Received: 23 June 2023 | Revised: 7 November 2023 | Accepted: 7 November 2023

DOI: 10.1002/cpp.2938

COMPREHENSIVE REVIEW

WILEY

Therapeutic games to reduce anxiety and depression in young people: A systematic review and exploratory meta-analysis of their use and effectiveness

Zarah Eve[1] | Martin Turner[1] | Daniela Di Basilio[2] | Benjamin Harkin[1] | Alan Yates[3] | Sofia Persson[4] | John Henry[5] | Ashley Williams[5] | Geoff Walton[6] | Marc V. Jones[1] | Charlotte Whitley[1] | Nathan Craddock[7]

[1]Department of Psychology, Manchester Metropolitan University, Manchester, UK

[2]Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[3]Department of Education, Manchester Metropolitan University, Manchester, UK

[4]Department of Psychology, Leeds Beckett University, Leeds, UK

[5]Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, UK

[6]Department of Information and Communications, Manchester Metropolitan University, Manchester, UK

[7]Department of People and Performance, Manchester Metropolitan University, Manchester, UK

Correspondence
Zarah Eve, Department of Psychology, Manchester Metropolitan University, Manchester, UK
Email: z.eve@mmu.ac.uk

Abstract

Objective: The development of serious games for mental wellbeing is a topic of growing interest. The increase in acceptance of games as a mainstream entertainment medium combined with the immersive qualities of games provides opportunities for meaningful support and intervention in mental wellbeing.

Method: We conducted a systematic review and exploratory meta-analysis to examine if aspects of the interventions influenced outcomes as measured via overall effect sizes. We employed a multilevel meta-analytic approach to accommodate the interdependency of effect sizes (18 effect sizes from 14 studies, with 2027 participants).

Results: Overall, the main effect for gaming interventions on any outcome variable was small to medium sized, $d = .35$ (confidence interval [.23, 47], $p < .001$). Results revealed that the only significant moderator was the nature of the intervention. Specifically, only interventions that included a rational emotional behavioural focus significantly predicted an improvement in depression and/or anxiety in participants.

Conclusion: The findings reveal promising effects for therapeutic games for mental health, but replications are needed, alongside the addressing of methodological and procedural concerns.

KEYWORDS
anxiety, depression, meta-analysis, serious games, systematic review

1 | INTRODUCTION

Both retrospective and prospective research have identified that most adult mental health problems begin in childhood and adolescence (Kessler et al., 2007). NHS digital mental health statistics show that over 340,000 children and young people were in contact with mental health services during a 1-month period (July 2021; NHS Digital, 2021a). Rates of probable mental health disorders have increased in the United Kingdom since 2017, from 1 in 9 to 1 in 6 in 6- to 16-year-olds, and from 1 in 10 to 1 in 6 in 17- to 19-year-olds (NHS Digital, 2021b). Anxiety and depression have been identified as the most common mental health conditions, particularly within teenage populations: 1 in 12 teens report experiencing anxiety, and 1 in 20 report experiencing depression (Radez et al., 2021). Understandably, prevalence rates of mental health problems have been exacerbated by the COVID-19 pandemic, with 18%–60% of children and young people

scoring above the threshold for anxiety and depressive symptoms throughout the period of restrictions (Viner et al., 2021). Pre-adolescents (4–10 years) were identified as having the greatest change in mental health difficulties in line with restrictions, thought to be linked to the lack of social spaces and companionship available for younger children (Creswell et al., 2021).

With continuous technological advances being made, it is important to consider approaches and modalities that children and young people are inclined to use and engage positively with in order to address growing mental health concerns. The benefits of digital health interventions are vast, including increased accessibility, anonymity and cost-effectiveness (O'Connor et al., 2016). Digital solutions can also scale in a way that analogue (i.e., face to face) cannot. While face-to-face interventions (delivered either in person or online) have demonstrated overall efficacy on a range of conditions affecting young people (Krzyzaniak et al., 2021; Reynolds et al., 2012), there are a range of barriers to accessing adequate support, including concerns about confidentiality, embarrassment and stigma, particularly in younger cohorts (Radez et al., 2021). Additionally, financial implications and limited access to services are noted across populations and constitute an ongoing issue in the United Kingdom particularly for younger generations in need of psychological support (Punton et al., 2022). As such, many interventions have been adapted to be delivered using digital tools, including apps, virtual reality programmes and websites. While mental health literacy may be lower in younger children and adolescents (Radez et al., 2021), digital health literacy in younger people has increased exponentially over recent decades (Paige et al., 2018). As a result, engaging people in digital modalities can help break down barriers within engagement and understanding and reduce disengagement from treatment (Liverpool et al., 2020).

In light of the utility of combining digital technologies with traditional interventions, there are emergent delivery methods that are gradually becoming well-established alternatives to, or integrations of, clinician-delivered face-to-face sessions. These include, for example, telephone therapy (Brenes et al., 2015), therapy delivered via texting (Snoswell et al., 2021) and internet-based therapy (Barak et al., 2008; Välimäki et al., 2017). As indicated in Barak et al. (2008), online psychological interventions demonstrated comparative effectiveness with more traditional methods of therapeutic activity.

Nonetheless, clinician-delivered digital interventions are not the only form of therapeutic support available, as there is a growing interest, particularly in younger generations, in new forms of digital interventions, including 'blended interventions' (encompassing live interactions with therapists and e-contents) and application-based therapeutic activities, also known as 'apps for mHealth' (mobile health) (Rasing, 2021). Smartphone apps, in specific, appear to be an increasingly popular digital way of accessing readily available mental health support, due to their 24-h availability, unique characteristics and increased connectivity, have been suggested to be a beneficial method for therapeutic delivery (Grist et al., 2017). Data from the United Nations indicated that over 90% of the population in developed countries utilise apps daily (International Telecommunications Union, 2016) and their use is also growing within low- and middle-income countries (Huang et al., 2019).

Key practitioner message

- Based on the evidence presented in the current paper, there could be some benefits to adopting serious game for the purposes of ameliorating anxiety and depression.
- Interventions that included a rational emotional behavioural focus significantly predicted an improvement in depression and/or anxiety.
- However, caution should be exercised due to the methodological issues apparent in the empirical studies conducted in this area.

Specific therapeutic approaches have been highlighted as more suitable for digital delivery, with cognitive-behavioural therapeutic (CBT) approaches and components germane to general CBT (e.g., behavioural activation) being widely used in digital interventions (Christ et al., 2020) mostly due to their structure, adaptability, and ability for personalisation (Cuijpers et al., 2016). In spite of the growth in the range of e-health apps available, there is inconclusive evidence on their clinical effectiveness, which is required to carefully determine quality and applicability (Neary & Schueller, 2018).

Video games are now the most popular form of entertainment in the United Kingdom (Dealessandri, 2023). In a recent survey of 16- to 24-year-olds in the United Kingdom, 92% of respondents in 2020–2021 had played a video game, up from 73% in the previous year (Ofcom, 2022). Video game users in the United Kingdom increased from 36 million in 2020 to 44.3 million in 2021, a potential impact of quarantine measures imposed by the COVID-19 pandemic (Clement, 2022). It has been projected that users will increase to 51.88 million by 2025. Moreover, the number of platforms available for game play has increased rapidly, enabling a wider reach. Distribution of game play across media varies for UK users, with mobile phones, game consoles and tablets remaining the most popular across age ranges (Ofcom, 2022). In sum, video gaming is an extremely popular activity with a growing market.

Play can be viewed as a window into inner experiences, which is a potential justification for current therapeutic serious games (Martins et al., 2011). Wilkinson (2016) notes that for a young person in particular, play is a natural behaviour, and thus engaging in therapeutic activity through play allows for a naturalistic approach that may put a young person at ease (Landreth, 2012). Due to the familiarity and comfort with gaming, integrating therapeutic approaches with gaming, as is attempted in some serious games, may allow therapeutic content to be taken up to engage groups more readily (Brown, 2013). In sum, drawing on people's natural proclivity to play, therapeutic games (TGs), also known as serious therapeutic games (Hudlicka, 2016), may provide a powerful way to engage people in psychological support.

Similar to apps, but more recent in their emergence in the field of digital solutions for mental health, 'serious games' represent a novel yet promising alternative to more 'standard' forms of online interventions (Fleming et al., 2017). Serious games demonstrate increased efficacy and effectiveness across a spectrum of conditions (Lau et al., 2017;

McCallum, 2012) and have the potential to empower vulnerable groups as they can be used as an independently accessible, low-cost solution to access mental health support (van der Lubbe et al., 2021). Although 'serious games' and 'gamification' are often used interchangeably, they relate to specific concepts. Gamification integrates characteristics of games into non-game activities, for example, through the inclusion of points, badges and leaderboards when levels are completed (Sicart, 2014). Often used in conjunction with exercise games (synonymously known as exergames, e.g., Pokémon GO), combining gaming notions within applications and websites has been found to aid adherence to health activities (Miller et al., 2014). Often gamification activities are developed as traditional entertainment products, with game elements added after; thus, they do not operate as full games (Sailer et al., 2017). However, the incorporation of gamification elements is often developed without consideration of motivation, engagement and behaviour change theories (Seaborn & Fels, 2015). In contrast, serious games are developed specifically with a focus beyond traditional leisure activities (Fleming et al., 2017). Serious games used for health utilise specific game design elements (e.g., story, mechanics and platforms) to foster healthcare benefits. As Protopsaltis et al. (2011) noted, while serious games could be motivating, users should be encouraged to become involved in the game in a personal, emotional and cognitive way.

2 | RATIONALE

There have been other reviews that pertain to serious games for mental health, on which we wish to build. Abd-Alrazaq, Alajlani, et al. (2022) and Abd-Alrazaq, Al-Jafar, et al. (2022) explored the impact of a broader notion of serious games, encompassing all games that were developed for 'serious' purposes, separately on anxiety (Abd-Alrazaq, Alajlani, et al., 2022) and depression (Abd-Alrazaq, Al-Jafar, et al., 2022). The purpose of these games included brain training, social skills development, biofeedback games, distraction of attention and exergames along with TGs such as computerised CBT. Research included games developed for education, prevention and treatment of chronic conditions and mental health, encompassing a broad range of focus. While Abd-Alrazaq, Alajlani, et al. (2022) included 33 studies, only 8 focused specifically on anxiety, and while Abd-Alrazaq, Al-Jafar, et al. (2022) include 27 studies, only 3 of the included in focused specifically on depression. Furthermore, Abd-Alrazaq, Alajlani, et al. (2022) did not include any follow-up data within their meta-analysis, limiting the understanding as to the effectiveness of serious games.

All systematic reviews have limitations that serve to inform future researchers in their work. The reviews by Abd-Alrazaq et al. are highly valuable pieces of work; however, there are aspects of their work that we aim to build upon. For example, Abd-Alrazaq, Alajlani, et al. (2022) and Abd-Alrazaq, Al-Jafar, et al. (2022) have a broad notion of serious games with purposes that had been adapted for therapeutic purposes, but were not developed initially for therapeutic use, alongside games specifically designed and developed for therapeutic purposes. In our systematic review, we focus only on games that have been developed for therapeutic purposes (i.e., TGs) so that we can assess the effectiveness of those games against the purposes they were designed for. Also, Abd-Alrazaq et al. incorporate other forms of serious games, such as exergames (i.e., games specifically developed for physical health purposes, which address mental health as secondary concerns). We focus our review on TGs which have been developed specifically and primarily to address mental health outcomes in order to examine whether those games deliver on their proposed purpose. Furthermore, Abd-Alrazaq et al. include studies which include mental health as secondary outcomes (e.g., interventions targeting physical health conditions that contribute to poor mental health). In our review, mental health outcomes are the primary outcome because we are specifically interested in examining the evidence for the use of TGs on those outcomes.

In sum, we aim to build upon the work of Abd-Alrazaq, Alajlani, et al. (2022) and Abd-Alrazaq, Al-Jafar, et al. (2022) by focusing our review on the effectiveness of TGs on the mental health outcomes they are proposed to alleviate. In doing so, we hope to be able to offer conclusions concerning the evidence for the use of TGs for mental health and provide recommendations for the future development and study of TGs. By providing a contemporary review that is specifically focused upon serious games that have been developed to therapeutically address these conditions, we hope to inform researchers, games developers, clinicians and mental health services as to the current state of knowledge in this growing field. We need to capture the state of the extant evidence with regard to serious games so that we can inform key stakeholders of adolescent mental health with the benefits and drawbacks of TGs for mental health.

3 | METHOD

We conducted a systematic review and meta-analysis in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The protocol was registered on PROSPERO (International Prospective Register for Systematic Reviews; ID: CRD42022343882).

3.1 | Search strategy

A literature search of PubMed, Web of Science, PsycINFO and CINAHL databases was conducted. The databases were searched on 5 April 2022, with an updated search conducted on 30 March 2023, by the first author. Forward and backward reference list checking was additionally conducted (screening reference lists of included studies and screening articles that cite back to included studies). Two experts in literature searching and computer games were consulted before developing the search string for this review. A scoping search was conducted in March 2022, wherein subsequent adaptations were made to the search string, ensuring relevant research was identified. The search string was a combination of games-related terms including 'game' and 'computer' mental health-related terms, including 'intervention', 'therapy' and 'psychotherapy', and outcome-related terms, including 'effect', 'efficacy' and 'outcome'. Boolean operators and truncation were used to broaden and narrow the search, ensuring specificity.

3.2 | Study eligibility criteria

Quantitative studies that assessed the effectiveness of TGs in alleviating mental health symptoms were included in this review. Specifically, the target intervention in this review was serious games that were developed specifically for the purposes of therapeutic or prevention of mental health symptoms (i.e., TGs). TGs were required to be delivered on digital platforms including computers, mobile phones and tablets. Non-digital games, along with those used for other purposes, including monitoring, diagnosis or screening, were excluded. The outcomes of interest were anxiety and depression; thus, outcome data measured immediately after intervention, along with associated follow-up data, were included.

Research in the English language were eligible for inclusion, with non-English articles being excluded. Research that focused on people with comorbidities or physical health conditions where mental ill health was a secondary condition or experience were excluded. Conference abstracts and posters, proposals, commentaries and editorials were excluded. No restrictions related to the population, country of publication, year of publication, comparator or study settings were applied.

3.3 | Study selection

We identified relevant studies through the following steps. First, retrieved studies were exported into Rayyan software to identify and eliminate duplicate entries. In the second step, seven reviewers (authors 1, 2, 3, 7, 8, 9 and 10) screened all retrieved papers for title inclusion. Subsequently, three reviewers (authors 1, 2 and 3) screened the title and abstracts of the included studies. Finally, the full texts of the studies included in the previous step were screened by three reviewers. The three reviewers resolved disagreement through discussion, with an independent fourth reviewer available if consensus could not be reached.

3.4 | Data extraction and synthesis

Three authors (authors 1, 11 and 12) used Microsoft Excel to extract data from the included studies. We used a narrative and statistical approach to synthesise the extracted data. Within our narrative synthesis, we describe the characteristics of included studies, population, intervention, comparator and outcome measures using text and tables.

3.5 | Quality assessment

The quality of research was assessed using the Effective Public Health Practice Project (EPHPP) Quality Assessment Tool for Quantitative Studies (Thomas et al., 1998). The tool appraises the quality of research in six areas: selection bias, study design, confounders, blinding, data collection methods, and withdrawals and dropouts. The overall quality of the meta-analysed research was appraised separately by three reviewers (authors 1, 2 and 3), and any differences in decisions were addressed through discussion (see Table 1).

3.5.1 | Results of systematic review

Search results

As shown in Figure 1, we identified 18,351 records by searching four electronic databases: Web of Science, PubMed, CINAHL and PsycINFO. Of these records, we identified and removed 5567 duplicates using Rayyan software. The screening of the titles and abstracts of the remaining 12,784 records led to the exclusion of 12,613 citations. Reading the full text of the remaining 171 publications led to the exclusion of 157 publications for the following reasons: (1) the intervention was not TG focused ($n = 47$), (2) it was not specific to anxiety or depression ($n = 42$), (3) the intervention was not a TG (i.e., a game not specifically designed for therapeutic use) ($n = 20$), (4) the article did not test effectiveness ($n = 11$), (5) protocol only ($n = 9$), (6) qualitative research ($n = 9$), (7) the intervention was a non-digital game ($n = 9$), (8) the intervention was virtual reality focused ($n = 6$), (9) conference abstract ($n = 3$), and (10) the article was not in English language ($n = 1$). A total of 14 articles were included in this review. We then conducted an exploratory meta-analysis (authors 4, 5, and 6), extracting 19 effect sizes from these 14 studies.

Characteristics of included studies

The included studies were published between 2012 and 2022 (table 1). The year that witnessed the largest number of included studies was 2018 ($n = 4$), followed by 2016 ($n = 3$). The included studies were conducted in five different countries, as shown in Table 1. The country that published the largest number of included studies was the Netherlands ($n = 5$), followed by New Zealand ($n = 3$) and Romania ($n = 3$). All included studies were published in peer-reviewed journals. While not explicitly included in the inclusion criteria, all included studies were randomised controlled trials ($n = 14$).

The sample size in the included studies ranged from 32 to 283, with an average of 144.8 (*SD* 66.6). The target participants were children (aged 7–13 years) in four studies, adolescents (aged 13–20 years) in nine studies and adults (aged >20 years) in one study. Specifically, the mean age of participants reported in the 14 studies included in this review ranged between 8.87 and 42.28 years, with an average of 15.44 (*SD* 8.2) years. The percentage of female participants across the studies reviewed ranged from 10.5% to 100%, with an average of 56.78% (*SD* 18.9%). The participants' mental health conditions varied, with anxiety and depression being explored as the primary condition equally across studies ($n = 7$). Depression was also explored as a secondary condition in four included studies.

As shown in Table 2, the duration of gameplay per session ranged from 10 to 60 min, with an average duration of 40 min (*SD* 15.6 min); however, it was most commonly reported as 30 min ($n = 5$). The total

TABLE 1 Characteristics of studies and populations.

Study	Year	Quality assessment	Country	Study type	Sample size, *n*	Age (years), mean	Sex (female), %	Health condition	Population setting
McCashin et al.	2022	Moderate	Ireland	RCT	122	13.87	42	Depression	Outpatient
David & Fodor	2022	Moderate	Romania	RCT	165	13.43	59	Depression	Educational
Schuurmans et al.	2018	Strong	Netherlands	RCT	37	15.60	16	Depression	Residential care
Schoneveld et al.	2018	Strong	Netherlands	RCT	174	12.75	59	Depression	Educational
David et al.	2018b	Strong	Romania	RCT	165	17.54	64	Anxiety	Educational
David et al.	2018a	Weak	Romania	RCT	165	12.98	64	Anxiety	General population
Sanchez et al.	2017	Moderate	USA	RCT	69	12.89	41	Depression	General population
Kuosmanen et al.	2017	Weak	New Zealand	RCT	146	13.97	39	Anxiety, depression	Educational
Schoneveld et al.	2016	Moderate	Netherlands	RCT	136	13.02	55	Depression	Educational
Scholten et al.	2016	Weak	Netherlands	RCT	138	9.95	65	Depression	Educational
Poppelaars et al.	2016	Strong	Netherlands	RCT	208	8.91	100	Anxiety	Educational
Roepke et al.	2015	Moderate	USA	RCT	283	13.35	71	Anxiety, depression	General population
Merry et al.	2012	Strong	New Zealand	RCT	187	9.97	66	Anxiety, depression	Outpatient
Fleming et al	2012	Moderate	New Zealand	RCT	32	14.90	44	Anxiety, depression	Educational

duration of the games included in the included studies ranged from 210 to 360 min, with an average total duration of 285 min (*SD* 64.1 min). Frequency of game play varied across studies, with an average of 8.57 sessions (*SD* 6.2 sessions) being played. Style of guidance within the TG varied across studies, including 1st person gameplay ($n = 3$), 3rd person gameplay ($n = 6$), guided gameplay ($n = 4$) and informational guidance ($n = 1$). We identified five types of TG, based on the psychotherapeutic focus of the game elements: general CBT (non-specific mode of CBT) games ($n = 7$), rational emotive behaviour therapy (REBT) (a specific mode of CBT) ($n = 3$), biofeedback game response ($n = 2$), social skill development ($n = 1$) and positive-CBT (blend of positive psychology and CBT principles) ($n = 1$). While biofeedback was specifically focused in two studies, it was used as a tool in four studies. For the games informed by general CBT, extant literature concerning game development does not point to a specific mode of CBT. The exception to this is REThink, where developers specifically state that the game is underpinned by a specific form of CBT, namely, REBT (Ellis, 1995).

As shown in Table 3, the comparison groups received inactive interventions in 9 studies and active interventions in 10 studies. Active comparator interventions included REBT ($n = 3$), Coping Cat ($n = 1$), treatment as usual ($n = 2$), school-based general CBT programme ($n = 1$) and non-TG play ($n = 3$). Note that the numbers do not add up to total studies for comparators, as certain studies utilised two comparator groups, involving both inactive and active comparators ($n = 5$). The duration of the active comparators ranged from 210 to 540 min. The outcome of interest (anxiety or depression level) was measured using 10 different tools, but the most common tools used by the included studies were the Spence Childhood Anxiety Scale ($n = 6$), and the Reynolds Adolescent Depression Scale-2 (RADS-2) ($n = 6$). The outcome of interest was immediately measured after the intervention in all included studies ($n = 14$), and the most common follow-up period was 3 months ($n = 7$), followed by 6 months ($n = 5$). Participant attrition was reported in 14 studies, ranging from 3.92% to 78.49%, with an average of 21.2% (*SD* 23.3%) attrition of participants.

4 | META-ANALYSIS

4.1 | Coding of study characteristics

For each study, we coded the following methodological characteristics: (a) bibliographic information (e.g., author, year and country of

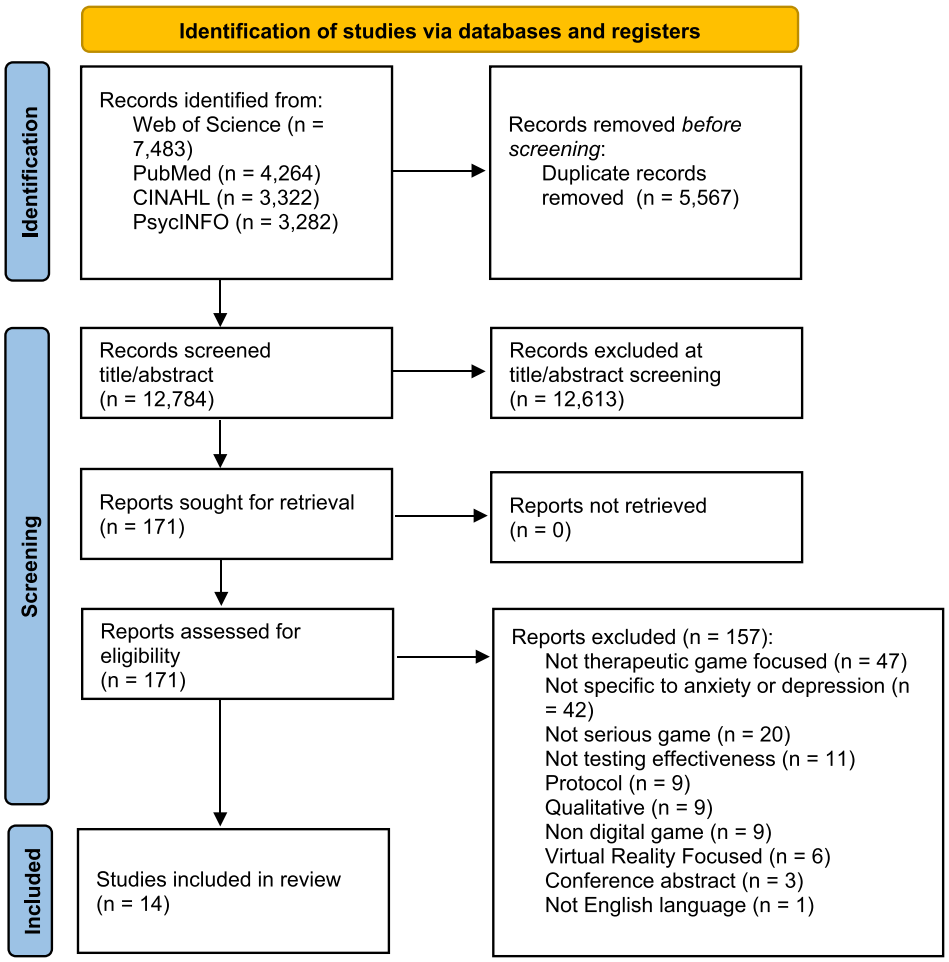

FIGURE 1 PRISMA flowchart of the study selection process.

publication); (b) the nature of the intervention (e.g., REBT and biofeedback); (c) specific outcome measures (e.g., anxiety and/or depression) and (d) methodological quality. We also coded specific features of the intervention sample: (a) type of sample (e.g., general population, outpatient and residential patient); (b) average age; (c) proportion females; (d) primary disorder (e.g., anxiety or depression); and (e) medication status. Last, specific to the interventions we coded the following: (a) whether biofeedback was included as an additional intervention component; (b) avatar presentation or style (e.g., 1st person, 3rd person, guided and instructional); (c) the specific interventional focus (e.g., psychological resilience, self-regulation; replacing negative automatic thoughts); (d) who delivered the intervention (e.g., therapist, teacher, psychologist and research assistant); (e) dosage of the intervention (i.e., calculated as a sum of total sessions multiplied by the length of the session to give a total dosage time in minutes); and (f) drop-out rates (i.e., a percentage of total participants who drop-out from the beginning to the end of the intervention). With the scope of the research question being broad, no previous reviews or meta-analyses being conducted in this area and limited number of included data points, this meta-analysis is exploratory in nature.

For reliability purposes, a two-step interrater analysis was conducted on the categorical variables (medication, comorbidities and quality) and the continuous variables (effect size and standard error). Step 1: Extraction of medication, comorbidities and quality were coded blindly and independently by a second coder (author 5) at 100% duplication. An interrater reliability analysis using Cohen's Kappa statistic for combined categorical variable (medication, comorbidities and quality) showed moderate agreement $\kappa = .558$ (95% confidence interval [.199, .852], $p < .001$). Any differences were resolved through discussion and final agreement within the team, with virtually all discretions attributable to different perceptions on the quality of some articles. A slight revision on categorising the use of medication was also made with healthy participants assumed not to medicate. Step 2: Extraction of the effect size (ES) and standard error (*SE*) data were coded blindly and independently at 100% duplication. A two-way mixed-effect intraclass correlation coefficient (ICC: type and absolute agreement) was conducted for the combined continuous variables standard errors and effect sizes. The average measure ICC was .99 (95% CI [.998, .999]), $p < .001$, indicating excellent agreement, with discretions attributable to slight variation in calculations computed.

TABLE 2 Characteristics of interventions.

Study	Serious game name	Serious game type	Platform	Duration per session (min)	Number of sessions	Total duration of intervention (min)
McCashin et al. (2022)	Pesky gNATs	General CBT game	Computer and phone	45	7	315
David and Fodor (2022)	REThink	REBT game	iPad	50	7	350
Schuurmans et al. (2018)	Dojo	Biofeedback game	Computer	30	8	240
Schoneveld et al. (2018)	MindLight	General CBT game	Xbox 360	60	6	360
David et al. (2018)	REThink	REBT game	iPad	50	7	350
David et al. (2018)	REThink	REBT game	iPad	50	7	350
Sanchez et al. (2017)	Adventures aboard the S.S. GRIN	Social skill development	Computer	25	9	225
Kuosmanen et al. (2017)	SPARX-R	General CBT game	Computer	30	7	210
Schoneveld et al. (2016)	MindLight	General CBT game	Xbox 360	60	5	300
Scholten et al. (2016)	Dojo	Biofeedback game	Computer	60	6	360
Poppelaars et al. (2016)	SPARX	General	CD-ROM	30	7	210
Roepke et al. (2015)	SuperBetter	Positive-CBT game	Phone	10	30	300
Merry et al. (2012)	SPARX	General CBT game	Computer or phone	30	7	210
Fleming et al. (2012)	SPARX	General CBT game	Computer	30	7	210

4.2 | Computing effect sizes

Cohen's *d* was calculated as a measure of effect size for the difference in outcome measures between the intervention and control condition. A positive effect indicated an improvement in the outcome compared to the control condition. Where Cohen's *d* was not reported in the original study, effect sizes were either converted from other effects (e.g., *F*). As recommended by Assink and Wibbelink (2016), the variance was calculated as SE^2. As we included five studies that provided two effect sizes, this violates the normal requirement of independent effect size measures in meta-analysis (Cheung, 2014). Dependency of effect sizes normally means that effect sizes within studies are correlated (as these are expected to show a certain similarity); this creates an overlap of information and inflates information produced by the analysis, which can result in an overconfidence in its results (Assink & Wibbelink, 2016). Although it is possible to conduct subgroup analysis or aggregating effect sizes, this reduces the number of effect sizes analysed in a set, therefore limiting the power of the analysis, something that is a particular concern for an exploratory meta-analysis with a limited number of data points such as ours (Assink & Wibbelink, 2016).

Where correlations between effect sizes are not known, it is possible to fit a three-level meta-analytical structure. This analysis considers three levels of variance components distributed across the model, including variance between effect sizes from the same study and variance between studies; this therefore allows for an examination of how effect sizes vary between participants (Level 1), outcomes (Level 2) and studies (Level 3; Assink & Wibbelink, 2016). This approach produces a robust analysis and has been successfully implemented in recent meta-analytical research (Harkin et al., 2023; Persson et al., 2021).

The current analysis was conducted using the rma.va function in the metafor package for the statistical software environment R (R Core Team, 2018) and recommendations of Viechtbauer (2010). A mixed-effects model was fitted, and estimation was based on the restricted maximum likelihood estimator. The analysis examined the variance distribution over the three levels, the overall effect (i.e., outcome measures for the intervention vs. the control condition) and the effects of a number of moderating variables. As recommended by Hox (2010) and Assink and Wibbelink (2016), moderators were first examined individually and then combined into one analysis. This allows for initial significance screening while also accounting for

TABLE 3 Characteristics of comparators and outcomes.

Study	Comparator (active vs. inactive)[a]	Total duration (min)[b]	Number of sessions[b]	Outcome measures	Follow up	Attrition, %
McCashin et al. (2022)	Waitlist (−)	—	—	CBCL	Postintervention, follow up	31.03
David and Fodor (2022)	Waitlist (−), REBT (+)	350	7	EATQ-R	Postintervention, 6-month follow up	11.11
Schoneveld et al. (2018)	Non-therapeutic game play (+)	300	5	SCAS	Postintervention, 3-month follow up	10.14
Schuurmans et al. (2018)	Treatment as usual (+)	—	—	SCAS	Postintervention	10
Schoneveld et al. (2018)	Coping cat (+)	540	8	SCAS	Postintervention, 3-month follow up, 6-month follow up	16
David et al. (2018)	Waitlist (−); REBT (+)	350	7	EATQ-R	Postintervention	11.11
David et al. (2018)	Waitlist (−); REBT (+)	350	7	EATQ-R	Postintervention	11.11
Sanchez et al. (2017)	Waitlist (−)	—	—	SASC-R	Postintervention	26
Kuosmanen et al. (2017)	Waitlist (−)	—	—	SMFQ, GAD-7	Postintervention	67.4
Scholten et al. (2016)	Non-therapeutic game play (+)	360	6	SCAS	Postintervention, 3-month follow up	6.5
Poppelaars et al. (2016)	Waitlist (−); school based CBT (+)	210	7	RADS-2	Postintervention, 3-month follow up, 6-month follow up, 12-month follow up	3.92
Roepke et al. (2015)	Waitlist (−); Non-therapeutic game play (+)	300	30	CES-D, GAD-7	Postintervention, 2 week follow up	78.49
Merry et al. (2012)	Treatment as usual (−)	—	—	CDRS-R, SCAS	Postintervention, 3-month follow up	9
Fleming et al. (2012)	Delayed intervention (−)	—	—	CDRS, SCAS	Postintervention	5

[a]Active comparator (+), inactive comparator (−).
[b]Information available for active comparators.

the possibility of variables of interest being intercorrelated, producing multicollinearity in the analyses.

5 | RESULTS OF META-ANALYSIS

5.1 | Data preparation

Based on Snyder et al. (2015), effect sizes 3 *SD* above or below the mean effect size ($d = .50$) were considered outliers and thus excluded. Based on this, one effect size from Fleming et al. (2012) was removed from further analysis. Thus, in total, 13 independent studies were included, totalling 18 effect sizes, with the mean number of effect sizes for each study at 1.4. It is important to note that within each of the moderators only the categories with three or more effect sizes were entered into the meta-analysis (see Harkin et al., 2016).

5.2 | Main and heterogeneity analyses

The first step of the analysis estimates the overall effect size for the difference between the TG intervention and the control condition, including 19 effect sizes from 13 individual studies. Overall, the main effect for gaming interventions on any outcome variable was small to medium sized, $d = .35$ (CI [.23, 47], $p < .001$). The second step of the analysis estimated the difference between within-study (Level 2) and between-study (Level 3) variance components, an important aspect of a three-level meta-analysis. This is assessed through two separate log-likelihood-ratio tests, where the original model (with freely estimated variance at Levels 2 and 3, respectively) is compared with one where the variance at each of the levels is fixed. The analyses suggested that there was significant variability ($p = .001$) between effect sizes (Level 2) and between studies (Level 3), indicating that moderator analyses should be conducted (Assink & Wibbelink, 2016). Based on formulas by Cheung (2014), the total variance distribution is as follows: Level

TABLE 4 Main and moderator analyses.

Variable	*k*	*d (SE)*	*p*	C−, C+	Q (*p*)
Main analysis	17	.35 (.06)**	<.001	.23, .47	461.14 (<.001)
Variable (categorical)					
Disorder type	17				
Depression	9	.40 (.08)[a]	<.001	.21, .58	451.27 (<.001)
Anxiety	8	.310 (.08)[a]	<.01	.13, .49	451.27 (<.001)
Avatar	17				
Embodied	11	.34 (.08)[a]	<.001	.18, .50	448.30 (<.001)
Guided/informational	6	.37 (.10)[a]	.002	.16, .59	448.30 (<.001)
Biofeedback	17				
Present	4	.18 (.10)	.09	−.03, .39	357.07 (<.001)
Not present	13	.42 (.60)[a]	<.001	.29, .55	357.07 (<.001)
Delivery	12				
Professional	2	.52 (.16)[a]	.01	.15, .88	316.98 (<.001)
Non-professional	9	.32 (.08)[a]	.004	13, 51	316.98 (<.001)
Nature of intervention	17				
RBE	5	.58 (.09)**	<.001	.39, .78	221.60 (<.001)
CBT	7	.26 (.07)*	.015	.12, .41	221.60 (<.001)
Other	5	.28 (.09)*	.031	.09, .47	221.60 (<.001)
Control condition	17				
Active	9	.28 (.06)	<.001	.15, .41	359.86 (<.001)
Passive	8	.43 (.06)	<.001	.29, .57	359.86 (<.001)
Variable (continuous)	*k*	*β (SE)*	*p*	C−, C+	Q (*p*)
Age	17	−.001 (.008)	.88	−.02, .02	459.30 (<.001)
Females (%)	17	−.003 (.003)	.42	−.009, .004	376.32 (<.001)
Total dosage	17	.001 (.001)	.57	-.00w, .003	449.23 (<.001)

Note: k = total number of studies included for each task. d (*SE*) = effect size in Cohen d (standard error); β = standardised beta (standard error). p = significance. C−, C+ = confidence intervals. Q (p) = q statistic.

[a]*F*-test non-significant (variable is only significantly different from zero).

*Significantly different from RBE at <.05.

**Significantly different from zero at <.001.

1, 3.92%; Level 2, 25.44%; Level 3, 70.64%. As recommended by Assink and Wibbelink (2016), moderation analyses should be conducted if less than 75% of the variance can be attributed to Level 1.

5.3 | Moderator analyses

The moderator analyses followed the following two-step approach. First, we first tested if the full moderation model for each moderator was significant, by entering all but one of the categories for each moderator—the moderator left out is used as the reference category, and all other categories are compared against this one.

Second, if the full model was significant (*F*-test), we can then look at each individual variable and see whether they differ from the reference category. If the full moderation model fails to attain significance, we do not look at the individual variables. We indicate significance of individual variables (in comparison to a reference category as specified) within a significant full moderation model by * and ** for a significant reference category in the *d* (*SE*) column of Table 4, respectively. This approach allows us to identify the moderators that have variables which differ from one another and ensures statistical integrity and avoidance of over-concluding on significant individual variables in the absence of a significant full moderation model.

The only moderator to reveal significant differences was the nature of intervention (coded: REBT, general CBT and 'other'). The overall moderation *F*-model is significant ($p = .03$), with mean effects of REBT = .58 (significantly larger than overall mean effect size; $p < .001$) whereas general CBT = .26 (significantly smaller than mean effect size; $p < .001$) and other interventions = .28 (significantly smaller than mean effect; $p = .015$). In other words, REBT had the largest mean effect, and the other two effects for general CBT and other interventions are significantly different to this. All other moderators that satisfied our analysis criteria were not significant ($p = .031$), that is, a type of disorder, total dosage, proportion females, age, who delivered intervention, biofeedback and avatar type.

5.4 | Publication bias and study quality

The Egger's regression coefficient using variance and standard error as independent moderators were not significant ($p < .05$), indicating no evidence of publication bias in this set of data points. To examine whether study quality was associated with overall effect sizes, a moderator analysis with methodological quality score was conducted and indicated that it was not associated with overall effect sizes ($p < .05$).

6 | DISCUSSION

For the first time in the literature, this systematic review has summarised the evidence for the use and effectiveness of TGs to reduce and prevent anxiety and depressive symptomology. We identified 14 studies encompassing seven TGs, involving general CBT (non-specific mode of CBT), REBT (a specific mode of CBT), social skill development and biofeedback techniques.

This review provides a unique contribution by conducting a meta-analysis that explores the previously unexamined link between specific aspects of TGs and their impact on psychological outcomes. The results indicate that while most existing TGs did not show significant improvement in mental health symptoms, there was a small to medium effect for TGs overall, suggesting that TGs focused on mental health could be beneficial.

Specifically, we found that TGs with a REBT focus outperformed CBT and other TGs in effectively reducing symptoms of depression and/or anxiety. In comparison to TGs based on a broader notion of CBT (i.e., general CBT), the games centred around REBT were highly specific to REBT theory and methods, explicitly targeting cognitive antecedents to and concomitants of (i.e., irrational beliefs; Ellis, 1994) anxiety and depression. It could be that the direct nature of the REBT games drove the superior effects of these games. REBT is also highly structured (Turner, 2022) and perhaps could be argued lends itself to being integrated into a serious game. This is an interesting finding and not one that we expected to find. However, there are some plausible reasons for this. First, REBT is a very particular CBT that focuses specifically on the articulation and weakening of irrational beliefs, and strengthening of rational beliefs. It also has a particular view on emotion, in that negative emotions can be adaptive or maladaptive depending on the underpinning beliefs and concomitant action tendencies. As such, the TGs that use REBT as their underlying framework have a specific focus and as such are perhaps more direct in their treatment of the theoretical antecedents to mental health issues. Second, the REBT interventions were on the higher end of the total duration. Indeed, each was 350 min, a high relative dose compared to the other interventions. Third, all REBT interventions were completed using an iPad. This may indicate a more sinuous exploitation method for participants to engage with compared to interventions using computers and traditional video gaming devices. Lastly, the REBT interventions used the same chief outcome measure, the EATQ-R, and therefore, we cannot discount that magnitude of effects could be influenced by the way the effect is measured.

Additionally, intriguing patterns emerged from the data, revealing that interventions delivered by professionals, that did not include biofeedback, as well as those utilising a passive control condition, yielded larger effect sizes compared to non-professionally delivered interventions, those employing biofeedback and active control conditions, respectively. The two studies which involved biofeedback utilised specific biofeedback techniques, namely, deep breathing, progressive muscle relaxation, positive thinking and guided imagery. Utilising biofeedback hardware (IOM, Wild Divine) to monitor heart rate variability, participant's heart rate directly impacted gameplay, with higher heart rates resulting in increasingly difficult game play. It must be noted that overall, the procedures that describe the development, utilisation and effectiveness of the included TGs lacked clarity. It is important to note that these patterns did not reach statistical significance, and we highlight them as potential fruitful areas for future research to investigate further. Additionally, it is crucial to highlight that the procedures detailing the development, utilisation and effectiveness of the included serious games lacked clarity.

We examined the effect of several moderators on the relationship between TG usage and psychological outcomes. The moderator analysis indicated that gender ratio did not influence the effect of TGs on psychological functioning. This is inconsistent with previous research which identified that females have greater favourable attitudes towards psychological treatment than males (Sheu & Sedlacek, 2004). A potential explanatory factor could be a result of the modality used within the therapeutic delivery, with research indicating males consistently engage more with game-based technologies (Clement, 2022). As a result, males may have had increased engagement with the therapeutic intervention through game-based methods than traditional psychotherapeutic delivery styles. Positively, the lack of effect of gender could indicate that TGs used for therapeutic benefit can aid all genders, particularly in light on increasing numbers of female gameplayers. We suggest that further research is needed to examine gender differences in the qualitative experiences and expectations of game-based interventions.

These findings are in line with earlier reviews of Abd-Alrazaq, Alajlani, et al. (2022) and Abd-Alrazaq, Al-Jafar, et al. (2022). While the findings were supported, the former reviews compared computerised CBT games to 'traditional' exergames, potentially conflating the specific therapeutic mechanisms embedded within 'serious games'. The current review centred mental health as the primary focus of included studies and only analysed research utilising TGs as opposed to the wider notion of serious games. The broader spectrum of game use includes exergames and massive multiplayer online games. Such games have been found to increase social wellbeing, reducing loneliness and increasing people's ability to form connections (Jones et al., 2014; Li et al., 2018). However, any improvement in participant's mental health were viewed as secondary outcomes, with the game content and underlying mechanisms not being designed for the purpose of therapeutic intervention.

The studies included in our review highlight several directions for future development of serious games as a therapeutic tool. Most of the games considered appeared to focus predominantly on one

specific domain (cognitive, emotional or behavioural). For example, 'Pesky gNATs' seeks to address negative automatic thoughts, while 'Dojo' is classified as 'emotion management game' (Schuurmans et al., 2015, p. 401). To produce tangible, long-lasting effects on anxiety and depressive symptoms, TGs will need to include content that encompasses all areas (on a cognitive, emotional and behavioural level) underpinning the current symptomology. That is, to be optimally effective, TGs will need to offer gamified content aimed at stimulating awareness and eliciting change in cognitive, emotional and behavioural patterns that underpin anxiety and/or depression.

Moreover, literature on internet-based CBT (e.g., Newby et al., 2021) has pointed out the need for digital CBT interventions to include a variety of CBT techniques and strategies to promote symptomatic reduction and relapse prevention, including but not limited to self-monitoring (e.g., monitoring of thoughts, feelings and tasks), challenging negative automatic thoughts and problem-solving activities. To the best of our knowledge, to date, there is no CBT-based serious game that offers users a wide range of CBT contents, making this one of the directions for future development of CBT-based games.

One important procedural inconsistency of the studies included in this review that needs to be considered is the wide variety in game dosages employed. There was large variation in terms of game length and number of game sessions, which calls into question the dosage needed, or recommended, for TGs to be effective. It is not as simple as 'more is better' because there is a limit to the extent to which games will remain interesting to game users (i.e., patients). Based on the studies included, seven sessions of around 30–45 min appears to be the typical dose, but it is not possible from our analyses to discern how much of that 210–315 min is constituted of therapeutic content (vs. world building, exploring and action). As such, it is not clear as to the specific dosage required to maximise the ameliorating effects of TGs on anxiety and depression.

An additional aspect that needs to be addressed further is the need to tailor the game content and layout to the needs of neurodiverse populations. For example, the general CBT-based game for anxiety reduction 'MindLight' showed a decrease in child-rated anxiety symptoms when tested against a control condition in neurotypical participants, but no differences in anxiety levels were found in a sample of children with autism spectrum disorder (Wijnhoven et al., 2022). The need to adapt TGs to different neurocognitive profiles and abilities has been advocated by Terras et al. (2018) who mentioned the need for serious games to be 'psychologically accessible' for users presenting with learning disabilities and/or neurodiversity. One of the proposed ways in which this can be achieved is the involvement, in the early stages of game ideation and throughout game development, of 'experts by experience' whose inclusion can not only foster empowerment and inclusion but also improve game usefulness and usability (Politis et al., 2017). This is in line with what Fleming et al. (2017) suggested as they outline the role of co-production as a crucial aspect to maximise the potential of serious games in clinical settings.

Various explanations are proposed as to the lack of significant impact in regard to game design, the intervention deliverer (e.g., teacher or self-delivery), and style of avatar within gameplay on psychological outcomes. Potentially, the most important is the lack of a therapeutic alliance developed as is commonly effective within traditional psychotherapeutic delivery (Catty, 2004). Recent narrative reviews have highlighted the importance of combining digital health interventions with human support in a bid to cultivate digital alliances, thereby promoting engagement and effectiveness (Tremain et al., 2020). Furthermore, the lack of ability to personalise the TG potentially resulted in a reduction in engagement, thereby minimising the impact of the serious game overall as users' ability to relate to the avatars is limited overall. Future TG design would benefit from developing adaptable games to improve relatedness to the intervention.

6.1 | Limitations and future research

The current meta-analysis identified several shortcomings in the available studies. Overall, the quality of the included studies was medium-to-poor (Table 1), with limited information available regarding how ingrained the therapeutic modalities were to the design of the TG. Furthermore, the research design resulted in the intervention being used in a 'noisy' environment (e.g., poor intervention methods, multiple potential confounds and low study numbers) which reduced the impact of the intervention on the outcomes. For example, we noted that effect sizes were greater when interventions were administered by professionals rather than non-professionals and when active control conditions were employed instead of passive control conditions. Such confounding factors pose challenges in isolating the influence of individual moderators in relation to others, underscoring the necessity for future research to precisely identify and control specific variables in randomised control trials and laboratory settings. These controlled methods will hopefully help isolate key variables such as those we identify and ascertain straightforward causal mechanisms. Future serious game development would benefit from simple interventions which monitor progress, frequency of intervention use (including dosage) and repeatedly targeting core areas of cognition. Overall, the findings suggested that psychotherapeutic interventions utilising TGs for mental ill health are promising, but replications are warranted, with further attention being placed on addressing methodological concerns within the research.

CONFLICT OF INTEREST STATEMENT

No potential conflict of interest was reported by the authors.

ORCID

Zarah Eve https://orcid.org/0000-0002-3945-2042
Daniela Di Basilio https://orcid.org/0000-0003-3786-442X

REFERENCES

Abd-Alrazaq, A., Alajlani, M., Alhuwail, D., Schneider, J., Akhu-Zaheya, L., Ahmed, A., & Househ, M. (2022). The effectiveness of serious games in alleviating anxiety: Systematic review and meta-analysis. *JMIR Serious Games*, *10*(1), e29137. https://doi.org/10.2196/29137

Abd-Alrazaq, A., Al-Jafar, E., Alajlani, M., Toro, C., Alhuwail, D., Ahmed, A., Reagu, S. M., Al-Shorbaji, N., & Househ, M. (2022). The effectiveness of serious games for alleviating depression: Systematic review and meta-analysis. *JMIR Serious Games*, *10*(1), e32331. https://doi.org/10.2196/32331

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. (2008). A comprehensive review and meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. *Journal of Technology in Human Services*, *26*(2–4), 109–160. https://doi.org/10.1080/15228830802094429

Brenes, G. A., Danhauer, S. C., Lyles, M. E., Hogan, P. E., & Miller, M. E. (2015). Telephone-delivered cognitive-behavioral therapy and telephone-delivered nondirective supportive therapy for rural older adults with generalised anxiety disorder: A randomized clinical trial. *JAMA Psychiatry*, *72*(10), 1012–1020. https://doi.org/10.1001/jamapsychiatry.2015.1154

Brown, D. (2013). Overcoming the barriers to uptake: A study of 6 Danish health-based serious games projects. *International Journal of Game-Based Learning*, *3*(3), 1–9. https://doi.org/10.4018/ijgbl.2013070101

Catty, J. (2004). 'The vehicle of success': Theoretical and empirical perspectives on the therapeutic alliance in psychotherapy and psychiatry. *Psychology and Psychotherapy: Theory, Research and Practice*, *77*(2), 255–272. https://doi.org/10.1348/147608304323112528

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modelling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968

Christ, C., Schouten, M. J., Blankers, M., Van Schaik, D. J., Beekman, A. T., Wisman, M. A., Stikkelbroek, Y. A., & Dekker, J. J. (2020). Internet and computer-based cognitive behavioral therapy for anxiety and depression in adolescents and young adults: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *22*(9), e17831. https://doi.org/10.2196/17831

Clement, J. (2022). *Video gaming audiences in the United Kingdom—Statistics and facts. (online, available from video gaming audiences in the United Kingdom (UK)—Statistics & facts|Statista)* [Accessed 1/3/2023].

Creswell, C., Shum, A., Pearcey, S., Skripkauskaite, S., Patalay, P., & Waite, P. (2021). Young people's mental health during the COVID-19 pandemic. *The Lancet: Child and Adolescent Health*, *5*(8), 535–537. https://doi.org/10.1016/S2352-4642(21)00177-2

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behaviour therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry*, *15*(3), 245–258. https://doi.org/10.1002/wps.20346

David, O. A., Cardoş, R. A., & Matu, S. (2018). *Is REThink therapeutic game effective in preventing emotional disorders in children and adolescents? Outcomes of a randomized clinical trial* (Vol. 28) (pp. 111–122). European Child & Adolescent Psychiatry. https://doi.org/10.1007/s00787-018-1192-2

David, O. A., Cardoş, R. A. I., & Matu, S. (2018). Changes in irrational beliefs are responsible for the efficacy of the REThink therapeutic game in preventing emotional disorders in children and adolescents: Mechanisms of change analysis of a randomized clinical trial. *European Child & Adolescent Psychiatry*, *28*, 307–318. https://doi.org/10.1007/s00787-018-1195-z

David, O. A., & Fodor, L. A. (2022). Are gains in emotional symptoms and emotion-regulation competencies after the REThink therapeutic game maintained in the long run? A 6-month follow-up. *European Child & Adolescent Psychiatry*, *32*, 1853–1862. https://doi.org/10.1007/s00787-022-02002-w

Dealessandri, M. (2023). ERA: Games remained biggest home entertainment industry in the UK in 2022. In *Games industry (online, available from ERA: Games remained biggest home entertainment industry in the UK in 2022)*, GamesIndustry.biz. Accessed 1/3/23.

Ellis, A. (1994). *Reason and emotion in psychotherapy*. Birch Lane Press.

Ellis, A. (1995). Changing rational-emotive therapy (RET) to rational emotive behavior therapy (REBT). *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *13*(2), 85–89. https://doi.org/10.1007/BF02354453

Fleming, T. M., Bavin, L., Stasiak, K., Hermansson-Webb, E., Merry, S. N., Cheek, C., Lucassen, M., Lau, H. M., Pollmuller, B., & Hetrick, S. (2017). Serious games and gamification for mental health: Current status and promising directions. *Frontiers in Psychiatry*, *7*, 215. https://doi.org/10.3389/fpsyt.2016.00215

Fleming, T., Dixon, R., Frampton, C., & Merry, S. (2012). A pragmatic randomized controlled trial of computerized CBT (SPARX) for symptoms of depression among adolescents excluded from mainstream education. *Behavioral and Cognitive Psychotherapy*, *40*, 529–541. https://doi.org/10.1017/S1352465811000695

Grist, R., Porter, J., & Stallard, P. (2017). Mental health mobile apps for preadolescents and adolescents: A systematic review. *Journal of Medical Internet Research*, *19*(5), e176. https://doi.org/10.2196/jmir.7332

Harkin, B., Persson, S., Yates, A., Zinkunegi, A., & Kessler, K. (2023). Top-down and bottom-up contributions to memory performance in OCD: A multilevel meta-analysis with clinical implications. *Journal of Psychopathology and Clinical Science.*, *132*, 428–444. https://doi.org/10.1037/abn0000793

Harkin, B., Webb, T. L., Chang, B. P. I., Prestwich, A., Conner, M., Kellar, I., Benn, Y., & Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence? *Psychological Bulletin*, *142*(2), 198–229. https://doi.org/10.1037/bul0000025

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780203852279

Huang, K. Y., Lee, D., Nakigudde, J., Cheng, S., Gouley, K. K., Mann, D., Schoenthaler, A., Chokshi, S., Kisakye, E. N., Tusiime, C., & Mendelsohn, A. (2019). Use of technology to promote child behavioral health in the context of pediatric care: A scoping review and applications to low-and middle-income countries. *Frontiers in Psychiatry*, *10*, 806. https://doi.org/10.3389/fpsyt.2019.00806

Hudlicka, E. (2016). Virtual affective agents and therapeutic games. In *Artificial intelligence in behavioural and mental health care* (pp. 81–115). Academic press. https://doi.org/10.1016/B978-0-12-420248-1.00004-0

International Telecommunications Union. (2016). *ICT facts and figures 2016*. Retrieved from https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2016.pdf

Jones, C. M., Scholes, L., Johnson, D., Katsiktis, M., & Carras, M. C. (2014). Gaming well: Links between videogames and flourishing mental health. *Frontiers in Psychology*, *5*(260), 1–8, 260. https://doi.org/10.3389/fpsyg.2014.00260

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Ustun, T. B. (2007). Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry*, *20*(4), 359–364. https://doi.org/10.1097/YCO.0b013e32816ebc8c

Krzyzaniak, N., Greenwood, H., Scott, A. M., Peiris, R., Cardona, M., Clark, J., & Glasziou, P. (2021). The effectiveness of telehealth versus face-to face interventions for anxiety disorders: A systematic review and meta-analysis. *Journal of Telemedicine and Telecare*, 1357633X211053738. https://doi.org/10.1177/1357633x211053738

Kuosmanen, T., Fleming, T. M., Newell, J., & Barry, M. M. (2017). A pilot evaluation of the SPARX-R gaming intervention for preventing depression and improving wellbeing among adolescents in alternative education. *Internet Interventions*, *8*, 40–47. https://doi.org/10.1016/j.invent.2017.03.004

Landreth, G. L. (2012). *Play therapy: The art of the relationship*. Routledge. https://doi.org/10.4324/9780203835159

Lau, H. M., Smit, J. H., Fleming, T. M., & Riper, H. (2017). Serious games for mental health: Are they accessible, feasible, and effective? A systematic review and meta-analysis. *Frontiers in Psychiatry*, *7*, 209. https://doi.org/10.3389/fpsyt.2016.00209

Li, J., Erdt, M., Chen, L., Cao, Y., Lee, S.-Q., & Theng, Y.-L. (2018). The social effects of exergames on older adults: Systematic review and metric analysis. *Journal of Medical Internet Research*, *20*(6), e10486. https://doi.org/10.2196/10486

Liverpool, S., Mota, C. P., Sales, C. M., Čuš, A., Carletto, S., Hancheva, C., Sousa, S., Cerón, S. C., Moreno-Peral, P., Pietrabissa, G., Moltrecht, B., Ulberg, R., Ferreira, N., & Edbrooke-Childs, J. (2020). Engaging children and young people in digital mental health interventions: Systematic review of modes of delivery, facilitators, and barriers. *Journal of Medical Internet Research*, *22*(6), e16317. https://doi.org/10.2196/16317

Martins, T., Carvalho, V., Soares, F., & Moreira, M. F. (2011). Serious game as a tool to intellectual disabilities therapy: Total challenge. In *Serious games and applications for health (SeGAH), 2011 IEEE 1st international conference on* (pp. 1–7). IEEE.

McCallum, S. (2012). Gamification and serious games for personalized health. *PHealth*, *177*, 85–96.

McCashin, D., Coyle, D., & O'Reilly, G. (2022). Pesky gNATs for children experiencing low mood and anxiety—A pragmatic randomised controlled trial of technology-assisted CBT in primary care. *Internet Interventions*, *27*, 100489. https://doi.org/10.1016/j.invent.2021.100489

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. G. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *BMJ*, *344*(e2598), 1–16, e2598. https://doi.org/10.1136/bmj.e2598

Miller, A. S., Cafazzo, J. A., & Seto, E. (2014). A game plan: Gamification design principles in mHealth applications for chronic disease management. *Health Information Journal*, *22*, 184–193. https://doi.org/10.1177/1460458214537511

Neary, M., & Schueller, S. M. (2018). State of the field of mental health apps. *Cognitive and Behavioral Practice*, *25*(4), 531–537. https://doi.org/10.1016/j.cbpra.2018.01.002

Newby, J., Mason, E., Kladnistki, N., Murphy, M., Millard, M., Haskelberg, H., Allen, A., & Mahoney, A. (2021). Integrating internet CBT into clinical practice: A practical guide for clinicians. *Clinical Psychologist*, *25*(2), 164–178. https://doi.org/10.1080/13284207.2020.1843968

NHS Digital. (2021a). *Mental health services monthly statistics*. https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-services-monthly-statistics

NHS Digital. (2021b). Mental health of children and young people in England 2021—Wave 2 follow up of the 2017 survey. In *Mental health of children and young people surveys*. https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2021-follow-up-to-the-2017-survey

O'Connor, S., Hanlon, P., O'Donnell, C. A., Garcia, S., Glanville, J., & Mair, F. S. (2016). Understanding factors affecting patient and public engagement and recruitment to digital health interventions: A systematic review of qualitative studies. *BMC Medical Informatics and Decision Making*, *16*, 120. https://doi.org/10.1186/s12911-016-0359-3

Ofcom. (2022). *Adult media literacy core survey – 2021*. (Online, available from Adults' Media Literacy Core survey 2021 data tables (ofcom.org.uk) [Accessed 2/3/2023]).

Paige, S. R., Miller, M. D., Krieger, J. L., Stellefson, M., & Cheong, J. (2018). Electronic health literacy across the lifespan: Measurement invariance study. *Journal of Medical Internet Research*, *20*(7), e10434. https://doi.org/10.2196/10434

Persson, S., Yates, A., Kessler, K., & Harkin, B. (2021). Modeling a multidimensional model of memory performance in obsessive-compulsive disorder: A multilevel meta-analytic review. *Journal of Abnormal Psychology*, *130*(4), 346–364. https://doi.org/10.1037/abn0000660

Politis, Y., Robb, N., Yakkundi, A., Dillenburger, K., Herbertson, N., Charlesworth, B., & Goodman, L. (2017). People with disabilities leading the design of serious games and virtual worlds. *International Journal of Serious Games*, *4*(2), 63–73. https://doi.org/10.17083/ijsg.v4i2.160

Poppelaars, M., Tak, Y. R., Lichtwarck-Aschoff, A., Engels, R. C., Lobel, A., Merry, S. N., Lucassen, M. F., & Granic, I. (2016). A randomized controlled trial comparing two cognitive-behavioral programs for adolescent girls with subclinical depression: A school-based program (Op Volle Kracht) and a computerized program (SPARX). *Behaviour Research and Therapy*, *80*, 33–42. https://doi.org/10.1016/j.brat.2016.03.005

Protopsaltis, A., Pannese, L., Pappa, D., & Hetzner, S. (2011). Serious games and formal and informal learning. In *E-learning papers* 1887, 1542.

Punton, G., Dodd, A. L., & Mcneill, A. (2022). 'You're on the waiting list': An interpretive phenomenological analysis of young adults' experiences of waiting lists within mental health services in the UK. *PLoS ONE*, *17*(3), e0265542. https://doi.org/10.1371/journal.pone.0265542

R CoreTeam. (2018). *R: A language and environment for statistical computing*.

Radez, J., Reardon, T., Creswell, C., Lawrence, P. J., Evdoka-Burton, G., & Waite, P. (2021). Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. *European Child & Adolescent Psychiatry*, *30*, 183–211. https://doi.org/10.1007/s00787-019-01469-4

Rasing, S. P. A. (2021). Blended treatment for depressive disorders in youth: A narrative review. *International Journal of Cognitive Therapy*, *14*(1), 47–58. https://doi.org/10.1007/s41811-020-00088-1

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, *32*(4), 251–262. https://doi.org/10.1016/j.cpr.2012.01.005

Roepke, A. M., Jaffee, S. R., Riffle, O. M., McGonigal, J., Broome, R., & Maxwell, B. (2015). Randomized controlled trial of SuperBetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. *Games for Health Journal*, *235-246*, 235–246. https://doi.org/10.1089/g4h.2014.0046

Sailer, M., Hence, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behaviour*, *69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Sanchez, R., Brown, E., Kocker, K., & DeRosier, M. (2017). Improving children's mental health with a digital social skills development game: A randomized controlled efficacy trial of adventures aboard the S.S. GRIN. *Games for Health Journal*, *19-27*, 19–27. https://doi.org/10.1089/g4h.2015.0108

Scholten, H., Malmberg, M., Lobel, A., Engels, R. C., & Granic, I. (2016). A randomized controlled trial to test the effectiveness of an immersive 3D video game for anxiety prevention among adolescents. *PLoS ONE*, *11*(1), e0147763. https://doi.org/10.1371/journal.pone.0147763

Schoneveld, E. A., Lichtwarck-Aschoff, A., & Granic, I. (2018). Preventing childhood anxiety disorders: Is an applied game as effective as a cognitive behavioral therapy-based program? *Preventative Science*, *19*, 220–232. https://doi.org/10.1007/s11121-017-0843-8

Schoneveld, E. A., Malmberg, M., Lichtwarck-Aschoff, A., Verheijen, G. P., Engels, R. C., & Granic, I. (2016). A neurofeedback video game (MindLight) to prevent anxiety in children: A randomized controlled trial. *Computers in Human Behavior*, *16*, 321–333. https://doi.org/10.1016/j.chb.2016.05.005

Schuurmans, A. A. T., Nijhof, K. S., Engels, R. C., & Granic, I. (2018). Using a videogame intervention to reduce anxiety and externalizing problems among youths in residential care: An initial randomized controlled trial. *Journal of Psychopathological Behaviour Assessment*, *40*, 344–354. https://doi.org/10.1007/s10862-017-9638-2

Schuurmans, A. A., Nijhof, K. S., Vermaes, I. P., Engels, R. C., & Granic, I. (2015). A pilot study evaluating "Dojo", a videogame intervention for youths with externalising and anxiety problems. *Games for Health Journal*, *4*(5), 401–408. https://doi.org/10.1089/g4h.2014.0138

Seaborn, K., & Fels, D. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Sheu, H. -B., & Sedlacek, W. H. (2004). An exploratory study of help-seeking attitudes and coping strategies among college students by race and gender. *Measurement and Evaluation in Counseling and Development*, *37*(3), 130–143. https://doi.org/10.1080/07481756.2004.11909755

Sicart, M. (2014). *Playing the good life: Gamification and ethics. The Gameful world* (pp. 228–244). MIT Press.

Snoswell, C. L., Chelberg, G., De Guzman, K. R., Haydon, H. H., Thomas, E. E., Caffery, L. J., & Smith, A. C. (2021). The clinical effectiveness of telehealth: A systematic review of meta-analyses from 2010 to 2019. *Journal of Telemedicine and Telecare*.

Snyder, H. R., Kaiser, R. H., Warren, S. L., & Heller, W. (2015). Obsessive-compulsive disorder is associated with broad impairments in executive function: A meta-analysis. *Clinical Psychological Science*, *3*(2), 301–330. https://doi.org/10.1177/2167702614534210

Terras, M. M., Boyle, E. A., Ramsay, J., & Jarrett, D. (2018). The opportunities and challenges of serious games for people with an intellectual disability. *British Journal of Educational Technology*, *49*(4), 690–700. https://doi.org/10.1111/bjet.12638

Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (1998). Effective public health practice project (EPHPP) quality assessment tool for quantitative studies. In *Effective public health practice project*. https://merst.ca/ephpp/

Tremain, H., McEnery, C., Fletcher, K., & Murray, G. (2020). The therapeutic alliance in digital mental health interventions for serious mental illnesses: Narrative review. *JMIR Mental Health*, *7*(8), e17204. https://doi.org/10.2196/17204

Turner, M. J. (2022). *The rational practitioner: The sport and performance psychologist's guide to practicing rational emotive behaviour therapy*. Routledge. https://doi.org/10.4324/9781003200437

Välimäki, M., Anttila, K., Anttila, M., & Lahti, M. (2017). Web-based interventions supporting adolescents and young people with depressive symptoms: Systematic review and meta-analysis. *JMIR mHealth and uHealth*, *5*(12), e8624. https://doi.org/10.2196/mhealth.8624

Van der Lubbe, L. M., Gerritsen, C., Klein, M. C. A., & Hindriks, K. V. (2021). Empowering vulnerable target groups with serious games and gamification. *Entertainment Computing*, *38*, 100402. https://doi.org/10.1016/j.entcom.2020.100402

Viechtbauer, W. (2010). Conducting meta-analyses in R with metfor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viner, R., Russell, S., Saulle, R., Croker, H., Stansfeld, C., Packer, J., Nicholls, D., Goddings, A.-L., Bonell, C., Hudson, L., Hope, S., Schwalbe, N., Morgan, A., & Minozzi, S. (2021). Impacts of school closure on physical and mental health of children and young people: A systematic review. *MedRxiv*. https://doi.org/10.1101/2021.02.10.21251526

Wijnhoven, L. A., Engels, R. C., Onghena, P., Otten, R., & Creemers, D. H. (2022). The addictive effect of CBT elements on the video game 'Mindlight' in decreasing anxiety symptoms of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *52*, 150–168. https://doi.org/10.1007/s10803-021-04927-8

Wilkinson, P. (2016). A brief history of serious games. In R. Dörner, S. Göbel, M. Kickmeier-Rust, M. Masuch, & K. Zweig (Eds.), *Entertainment computing and serious games* (pp. 17–41). https://doi.org/10.1007/978-3-319-46152-6_2

How to cite this article: Eve, Z., Turner, M., Di Basilio, D., Harkin, B., Yates, A., Persson, S., Henry, J., Williams, A., Walton, G., Jones, M. V., Whitley, C., & Craddock, N. (2023). Therapeutic games to reduce anxiety and depression in young people: A systematic review and exploratory meta-analysis of their use and effectiveness. *Clinical Psychology & Psychotherapy*, 1–14. https://doi.org/10.1002/cpp.2938