the DoBAt Study). JAACAP Open. 2025;3(4):1270-1283. [FREE Full text] [doi: 10.1016/j.jaacop.2025.06.009] [Medline: 41367987]
26. R Pozuelo J, Nabulumba C, Sikoti D, Davis M, Gumikiriza-Onoria JL, Kinyanda E, et al. A narrative-gamified mental health app (Kuamsha) for adolescents in Uganda: mixed methods feasibility and acceptability study. JMIR Serious Games. 2024;12:e59381. [FREE Full text] [doi: 10.2196/59381] [Medline: 39700489]
27. Lepe-Salazar F, Mejía-Romero F, Benicio-Rodríguez D, Hernández-Reyes A, Nakajima T, Salgado-Torres S. Game-based promotion of assertiveness to mitigate the effects of bullying in high school students: development and evaluation study. JMIR Serious Games. 2024;12:e58452. [FREE Full text] [doi: 10.2196/58452] [Medline: 39718822]
28. Pera MC, Poli C, Gnazzo M, Baldini V, Delsante L, Pacchioni M, et al. Design and usability testing of a novel internet-delivered cognitive behavioral therapy (iCBT) software platform for children with anxiety. Children (Basel). 2025;12(11):1535. [FREE Full text] [doi: 10.3390/children12111535] [Medline: 41300652]
29. Fleming T, Lucassen M, Frampton C, Parag V, Bullen C, Merry S, et al. Nationwide implementation of unguided cognitive behavioral therapy for adolescent depression: observational study of SPARX. J Med Internet Res. 2025;27:e66047-e66047. [doi: 10.2196/66047]
30. Nicolaidou I, Tozzi F, Antoniades A. A gamified app on emotion recognition and anger management for pre-school children. Int J Child Comput Interact. 2022;31:100449. [doi: 10.1016/j.ijcci.2021.100449]
31. Dietvorst E, Aukes MA, Legerstee JS, Vreeker A, Hrehovcsik MM, Keijsers L, et al. A smartphone serious game for adolescents (Grow It! App): development, feasibility, and acceptance study. JMIR Form Res. 2022;6(3):e29832. [doi: 10.2196/29832]
32. Pine R, Mbinta J, Te Morenga L, Fleming T. A novel casual video game with simple mental health and well-being concepts (Match Emoji): mixed methods feasibility study. JMIR Serious Games. 2024;12:e46697. [FREE Full text] [doi: 10.2196/46697] [Medline: 38345848]
33. Grist R, Porter J, Stallard P. Acceptability, use, and safety of a mobile phone app (BlueIce) for young people who self-harm: qualitative study of service users' experience. JMIR Ment Health. 2018;5(1):e16. [FREE Full text] [doi: 10.2196/mental.8779] [Medline: 29475823]
34. Lockwood J, Williams L, Martin JL, Rathee M, Hill C. Effectiveness, user engagement and experience, and safety of a mobile app (Lumi Nova) delivering exposure-based cognitive behavioral therapy strategies to manage anxiety in children via immersive gaming technology: preliminary evaluation study. JMIR Ment Health. 2022;9(1):e29008. [doi: 10.2196/29008] [Medline: 35072644]
35. McCashin D, Coyle D, O'Reilly G. Pesky gNATs for children experiencing low mood and anxiety? A pragmatic randomised controlled trial of technology-assisted CBT in primary care. Internet Interv. 2022;27:100489. [FREE Full text] [doi: 10.1016/j.invent.2021.100489] [Medline: 35024346]
36. Nicolaidou I, Tozzi F, Kindynis P, Panayiotou M, Antoniades A. Development and usability of a gamified app to help children manage stress: an evaluation study. Ital J Educ Technol. 2019;27(2):105-120. [doi: 10.17471/2499-4324/1050]
37. Tark R, Metelitsa M, Akkermann K, Saks K, Mikkel S, Haljas K. Usability, acceptability, feasibility, and effectiveness of a gamified mobile health intervention (Triumf) for pediatric patients: qualitative study. JMIR Serious Games. 2019;7(3):e13776. [FREE Full text] [doi: 10.2196/13776] [Medline: 31573904]
38. Birrell L, Debenham J, Furneaux-Bate A, Prior K, Spallek S, Thornton L, et al. Evaluating a peer-support mobile app for mental health and substance use among adolescents over 12 months during the COVID-19 pandemic: randomized controlled trial. J Med Internet Res. 2023;25:e45216. [FREE Full text] [doi: 10.2196/45216] [Medline: 37756116]
39. Bohr Y, Litwin L, Hankey JR, McCague H, Singoorie C, Lucassen MFG, et al. Evaluating the utility of a psychoeducational serious game (SPARX) in protecting Inuit youth from depression: pilot randomized controlled trial. JMIR Serious Games. 2023;11:e38493. [FREE Full text] [doi: 10.2196/38493] [Medline: 36892940]

XSL•FO
RenderX

40. Chen Q, Lu Z, Liu B, Xiao Q, Chan KL. Effectiveness of digital game-based GISCC program on cyberbullying prevention among Chinese adolescents. Child Abuse Negl. 2025;161:107293. [FREE Full text] [doi: 10.1016/j.chiabu.2025.107293] [Medline: 39892274]
41. David OA, Cardo RAI, Matu S. Is REThink therapeutic game effective in preventing emotional disorders in children and adolescents? Outcomes of a randomized clinical trial. Eur Child Adolesc Psychiatry. 2019;28(1):111-122. [doi: 10.1007/s00787-018-1192-2] [Medline: 29992353]
42. David OA, Magurean S, Tomoiagă C. Do improvements in therapeutic game-based skills transfer to real life improvements in children's emotion-regulation abilities and mental health? a pilot study that offers preliminary validity of the hink in-game performance scoring. Front Psychiatry. 2022;13:828481. [FREE Full text] [doi: 10.3389/fpsyt.2022.828481] [Medline: 35386525]
43. Dietvorst E, Legerstee JS, Vreeker A, Koval S, Mens MM, Keijsers L, et al. The Grow It! app-longitudinal changes in adolescent well-being during the COVID-19 pandemic: a proof-of-concept study. Eur Child Adolesc Psychiatry. 2023;32(6):1097-1107. [FREE Full text] [doi: 10.1007/s00787-022-01982-z] [Medline: 35524826]
44. Hugh-Jones S, Pert K, Kendal S, Eltringham S, Skelton C, Yaziji N, et al. Adolescents accept digital mental health support in schools: a co-design and feasibility study of a school-based app for UK adolescents. Ment Health Prev. 2022;27:200241. [doi: 10.1016/j.mhp.2022.200241]
45. Keinonen K, Lappalainen P, Kotamäki-Viinikka S, Lappalainen R. Magis - a magical adventure: using a mobile game to deliver an ACT intervention for elementary schoolchildren in classroom settings. J Contextual Behav Sci. 2023;27:26-33. [FREE Full text] [doi: 10.1016/j.jcbs.2022.11.010] [Medline: 36471822]
46. Lincke L, Martin-Döring T, Daunke A, Sadkowiak A, Nolkemper DA, Sproeber-Kolb N, et al. Integration of a mental health app (e-MICHI) into a blended treatment of depression in adolescents: single-group, naturalistic feasibility trial. JMIR Form Res. 2025;9:e58427-e58427. [FREE Full text] [doi: 10.2196/58427] [Medline: 40313202]
47. Miranda C, Matheson B, Datta N, Whyte A, Yang H, Schmiedmayer P, et al. Enhancing distress tolerance skills in adolescents with anorexia nervosa through the BALANCE mobile app: feasibility and acceptability study. JMIR Form Res. 2025;9:e70278. [FREE Full text] [doi: 10.2196/70278] [Medline: 40019817]
48. Mittmann G, Barnard A, Krammer I, Martins D, Dias J. LINA - a social augmented reality game around mental health, supporting real-world connection and sense of belonging for early adolescents. Proc ACM Hum-Comput Interact. 2022;6(CHI PLAY):1-21. [doi: 10.1145/3549505]
49. Peuters C, Maenhout L, Cardon G, De Paepe A, DeSmet A, Lauwerier E, et al. A mobile healthy lifestyle intervention to promote mental health in adolescence: a mixed-methods evaluation. BMC Public Health. 2024;24(1):44. [FREE Full text] [doi: 10.1186/s12889-023-17260-9] [Medline: 38166797]
50. Pramana G, Parmanto B, Lomas J, Lindhiem O, Kendall PC, Silk J. Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: open clinical trial. JMIR Serious Games. 2018;6(2):e9. [FREE Full text] [doi: 10.2196/games.8902] [Medline: 29748165]
51. Stallard P, Whittle K, Moore E, Medina-Lara A, Morrish N, Rhodes S, et al. Acceptability, use and safety of the BlueIce self-harm prevention app: qualitative findings from the Beating Adolescent Self-Harm (BASH) randomised controlled trial. BMJ Ment Health. 2024;27(1):e300961. [FREE Full text] [doi: 10.1136/bmjment-2023-300961] [Medline: 38925663]
52. Williams C, Griffin K, Botvin C, Sousa S, Botvin G. Effectiveness of digital health tools to prevent bullying among middle school students. Adolescents. 2023;3(1):110-130. [doi: 10.3390/adolescents3010009]
53. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ. Lawrence Erlbaum Associates; 1988.
54. Sterne J, Egger M. Regression methods to detect publication and other bias in meta-analysis. In: Rothstein HR, Sutton AJ, Borenstein M, editors. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Hoboken. Wiley; 2006:99-110.
55. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med. 2013;46(1):81-95. [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
56. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res. 2019;21(9):e14567. [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
57. Atalatti S, Pawar U. Addictive interfaces: how persuasion, psychology principles, and emotions shape engagement. In: Interactive Media With Next-Gen Technologies and Their Usability Evaluation. Boca Raton. Chapman and Hall/CRC; 2024:47-66.
58. Srivastava G, Bag S, Rahman MS, Pretorius JHC, Gani MO. Examining the dark side of using gamification elements in online community engagement: an application of PLS-SEM and ANN modeling. Benchmarking. 2022;30(9):2921-2947. [doi: 10.1108/bij-03-2022-0160]
59. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol Bull. 1999;125(6):627-628. [doi: 10.1037/0033-2909.125.6.627] [Medline: 10589297]
60. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol. 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]

XSL•FO
RenderX

61. KIM-studie 2024. MPFS. URL: https://mpfs.de/studie/kim-studie-2024/ [accessed 2024-03-11]
62. Paulus FW, Joas J, Gerstner I, Kühn A, Wenning M, Gehrke T, et al. Problematic internet use among adolescents 18 months after the onset of the COVID-19 pandemic. Children (Basel). 2022;9(11):1724. [FREE Full text] [doi: 10.3390/children9111724] [Medline: 36360452]
63. Girela-Serrano BM, Spiers ADV, Ruotong L, Gangadia S, Toledano MB, Di Simplicio M. Impact of mobile phones and wireless devices use on children and adolescents' mental health: a systematic review. Eur Child Adolesc Psychiatry. 2024;33(6):1621-1651. [FREE Full text] [doi: 10.1007/s00787-022-02012-8] [Medline: 35705765]
64. Twenge JM, Campbell WK. Media use is linked to lower psychological well-being: evidence from three datasets. Psychiatr Q. 2019;90(2):311-331. [doi: 10.1007/s11126-019-09630-7] [Medline: 30859387]
65. Wang M. Developing digital self-regulation in adolescence: a developmental self-determination model of digital self-regulation. Educ Psychol Rev. 2026;38(1):86. [doi: 10.1007/s10648-026-10183-y] [Medline: 42273394]
66. Munzer T, Milkovich L, Madigan S, Tomopoulos S, Parga-Belinkie J, Ajumobi T, et al. Digital ecosystems, children, and adolescents: technical report. Pediatrics. 2026;157(2):e2025075321. [doi: 10.1542/peds.2025-075321] [Medline: 41556931]
67. Kazdin AE, Nock MK. Delineating mechanisms of change in child and adolescent therapy: methodological issues and research recommendations. J Child Psychol Psychiatry. 2003;44(8):1116-1129. [doi: 10.1111/1469-7610.00195] [Medline: 14626454]
68. Lakens D, Mesquida C, Rasti S, Ditroilo M. The benefits of preregistration and registered reports. Evidence-Based Toxicology. 2024;2(1). [doi: 10.1080/2833373x.2024.2376046]

Abbreviations

ACT: acceptance and commitment therapy
ADHD: attention-deficit/hyperactivity disorder
CBT: cognitive behavioral therapy
CESD-R: Centre for Epidemiologic Depression Scale-Revised
DASS-21: Depression Anxiety Stress Scale-21
DBT: dialectical behavior therapy
HRQoL: health-related quality of life
iCBT: internet-delivered cognitive behavioral therapy
LMIC: low- and middle-income country
MMAT: Mixed Methods Appraisal Tool
PHQ-A: Patient Health Questionnaire adapted for adolescents
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by I Steenstra; submitted 17.Jul.2026; peer-reviewed by T Baranowski, N Esfandiari, E Bonilla; comments to author 28.Aug.2026; revised version received 29.Sep.2026; accepted 29.Sep.2026; published 13.Nov.2026

Please cite as:
Baschab JF, Möhler E, Hussong J
Digital Gamified Interventions Targeting Internalizing Mental Health Outcomes in Children and Adolescents: Systematic Review and Meta-Analysis
J Med Internet Res 2026;28:e106933
URL: https://www.jmir.org/2026/1/e106933
doi: 10.2196/106933
PMID:

©Julia Franziska Baschab, Eva Möhler, Justine Hussong. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 13.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX