INPLASY

# Effectiveness of Applied Serious Games in Reducing Anxiety Symptoms Among School Students: A Systematic Review and Network Meta-Analysis

INPLASY202690112
doi: 10.37766/inplasy2026.9.0112
Received: 25 September 2026
Published: 25 September 2026

Yu, M; Wen, KH; Wang, Z; Pang, Y.

**Corresponding author:**
Yu Min

962972721@qq.com

**Author Affiliation:**
City University of Macau.

## ADMINISTRATIVE INFORMATION

**Support -** City University of Macau; no additional external research funding.

**Review Stage at time of this submission -** Preliminary searches.

**Conflicts of interest -** None declared.

**INPLASY registration number:** INPLASY202690112

**Amendments -** This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 25 September 2026 and was last updated on 25 September 2026.

## INTRODUCTION

**Review question / Objective** Title: The effectiveness of applied serious games for anxiety symptoms in school students: A systematic review and network meta-analysis.
Objective: This systematic review and network meta-analysis (NMA) evaluates the relative efficacy of applied serious games in reducing anxiety symptoms among school students, defined using the PICOS framework: Population—school students from elementary through tertiary education; Intervention—five categories of applied serious games (CBT-integrated narrative serious games, biofeedback-based serious games, immersive virtual-reality serious games, exergames, and gamified mental-health mobile applications); Comparator—four categories (waitlist, off-the-shelf leisure games, face-to-face CBT, and non-gamified psychoeducation); Outcome—change in anxiety symptom scale scores (e.g., SCAS, STAI, GAD-7); Study design—randomized controlled trials.
Specific objectives: (1) identify the most effective serious-game category; (2) compare relative efficacy across school education stages (elementary, secondary, tertiary); (3) assess whether benefits are sustained at short- and long-term follow-up. An NMA integrates direct and indirect evidence and ranks interventions by SUCRA values.

**Rationale** Rationale: Anxiety disorders are the most common mental-health conditions among school-aged populations, with a pooled prevalence of ~24.9% in children and adolescents and ~18%–20% clinically significant anxiety among university students. Unaddressed school-year anxiety predicts persistent academic underachievement, impaired social functioning, elevated adult mood-disorder risk, and reduced quality of life. Although face-to-face cognitive behavioral therapy (CBT) is first-line for childhood anxiety, fewer than 25% of anxious students receive evidence-based treatment, owing to shortages of trained

professionals, help-seeking stigma, and scheduling constraints in schools.
Digital interventions, particularly serious games that embed therapeutic content into developmentally matched game environments, are a key pathway to close this treatment gap. Existing core reviews leave gaps: none spans the full continuum from elementary to tertiary education; none systematically compares effect sizes across distinct game-des.

**Condition being studied** Condition(s) studied: This review focuses on anxiety symptoms in school students from elementary through tertiary education, spanning subclinical elevated-anxiety symptoms to clinically significant anxiety disorders.
Anxiety is an emotional state of excessive worry, fear, and tension in response to perceived threat or uncertainty, accompanied by somatic symptoms (palpitations, sweating, muscle tension) and cognitive symptoms (catastrophic thinking, poor concentration, irritability). The anxiety-disorder spectrum (DSM-5/ICD-11) includes generalized, social, separation anxiety disorder, specific phobias, and panic disorder. In children and adolescents, anxiety often presents as behavioral avoidance, social withdrawal, and somatic complaints—developmentally distinct from adult presentations.
Anxiety disorders are the most common mental-health problems in school-aged populations: pooled prevalence is ~24.9% in children and adolescents and ~18%–20% clinically significant anxiety among university students, with dis.

## METHODS

**Search strategy** Search strategy
Databases (inception to September 2026): Web of Science (Core Collection), PubMed, Embase, Cochrane CENTRAL (Cochrane Library), EBSCO-Medline, and ProQuest. Reference lists of included studies were snowball-searched, and grey literature was screened for additional records.
Search terms combined three concept groups with Boolean operators:
Concept 1 (serious games): "serious game", "applied game", "therapeutic game", "game-based intervention", "gamified"/"gamification", "video game", "computer game", "eHealth game", "exergame", "virtual reality game", "biofeedback game".
Concept 2 (anxiety): anxi*, worr*, phobi*, panic, "anxiety disorder", "generalized anxiety", "social anxiety", "separation anxiety".
Concept 3 (school-aged/students): student*, school*, child*, adolescent*, youth*, universit*, colleg*, elementary, secondary, tertiary.
Concept 4 (design): randomized, randomised, "controlled trial", "clinical trial", RCT.
Example PubMed query: (serious game* OR applied game* OR game-based int.

**Participant or population** Participants/population: This review includes school students across the full education continuum from elementary to tertiary level:
Elementary: primary-school children;
Secondary: middle- and high-school adolescents;
Tertiary: university/college undergraduate students.
Anxiety status: eligible participants present anxiety-related symptoms or diagnoses, comprising (1) subclinical elevated anxiety—scored high on standardized scales (e.g., SCAS, STAI, GAD-7, HADS-A) without meeting diagnostic criteria; and (2) clinically significant anxiety—meeting anxiety-disorder diagnostic criteria (DSM/ICD) or scoring above clinical cut-offs.
Key inclusion: studies must assess "anxiety symptoms" as an outcome among enrolled school students; anxiety may define the sample at screening or be measured at baseline.
Exclusion: non-student adults; populations defined primarily by a somatic condition (e.g., cancer, chronic illness, dental/surgical procedures) where anxiety is procedural or secondary (unless participants are school students with anxiety as an outcome); and studies with healthy promotion as the sole aim and no anxiety symptoms. No language restriction; classification is by education stage rather than a strict age cut-off.

**Intervention** Interventions: This review evaluates applied serious games—digital game environments embedding evidence-based therapeutic content for therapeutic/educational goals—in five categories:
1.CBT-integrated narrative serious games: CBT components (cognitive restructuring, behavioral activation, relaxation) woven into role-play/narrative games (e.g., SPARX, Secret Agent Society).
2.Biofeedback-based serious games: games driven by real-time physiological signals (heart-rate variability, EEG) to train relaxation and emotion regulation (e.g., MindLight, Dojo).
3.Immersive virtual-reality (VR) serious games: head-mounted VR exposure/situational training.
4.Exergames: games requiring active physical movement (e.g., VR-exercise, Switch-based games).

5.Gamified mental-health mobile applications: mobile apps packaging psychological content with gamified elements (points, levels, rewards).
Comparators: waitlist, off-the-shelf leisure games, face-to-face CBT, and non-gamified psychoeducation.
Key definitions: included studies must present the serious game as an identifiable, independently described intervention arm (not as an inseparable mixed component), delivered as self-guided or low-intensity supervised, with session dose, frequency, and follow-up recorded. Interventions form distinct network nodes to enable NMA comparison of relative efficacy across game categories and comparators.

**Comparator** Comparators: Four comparator conditions are applied to the target population as controls against serious-game arms:
1.Waitlist / no-intervention control: allocation to routine waiting or inert control with no active intervention—the common baseline for estimating absolute efficacy versus no treatment.
2.Off-the-shelf leisure games: non-therapeutic, commercially available entertainment games, distinguishing evidence-based therapeutic content from general game activity and controlling non-specific effects of game exposure.
3.Face-to-face cognitive behavioral therapy (CBT): standard professionally delivered CBT as an active evidence-based "gold-standard" comparator for assessing comparability with conventional first-line treatment.
4.Non-gamified psychoeducation: psychoeducational content (lectures, manuals, text materials) without gamification/interactive elements, testing whether the serious-game format adds benefit over equivalent non-gamified delivery.
Network role: These four comparators together with the five serious-game categories form the network nodes, enabling each game type to be compared against every comparator via direct and indirect evidence, with relative efficacy ranked by SUCRA. In each trial, serious-game arms and corresponding comparator arms are categorized accordingly.

**Study designs to be included** Study designs included: To achieve the NMA objectives, this review includes randomized controlled trials (RCTs):Parallel-group RCTs: individual randomization to serious-game and comparator arms; Cluster-randomized trials: randomization at school/class level;?Two- or multi-arm trials: parallel arms including multiple serious-game or comparator categories; Both non-inferiority and superiority randomized designs.Inclusion conditions: participants randomly allocated to intervention/control; anxiety symptoms as an outcome; and extractable effect-size data (means, SDs, sample sizes, or co.

**Eligibility criteria** Eligibility criteria (additional beyond PICOS)
Additional inclusion:
Publication type: peer-reviewed full-text studies; full papers prioritized over preprints/conference reports.
Extractable data: pre/post (and follow-up) anxiety-outcome effect sizes (group means, SDs, sample sizes, or convertible statistics); studies with incomplete unextractable data (and no author response) are excluded.
Identifiable intervention: the serious game must be an independent, clearly described arm, not an inseparable mixed component.
Additional exclusion:
Duplicate/overlapping datasets: where the same trial is published in multiple reports, only the most complete is retained.
Conference abstracts, study protocols, registry entries, and dissertations (unless formally published).
Non-randomized/observational, single-arm, and quasi-experimental designs.
Unclear intervention protocol (serious-game composition/delivery unascertainable).
Mixed interventions with inseparable effects.
Studies without anxiety as a primary or extractable outcome.
Populations without anxiety symptoms (pure health promotion).
Language: no language restriction; studies in English, Chinese, and other languages are eligible (data extracted by translation).

**Information sources** Information sources:
1.Electronic databases (inception to September 2026): Web of Science (Core Collection), PubMed, Embase, Cochrane CENTRAL (Cochrane Library), EBSCO-Medline, and ProQuest—six databases spanning medical, psychological, and interdisciplinary literature.
2.Citation tracking (snowballing): backward citation searching of reference lists of included studies and forward tracking via citation databases (e.g., Scopus, Google Scholar).
3.Trial registries and grey literature: ClinicalTrials.gov and WHO ICTRP to identify registered/ongoing RCTs and reduce publication bias; preprints and grey literature (theses, conference reports) screened where feasible.
4.Author contact: corresponding authors contacted by email when reported data are incomplete or unextractable, to obtain missing outcome data or methodological details.

All records are imported into reference-management software for deduplication and screened independently by two reviewers. Hits from each source are recorded and summarized in a PRISMA flow diagram.

**Main outcome(s)** Main outcomes: Change in anxiety symptoms measured by validated standardized anxiety scales (e.g., SCAS, STAI, GAD-7, HADS-A) via self- or proxy-report.
Effect measure: The standardized mean difference (SMD) between groups, entered into the network meta-analysis.
Time points: Two windows—(1) immediately post-intervention (primary), and (2) follow-up, stratified into short (<6 months) and long (≥6 months) to assess sustainability. Where multiple time points are reported, post-intervention and the longest follow-up are prioritized.
Ranking: Interventions ranked by SUCRA values.

**Additional outcome(s)** Additional outcomes: Where reported by included studies, the following secondary outcomes are also extracted:
Depressive symptoms: change on standardized scales (e.g., PHQ, CES-D, CDI);
Mental-health literacy and stigma: scale scores and help-seeking attitudes, if reported;
Quality of life / well-being: e.g., PedsQL scores;
Acceptability and adherence: completion/dropout rates and participant feedback, as feasibility evidence (not quantitatively pooled).
Additional outcomes are extracted only when reported and extractable, and serve to describe overall effects and implementation characteristics rather than as primary effect sizes in the NMA.

**Data management** Data management: Records and data are managed per systematic-review standards:
Reference management: all records imported into reference-management software (e.g., EndNote/Zotero), deduplicated automatically plus manually to form the screened pool.
Screening & extraction: two reviewers independently screen titles/abstracts and assess full texts, then independently extract data into a pre-designed standardized extraction form (predefined fields: study characteristics, population, intervention/comparator details, outcome time points and effect sizes, risk-of-bias items); discrepancies resolved by discussion or a third reviewer.
Storage & verification: extracted data entered into electronic tables with double-entry/cross-checking; outliers checked against source texts; all screening decisions, exclusion reasons, extracted values, and source citations archived to maintain an auditable trail.
Completeness: corresponding authors contacted for missing data; unextractable data reported as missing, never imputed.
Backup & versioning: search records, screening sheets, and extraction tables regularly backed up with version history for full reproducibility.

**Quality assessment / Risk of bias analysis** Risk-of-bias assessment: Each included RCT is assessed using the Cochrane RoB 2.0 tool across five domains:
D1 randomization process (sequence generation and allocation concealment);
D2 deviations from intended interventions (blinding and adherence);
D3 missing outcome data (completeness and handling);
D4 measurement of the outcome (blinding of assessors, reliability);
D5 selection of the reported result (reporting bias).
Each domain and overall judgment is rated low risk / some concerns / high risk.
Process: two reviewers independently assess, resolving disagreements by discussion or a third reviewer. Results are presented as a per-study domain matrix and a risk-of-bias summary figure.
Application: judgments inform interpretation; where evidence allows, a sensitivity analysis excluding high-risk studies tests robustness, and risk of bias is reported as a limitation.

**Strategy of data synthesis** Data synthesis: A frequentist random-effects network meta-analysis (NMA) is used to synthesize relative efficacy.
Effect measure/model: standardized mean difference (SMD) of anxiety scores (post-intervention/follow-up); random-effects model given between-study heterogeneity, quantified by $\tau^2$ and $I^2$.
Network: 9 nodes—five serious-game categories (CBT narrative, biofeedback, immersive VR, exergames, gamified apps) plus four comparators (waitlist, leisure games, face-to-face CBT, non-gamified psychoeducation); a network graph presents direct and indirect evidence.
Consistency: global consistency assessed by $\chi^2$; local consistency via node-splitting and loop-specific analysis; if global inconsistency is significant, an inconsistency model is used for interpretation.
Ranking: interventions ranked by SUCRA and PrBest.
Subgroup analyses: stratified by education stage (elementary/secondary/tertiary), follow-up duration (short/long), supervision mode (self-guided/low-intensity), and outcome instrument.

Sensitivity analysis: leave-one-out and, where evidence permits, exclusion of high-risk-of-bias studies.
Software: analyses performed in R (netmeta package).

**Subgroup analysis** Subgroup analyses: To test whether relative efficacy varies by population, intervention, and measurement characteristics, the following prespecified subgroups are analyzed:
1.Education stage: elementary, secondary, and tertiary, testing whether serious-game effects generalize across school stages ("cross-stage adaptation").
2.Follow-up duration: immediate post-intervention, short (<6 months), and long (≥6 months), assessing immediacy and sustainability.
3.Supervision mode: fully self-guided vs. low-intensity professional supervision, testing moderation by support intensity.
4.Outcome instrument: by scale family (SCAS, STAI, GAD-7, others), testing measurement sensitivity.
Subgroups are analyzed within the random-effects NMA framework, with effect sizes and SUCRA rankings estimated per stratum. Given limited studies in some subgroups, results are interpreted exploratorily, bounded by sample size and heterogeneity.

**Sensitivity analysis** Sensitivity analyses: To test robustness of the NMA results:
1.Leave-one-out: re-estimate relative efficacy and SUCRA rankings after sequentially excluding each study, testing whether pooled effects are driven by a single trial.
2.Risk-of-bias stratification: where evidence allows, re-analyze after excluding studies rated high risk of bias.
3.Model robustness: compare random-effects vs. fixed-effects estimates to assess the impact of model choice.
4.Comparator composition: where network structure permits, test whether inclusion of different comparator types (especially waitlist vs. active comparators) alters rankings.
5.Data-extraction assumptions: re-check converted/derived effect sizes to test the impact of extraction assumptions.
Results are judged robust if primary effect sizes and rankings remain consistent across most sensitivity scenarios; otherwise, affected conclusions are reported cautiously with evidence limitations.

**Language restriction** No language restriction. Search and screening include English, Chinese, and other languages; non-English studies are included via translation to ensure comprehensiveness and reduce language bias.

**Country(ies) involved** China.

**Keywords** serious games; applied games; anxiety; anxiety symptoms; network meta-analysis; school students; randomized controlled trial; digital mental health intervention; gamification; biofeedback;.

**Contributions of each author**
Author 1 - Yu Min.
Email: 962972721@qq.com
Author 2 - Wen, KH.
Author 3 - Wang, Z.
Author 4 - Pang, Y.