Child & Youth Care Forum (2020) 49:817–838
https://doi.org/10.1007/s10566-020-09566-1

REVIEW

How Effective are Serious Games for Promoting Mental Health and Health Behavioral Change in Children and Adolescents? A Systematic Review and Meta-analysis

Oana Alexandra David[1] · **Cristina Costescu**[1] · **Roxana Cardos**[1] · **Cristina Mogoaşe**[1]

Published online: 31 July 2020

Abstract

Background Children and adolescents are major computer and technology gadget users. While serious games hold important promises for promoting health-related behavioral change and mental health among children and adolescents, their efficacy is yet unclear.

Objectives We conducted a systematic review and meta-analysis aimed to offer a comprehensive picture on the evidence-based status of serious games for mental health promotion and health-related behavioral change in children and adolescents.

Method We included 34 clinical and experimental randomized studies investigating the efficacy of serious games on mental health promotion and health-related behavioral change in children and adolescents.

Results Results showed a small, but significant effect size with very high heterogeneity. Participants' age, number of sessions, the length of session, and study quality significantly moderated the effect size, with younger participants, fewer and shorter sessions, and lower quality of the study being associated with higher effect sizes. The effect size was not significant for follow-up measurements.

Conclusions The evidence supporting the use of serious games in children and adolescents for purposes of health promotion is limited. These conclusions must be considered with caution given the bias of publication. We need more adequately conducted studies testing well-specified serious games before we can draw clear conclusions.

Keywords Systematic review · Therapeutic games · Children and adolescent · Mental health

✉ Oana Alexandra David
oanadavid@psychology.ro

[1] Department of Clinical Psychology and Psychotherapy, The International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health, Babeș-Bolyai University, 37 Republicii St., 400015 Cluj-Napoca, Romania

Introduction

Given the high prevalence of psychopathology in children and adolescents (Collishaw et al. 2010), the long-term negative consequences of early onset psychopathology for later functioning (Goodman et al. 2011), and the gap between the number of youth in need for treatment and those who actually receive adequate treatment (Knopf et al. 2008; Merikangas et al. 2010), offering effective therapeutic interventions is a necessity. Moreover, it is known that children and adolescents presenting chronic medical illness (e.g. diabetes, asthma) have a significantly higher risk for the development of psychiatric disorders, compared to their healthy counterpart (Pao and Bosk 2011). Currently, cognitive-behavioral therapy (CBT) is considered the evidence-based choice for approaching a variety of mental health (such as depression, anxiety, attention deficit hyperactivity disorder—ADHD, or autism spectrum disorders etc.) in children and adolescents (Compton et al. 2004); see also The National Institute for Health and Care Excellence [NICE] guidelines). Yet, despite the fact that effective treatments do exist, most of the youth in need for therapeutic intervention will never receive adequate treatment due to access barriers (Farrell and Barrett 2007).

One of the most suitable and affordable way of increasing treatment access is to make use of the recent technological development and wide Internet spreading. Up to 20% of adolescents use the Internet to seek help for emotional problems (Gould et al. 2002). More specifically, Horgan and Sweeney (Horgan and Sweeney 2010) found that 267 out of 867 young people (30.8%) used the Internet to search for mental health information and 593 out of 872 (68%) indicated they would use the Internet if they needed to. Notably, seeking help from the Internet overcomes the gender divide, with boys being as likely to seek help as girls (although otherwise boys are significantly less likely to seek help; Coyle et al. 2009; Gould et al. 2002).

Computer-based and online cognitive-behavioral therapy (cCBT/iCBT) intervention protocols are already available and have been found to yield comparable results with traditional, face-to-face CBT, both in adult and youth samples (Newman et al. 2011; Przeworski and Newman 2006). iCBT involves online delivery of the CBT intervention protocol, via computer technologies (PC, mobiles, tablets), either at home or in healthcare settings (Cartreine et al. 2010).

Serious Games

Given all the iCBT's advantages and the encouraging results obtained with online interventions, researchers developed therapeutic (online) games. Therapeutic games are considered serious games (i.e. entertaining games with non-entertainment goals; (Ceranoglu 2010) that capitalize on the general fascination children and teenagers have for computer games in order to facilitate the implementation of therapeutic interventions for this age category. Online games rank as the most frequent entertainment form among adolescents (Gentile and Walsh 2002), with more than half of the eleven to fourteen years old playing games four times a week or more (McFarlane et al. 2002). In the United States alone, 91% of children aged 2 to 17 play video games, and 97% of them are playing for at least one hour per day (Granic et al. 2014; NDP Group 2011). Therefore, video-games should be considered a suitable media support to reach children and adolescents in need for therapeutic services.

Therapeutic computer games entail a series of advantages over the traditional psychotherapeutic approaches, including complex and interactive environments. As one of the most challenging task faced by clinicians working with children and adolescences is to engage youths in the therapeutic process and to maintain them motivated (Crenshaw 2008), the feature of interactivity incorporated in the gamification process might make therapeutic games uniquely equipped tools for promoting mental health among youths. Within therapeutic online games, children and adolescents can be exposed to contingency-based learning, as well as to the opportunity of acquiring procedural skills they can vicariously practice (Kato et al. 2008). In addition, therapeutic computer games have the advantages of being used both as a *prevention strategy* and/or as an *adjunctive clinical tool* in addressing psychopathology or health related behavioral change.

There are several reviews in the literature describing various initiatives in developing such therapeutic games for approaching different mental health and chronic medical conditions in children. Some of these reviews document promising initial results (Arns et al. 2009; Horne-Moyer et al. 2014; Primack et al. 2012). However, most of the existing reviews on therapeutic games for child psychopathology are qualitative (Granic et al. 2014; Wilkinson et al. 2008), due to the lack of rigorous studies and to the fact that the majority of articles present mostly preliminary results. A more encouraging situation can be found in the case of therapeutic games addressing health-related behaviors, where a wider literature of studies using sound methodologies is available (Baranowski et al. 2008; DeSmet et al. 2014; DeSmet et al. 2014; Primack et al. 2012). A recent meta-analysis including 54 studies (DeSmet et al. 2014) showed that serious games for promoting healthy lifestyle and clinical outcomes are as effective as other computer-delivered interventions, bringing small but significant post-treatment effects, which, with the exception of behavioral changes, are maintained in the long term. Reviews showed positive effects of therapeutic games for specific medical or psychological conditions, like diabetes (DeShazo et al. 2010), obesity (Guy et al. 2011), sexual health promotion (DeSmet et al. 2014), or ADHD (Arns et al. 2009). However, most reviews noted large variations between studies (Baranowski et al. 2008; DeShazo et al. 2010; Gamberini et al. 2008; Guse et al. 2012; Guy et al. 2011; Kato et al. 2008; Kharrazi et al. 2012; Lu et al. 2012; Papastergiou 2009; Primack et al. 2012; Rahmani and Boren 2012; Wilkinson et al. 2008), some suggesting that investigating further the game features associated with large effects could potentially explain the heterogeneity.

As noted elsewhere (Goh et al. 2008), the available research on serious gaming has largely focused on playability and acceptability of games, while omitting to consider their clinical utility in respect to behavioral change and mental health aspects. Most of the reviews published up to date are qualitative in nature. Among the quantitative reviews, moderation analyses showed that serious games are appealing to populations regardless of age and gender, and are effective in healthy lifestyle promotion whether stand-alone or part of multi-component programs (DeSmet et al. 2014). However, no review has specifically focused on the efficacy of serious games addressing various mental health and health-related conditions in youth population. This is unfortunate, as (1) children and adolescents are important consumers of computerized games in this digital era, (2) they are among the most vulnerable for mental disorders (as they do not own fully developed adaptive coping strategies), with important implications—in terms of emotional, financial, and social burden—for the rest of their life, and (3) therapeutic computerized games entail a series of advantages over the traditional psychotherapeutic approach for youth, including complex and interactive environments that intrinsically promotes engagement, platforms already familiar to players, wide accessibility and easiness to use. We believe that a special focus on the clinical utility of serious games for children and adolescents is timely and important, as the extent to which youth can benefit from serious games is highly dependent on their developmental particularities. Previous research work highlighted that little is known about the types of games design elements that appeal to children and adolescents during the developmental periods, although designing developmentally appropriate serious games has the potential to enhance the effectiveness and to facilitate transfer to real-life behaviors (Baranowski et al. 2016). To our knowledge, there are only two previous benchmark qualitative reviews of the clinical utility of serious games in youth. The first, authored by Granic and collaborators (Granic et al. 2014), provided evidence and theoretical rationale for the clinical utility of serious games for children and adolescents, highlighting promising research directions for testing the clinical potential of serious games for this age category. The second, authored by Baronowky and collaboratores (Baranowski et al. 2016), is a white paper highlighting the need for additional research to determine the game design and behavior change procedures that promote the effectiveness of serious games, while identifying and minimizing potential adverse effects. With this meta-analysis, we aimed to contribute to this desiderate, by employing a systematic approach and documenting the effect size of serious games for mental health and health-related behavior. In addition, we aimed to document the game features that might contribute to increased outcomes within this specific population. In this respect, we investigated a range of possible variables that can moderate the clinical utility of the serious games in children and adolescents, including age, game technology, type of game (educational versus therapeutic), type of sample (healthy or suffering from a specific health condition), and gaming length. Another important contribution of this meta-analysis resides in the fact that, in order to provide a clean overview of the serious games clinical utility for youth, we considered the methodological characteristics of the studies, as well as the overall studies quality. In terms of the methodological characteristics, we differentiated between experimental studies and clinical randomized studies, as the results of the latest might entail higher ecological validity. Regarding the studies quality, we believe it is critical for understanding their results and for guiding the further research progress in the field.

Method

Literature Search

We conducted a literature search, focused on intervention studies published between 1990 and January 2018 in the databases PubMed, PsycINFO, and CINAHL. We used the following search terms: "(computer AND therapeutic game) AND (children OR adolescents)"; "(video game AND psychotherapy) AND (children OR adolescents)"; "(computer game AND psychotherapy) AND (children OR adolescents)"; "(online game AND psychotherapy) AND (children OR adolescents)"; "(mobile game AND psychotherapy) AND (children OR adolescents)"; "(serious game AND psychotherapy) AND (children OR adolescents)"; "(gaming AND psychotherapy) AND (children OR adolescents)"; "(gaming AND psychological intervention) AND (children OR adolescents)"and "(online game AND psychological intervention) AND (children OR adolescents)". We also examined bibliographical references of recent systematic reviews, meta-analyses and primary intervention studies.

Selection of Studies

The search strategy produced a total of 447 recordings. After removing duplicates and irrelevant entries, a total of 135 full-text articles remained to be assessed for eligibility. To be included in this review, studies must have met the following criteria: (a) randomized controlled study, that (b) was aimed to examine the efficacy of a serious game intended to promote mental health or health-related behavioral change; (c) study participants were children and/or adolescents (i.e., up to 18 years old), and (d) receive no other active intervention during participating in the study; (e) the game was implemented in a technology-based format (computer, virtual reality system, tablet or smartphone); and (f) sufficient data were reported to allow effect size calculation. Based on these criteria, we included a total of 34 studies (please see Fig. 1 for the PRISMA Flow Diagram).

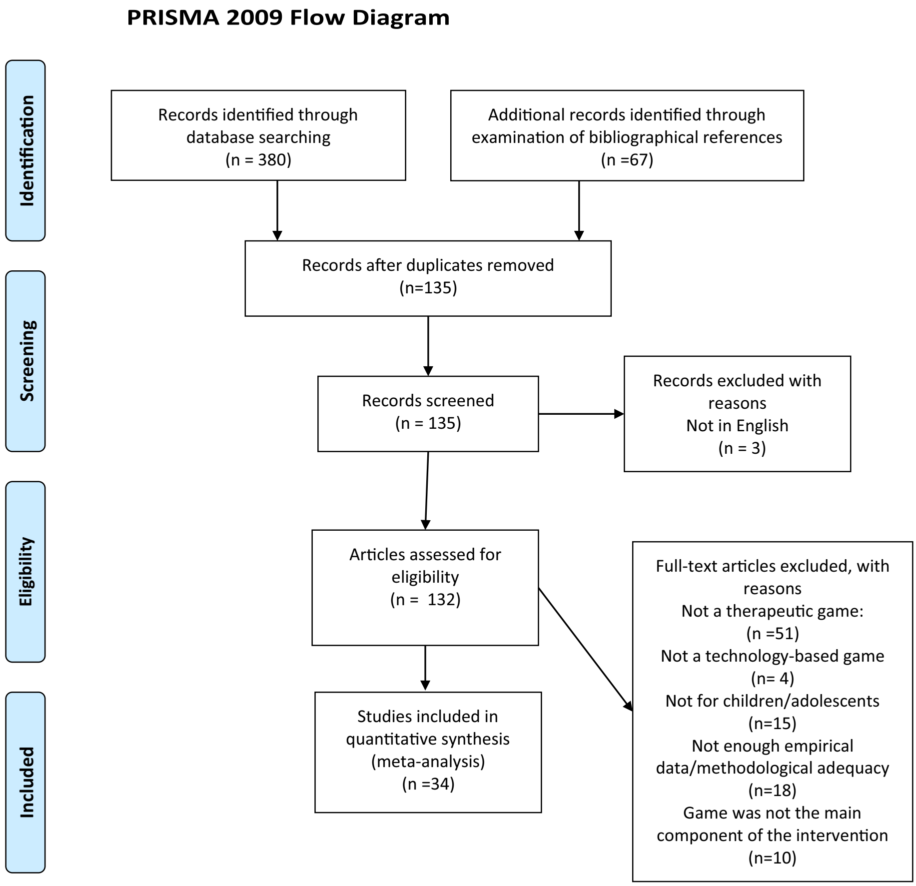

Fig. 1 The PRISMA Flow Diagram of the study

Data Extraction

For each of the included study we extracted the following information: identification data (author, year of publication), comparison condition, outcome, assessment time point (post, follow-up), effect size data, and several moderator variables, including methodological characteristics of the study, intervention (i.e., game) and sample characteristics. We coded the *specific comparison condition* (wait list, regular/unrelated game, treatment as usual/standard care, (psycho) education/group therapy, others). We grouped the *specific outcome measures* based on two independent criteria: (a) type of variable (emotional, cognitive, behavioral, and psychophysiological), and (b) type of measurement (self-report, parent-report, interview-based, and objective: performance-based or psychophysiological indices).

In terms of the *methodological characteristics* of the study, we grouped studies in (a) experimental studies, and (b) clinical randomized studies. The *intervention (game) characteristics* included: (a) game technology (computer, virtual reality system, tablet or smartphone); (b) type of game (educational, therapeutic); and (c) game length (number of sessions, session's length, and total duration of game use—in weeks). The *sample characteristics* included: (a) age and, (b) problem/pathology.

Data were extracted independently by two of the authors (author names removed for blind review). Minor disagreements were solved through discussion.

Quality Assessment

The quality of the included studies was assessed using eight criteria, derived on the basis of authoritative review of empirically supported psychotherapies (Chambless and Hollon 1998) and on the criteria proposed by the Cochrane Collaboration to assess the methodological validity of a study (Higgins et al. 2003; Higgins and Green 2011). A study was considered of high quality if: (1) participants were diagnosed using a standardized clinical interview; (2) a treatment manual has been used; (3) the therapists who treated patients were adequately trained; (4) treatment integrity have been checked; (5) data were analyzed based on intent-to-treat procedure; (6) the study has sufficient statistical power to detect at least a large effect size; (7) adequate randomization have been employed; and (8) assessors were blinded to the participant's condition. The same eight criteria have been used previously by (Cuijpers et al. 2010). Each of the eight criteria was judged as being met or unmet (we considered the criteria being unmet when study failed to match the quality criteria or insufficient information for judging the criteria as being met was provided). All the studies included in the meta-analysis were assessed for quality by two of the authors (author names removed for blind review); minor disagreements were resolved through discussion.

Meta-analytic Procedure

Effect Size Computation

We computed an effect size (ES) for each comparison between the group playing the serious game and the control group. ESs were computed by subtracting the average score of the group playing serious game from the average score of the control group, and dividing the result by the pooled standard deviations of the two groups. A value between 0.2 and 0.5 indicates a small ES, a value of 0.5–0.8 indicates a medium one, and values of 0.8 or larger indicates a large ES (Cohen 1988). Because several studies had small sample sizes, we adjusted the ES for small sample bias according to the procedure developed by Hedges (1981), respectively Hedges and Olkin (1985) (Hedges' *g*). Hedges's *g* is interpreted similarly to Cohen's *d*. Cohen's *d* was computed based on the following formula:

$$d = \frac{m1 - m2}{Swithin}$$

where *m1* is the mean of the experimental group, *m2* is the mean of the control group, and S_{within} is the within-group standard deviation, pooled across groups. Because Cohen's *d* estimates the variance in population based on the variance within samples, in case of small samples it tends to show a small bias, by overestimating the ES. Hedges' *g* removes this bias, by using a correction factor called *J* (Hedges 1981), computed based in the following formula:

$$J = 1 - \frac{3}{4df - 1}$$

where *df* is the degrees of freedom used to estimate S_{within}. Hedges' *g* is computed as the product of *J* and Cohen's *d*, i.e., $g = J \times d$ (Hedges 1981).

We computed separate ESs for different assessment time points (i.e., post intervention, follow-up). To compute ESs, we used means and standard deviations whenever they were available. When this was not the case, we used available statistics, such as *t* and *F* values and sample sizes, Chi Squared and total sample size, *p* values and degrees of freedom. We computed the mean ES for each study, by averaging all the individual ESs computed for each outcome reported in that study for a certain assessment time point. All the ESs were coded such that a positive value of Hedges's *g* indicated greater improvement in the group playing serious game compared with a specific control group. We also reported the obtained results at the effect size thresholds suggested by Ferguson, (Ferguson 2009) for what should be considered the "recommended minimum effect size representing a 'practically' significant effect for social science data" (RMPE; $d = 0.41$; ˇ or $r = .20$). We emphasize effect sizes when presenting and discussing the results.

We used study as the unit of the analysis. However, for certain moderation analysis that implied including different data coming from the same study in different categories (e.g., the case of studies including different types of outcome measures), we used the subgroup within the study as the unit of analysis. As we expected considerable heterogeneity among the studies (given the diversity of serious games and sample characteristics), we computed the mean ES using a random effect model, which assumes that studies come from a pool where the ESs vary (Riley et al. 2011). Homogeneity of ES was assessed with the *Q* statistic and the I^2index (Borenstein et al. 2009; Cooper et al. 2009). The *Q* statistic compares true heterogeneity to random error, with statistically significant values indicating true heterogeneity beyond random error. I^2 reflects the percentage of observed heterogeneity, and, unlike *Q*, does not depend on the number of studies (Borenstein et al. 2009). A value of 0% indicates no heterogeneity, with 25% as low and 50% as moderate, while 75% indicates high heterogeneity (Higgins et al. 2003).

Publication Bias

For each data set based on which we generated a pooled ES, we created and visually inspected funnel plots, which graphically contrasts standard error for each study (determined by sample size) against the study's ES (Light and Pillemer 1984). The assumption of the funnel plot is that studies with larger samples yield larger and more reliable ESs, thereby clustering toward the top of the plot, whereas studies with smaller sample size yield smaller ESs, more prone to error. Therefore, studies with smaller samples should scatter widely about the mean and cluster near the bottom of the funnel plot (Borenstein et al. 2009). Last, we applied Duval and Tweedie's (Duval and Tweedie 2000) trim-and-fill procedure to each data set. This procedure estimates the likely number of missing studies on the basis of asymmetry in the funnel plot, yielding corrected ESs and confidence intervals adjusted to account for these missing studies (Borenstein et al. 2009; Cooper et al. 2009; Duval and Tweedie 2000).

Categorical moderators were tested with a mixed-effect meta-analytic test, which pools the studies within a category using random effects model, whereas tests for significant differences between groups using a fixed effect model. For testing continuous moderators, we used unrestricted maximum likelihood meta-regression analysis, which, when obtaining a significant Z value, indicates a significant relationship between the continuous variable and the ES. All analyses were conducted using Comprehensive Meta-Analysis (Version 2.2.046; Biostat, Englewood, NJ).

Results

Study Characteristics

Characteristics of the studies included in the meta-analysis are summarized in Table 1. Out of the included studies, three were experimental studies (healthy participants), while the rest were interventional studies (randomized clinical trials with clinical interventions); 22 tested educational serious games, while 12 tested therapeutic serious games. Most of the studies tested serious games aimed to improve health-related aspects, while only 7 tested serious games that targeted mental health issues (i.e., ADHD and depression).

Quality Assessment

The quality of the studies varied. No study met all the eight quality criteria, to receive 8 points. A single one (Steiner et al. 2014) met 7 out of the eight criteria, another one met 6 criteria (Kato et al. 2008), and other two (Baranowski et al. 2011; Merry et al. 2012) met 5 quality criteria. Only 6 out of the 31 studies (excluding the experimental studies, for which the quality criteria were inappropriate) reported adequate randomization.

Quantitative Data Synthesis

A summary of the findings is provided in Table 1. Below we present in details the results we obtained.

Overall Effect Size

The overall ES post intervention, computed based on all the studies, indicated a medium and significant effect, Hedges's $g = 0.51$, 95% CI = [0.03; 0.98), $Z = 2.113$, $p = 0.03$, which exceeds the minimum recommended effect size for social sciences (RMPE). The

Table 1 Studies included in the meta-analysis, coding criteria and effect sizes for outcomes overall

Article	Design of the study	Type of measurement	Type of outcome	Problem/pathology	Game name	*Hedges's g* at post-test	*Hedges' g* at follow-up
Baranowski et al. (2003)	Randomized clinical trial	Interview	Behavioral	Healthy	Squire's quest	0.283	–
Bartholomew et al. (2000)	Randomized clinical trial	Parent-report Self-report	Behavioral Psychophysiological Cognitive	Asthma	Watch, Discover, Think and Act	0.315	–
Beale et al. (2007)	Randomized clinical trial	Self-report	Cognitive	Cancer	Re-Mission	0.226	–
Beaumont et al. (2008)	Randomized clinical trial	Objective Parent-report	Cognitive Behavioral	Others	The Junior Detective	0.863	–
Bosworth, et al. (2000)	Randomized clinical trial	Self-report	Behavioral Cognitive	Others	Smart Talk	0.078	–
Brown et al. (1997)	Randomized clinical trial	Objective Parent-report Self-report	Psychophysiological Behavioral Cognitive	Diabetes	Packy and Marlon	0.281	–
Clarke et al. (2012)	Randomized clinical trial	Self-report	Behavioral Cognitive Emotional	Others	PR:EPARe	0.150	–
Cole et al. (2012)	Experimental	Objective	Psychophysiological	Healthy	Re-Mission	0.289	–
Dias et al. (2011)	Experimental	Self-report	Behavioral Cognitive Emotional	Healthy	Advergame	0.475	–
Drechsler et al. (2007)	Randomized clinical trial	Objective Parent-report Teacher-report	Cognitive Behavioral	ADHD	Goefi	0.378	–
Fleming et al. (2012)	Randomized clinical trial	Self-report	Cognitive Emotional Quality of life	Depression	SPARX	0.638	–

Table 1 (continued)

Article	Design of the study	Type of measurement	Type of outcome	Problem/pathology	Game name	*Hedges's g* at post-test	*Hedges' g* at follow-up
Folkvord et al. (2013)	Randomized clinical trial	Objective	Behavioral	Healthy	Advergame	0.477	–
Fuchslocher et al. (2011)	Randomized clinical trial	Self-report	Cognitive Emotional	Diabetes	Balance	0.441	–
Gevensleben et al. (2009)	Randomized clinical trial	Parent-report Teacher-report	Behavioral Emotional	ADHD	Neurofeedback	0.481	–
Huss et al. (2003)	Randomized clinical trial	Self-report	Cognitive	Asthma	Magic School Bus	0.097	–
Kato et al. (2008)	Randomized clinical trial	Objective	Psychophysiological	Cancer	Re-Mission	0.019	0.074
Kumar et al. (2004)	Randomized clinical trial	Objective Parent-report Self-report	Behavioral Psychophysiological Cognitive	Diabetes	Daily	0.508	−0.109
Li et al. (2011)	Randomized clinical trial	Self-report	Emotional	Cancer	Various games	0.609	–
McPherson et al. (2006)	Randomized clinical trial	Parent-report Self-report	Behavioral Cognitive	Asthma	The Asthma Files	0.634	0.248
Merry et al. (2012)	Randomized clinical trial	Self-report	Cognitive Quality of life	Depression	SPARX	0.072	0.043
Verbeken et al. (2013	Randomized clinical trial	Parent-report Teacher-report	Behavioral Cognitive	ADHD	Braingame Brian	0.383	–
Panic et al. (2014)	Randomized clinical trial	Self-report	Cognitive	Others	Various games	0.251	–
Paperny et al. (1989)	Experimental	Self-report	Cognitive	Healthy	The Baby Game and Romance!	0.629	–
Pempek et al. (2009)	Randomized clinical trial	Objective	Cognitive	Others	Pac-Man arcade	3.333	–

Table 1 (continued)

Article	Design of the study	Type of measurement	Type of outcome	Problem/pathology	Game name	*Hedges's g* at post-test	*Hedges' g* at follow-up
Shames et al. (2004)	Randomized clinical trial	Objective Parent-report Self-report	Behavioral Cognitive Psychophysiological Quality of life	Asthma	Bronkie's Asthma Adventure	0.450	0.307
Shegog et al. (2001)	Randomized clinical trial	Self-report	Cognitive	Asthma	It's Your Game	0.788	–
Sherer (1995)	Randomized clinical trial	Self-report	Behavioral Cognitive Emotional	Others	ETGAR	0.519	–
Steiner et al. (2011)	Randomized clinical trial	Parent-report Self-report Teacher-report	Behavioral Behavioral (mixed) Cognitive	ADHD	Neurofeedback	0.232	–
Steiner et al. (2014)	Randomized clinical trial	Parent-report	Behavioral Cognitive	ADHD	Neurofeedback	0.200	0.271
Tanaka et al. (2010)	Randomized clinical trial	Objective	Cognitive	Others	Let's Face It!	– 0.097	–
Todirita et al. (2013)	Randomized clinical trial	Self-report	Cognitive	Others	Amazing chateau	– 0.208	–
Turnin et al. (2001)	Randomized clinical trial	Objective	Behavioral Cognitive	Nutrition/weight management	Various	3.951	–
Verbeken et al. (2013)	Randomized clinical trial	Objective	Cognitive	Nutrition/weight management	Braingame Brian	0.0445	0.225
Yien et al. (2011)	Experimental	Self-report	Cognitive	Nutrition/weight management	Various	0.279	–

heterogeneity was significant and very high, $Q(33) = 2139.39, p = 0.00$, $I^2 = 98.45$. We identified two outliers (Pempek and Calvert 2009; Turnin et al. 2001) with ESs deviating with more than 2 standard deviations from the mean ES. After eliminating them, Hedges's g decreased to 0.24, but was still significant: 95% CI = [0.12; 0.45], $Z = 3.390$, $p = 0.001$. The effect size does not exceed the RMPE and the heterogeneity also decreased slightly, but remained significant, $Q(31) = 203.00$, $p = 0.00$, $I^2 = 84.73$. We did not consider the outliers anymore in any of the subsequent analyses.

As three of the included studies were experimental studies, and the heterogeneity was still high, we computed an overall ES by excluding them. The results showed a similar value of ES, Hedges's $g = 0.24$, 95% CI = [0.06; 0.41], $Z = 2.698$, $p = 0.007$, not exceeding the RMPE, whithsimilar heterogeneity, $Q(30) = 179.52$, $I^2 = 84.40$. Therefore, we chose to consider all the studies, regardless of the design, for the subsequent moderation analysis.

Of the included studies, only 8 reported follow-up measurements. The overall ES computed for follow-up measurements indicated a null ES, Hedges's $g = -0.00$, 95% CI = [−0.49; 0.48], $Z = -0.002$, $p = 0.99$. The heterogeneity was high and significant, $Q(7) = 68.255$, $p = 0.00$, $I^2 = 89.74$. Given these results, we conducted moderation analysis only for the ES computed based on the post intervention measurements.

Moderation Analyses

When considering the type of outcome variable, we found no moderation effect: $Q(3) = 2.361$, $p = 0.50$. However, the computed ESs were significant only for behavioral outcomes ($k = 9$; Hedges's $g = 0.41$, 95% CI = [0.19; 0.62], small but significant practically effect size), and for cognitive outcomes ($k = 16$; Hedges's $g = 0.25$, 95% CI = [0.06; 0.44], not exceeding RMPE), while for emotional ($k = 4$; Hedges's $g = 0.63$, 95% CI = [−0.09; 1.36]) and psychophysiological outcomes ($k = 3$; Hedges's $g = -0.22$, 95% CI = [−1.78; 1.34]) were non-significant. No moderation effect was evident also when considering type of measurement, $Q(3) = 2.11$, $p = 0.54$. Only self-report ($k = 17$; Hedges's $g = 0.41$, 95% CI = [0.22; 0.60] measures showed significant ESs, with small but significant practically effect size, while parent-report ($k = 6$; Hedges's $g = 0.20$, 95% CI = [−0.08; 0.49]) and objective measures ($k = 8$; Hedges's $g = 0.08$, 95% CI = [−0.50; 0.67]) showed non-significant ESs.

In terms of game characteristics, we were unable to test game technology as a moderator, as the variability was quasi-inexistent (most of the games were computer-based and poorly specified). Type of the game did not moderate the ES, Q(1) = 0.361, $p = 0.54$ (educational games: $k = 20$, Hedges's $g = 0.32$, 95% CI = [0.09; 0.55]; therapeutic games: $k = 12$, Hedges's $g = 0.22$, 95% CI = [−0.01; 0.46]). However, game length moderated the ES. When the number of sessions was considered as a continuous moderator, we found that studies employing fewer sessions reported somewhat larger ESs, even if the practical effect size did not exceed the RMPE: slope = −0.01, Z = −2.93, $p = 0.00$, CI = [−0.02; −0.004],. Similarly, session length (in minutes) was found to moderate the ES, with studies using shorter session length reporting somewhat larger ESs: slope = −0.01, Z = -6.27, $p = 0.00$, CI = [−0.01; −0.00], but having no significant practical effect size. Usage of game (in weeks) did not moderated the ES, slope = −0.00, Z = −1.60, $p = 0.10$, CI = [−0.01; 0.00].

When considering sample characteristics as moderators, type of pathology did not moderate the ES, Q(7) = 10.567, $p = 0.159$. However, significant ESs were evidenced only for asthmatic ($k = 5$; Hedges's $g = 0.38$, 95% CI = [0.13; 0.64], not exceeding the RMPE) and healthy participants ($k = 5$; Hedges's $g = 0.57$, 95% CI = [0.32; 0.82], with small significant practical effect size), while in case of Attention Deficit and Hyperactivity Disorder (ADHD; $k=5$; Hedges's $g=0.18$, 95% CI = [−0.17; 0.54]), oncological conditions ($k=3$; Hedges's $g=0.35$, 95% CI = [−0.15; 0.85]), depression ($k=6$; Hedges's $g=0.68$, 95% CI = [−0.95; 2.32]), diabetes ($k=3$; Hedges's $g=0.26$, 95% CI = [−0.07; 0.53]), nutrition/weight management ($k=2$; Hedges's $g=-0.76$, 95% CI = [-2.39; 0.85]) or other problems ($k=7$; Hedges's $g=0.18$, 95% CI = [−0.08; 0.45]), the computed ESs were non-significant. Participants' age marginally moderated the ES, with younger participants benefiting the most from the serious games: slope = -0.01, Z = −1.82, $p=0.06$, CI = [−0.04; 0.00], not exceeding the RMPE. The quality of the studies also moderated the ES, with lower quality studies yielding higher ESs, but no significant practically effect size (not exceeding the RMPE): slope = −0.09, Z = −3.49, $p=0.00$, CI = [−0.14; −0.04].

Publication Bias

Duval and Tweedie's (2000) trim and fill procedure, estimated 3 missing studies to the left of the mean (based on a random effect model), that would have reduced the mean ES to 0.20, 95% CI = [0.14; 0.26] (see also the funnel plot, Fig. 2): as shown, the estimated mean ES remain very similar to the original one, and still significant.

Discussion

This is the first meta-analysis to assess the clinical utility of serious games in children and adolescents. We performed a systematic review and meta-analysis on 34 controlled studies examining the efficacy of serious games in reducing symptoms of mental disorders and promoting health-related behavioral change. We evaluated the quality of the studies, and computed effect sizes both for post intervention and for follow-up measurements. We then explored possible moderators of the effect size and evaluated potential publication bias.

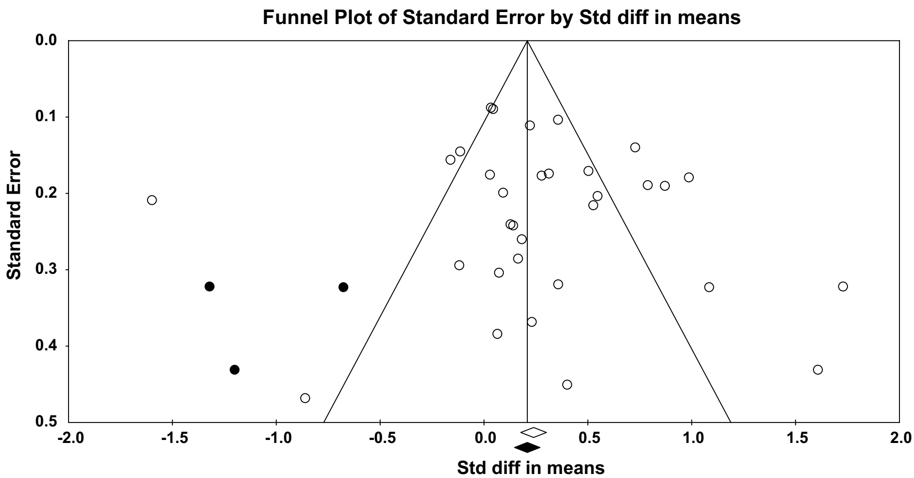

Fig. 2 The funnel plot of the publication bias

Main Effects

Overall, at post intervention, results showed a small main statistical and practical effect size favoring the group receiving serious game intervention compared with the control group. This effect remained significant after the exclusion of outlier studies and adjustment for publication bias. However, the “practically” significant effect was no longer significant after the exclusion of outliers, and the associated heterogeneity was very high, even after the exclusion of outlier studies. The reduction of the practically effect size after removing the outliers, suggests that the overall effects of the game are probably lower still, being artificially inflated due to the weak methodologies of the studies.

Although the high heterogeneity obtained is not unexpected (similar reviews on the topic of serious gaming reported high heterogeneity, too; e.g., DeSmet et al., 2014; Granic et al. 2014) it certainly limits the interpretation of the results we obtained. In our analysis, the high heterogeneity is most probably the result of the combination of two main source of variations: types of game used, and types of problem addressed. Most of the serious games tested in the studies we included in the meta-analysis were used with an educational purpose on youth population suffering from various medical conditions. However, different studies tested different poor specified games that apparently varied in purpose, content, complexity, time and context of usage, and so on. The outcomes measures varied also widely. Studies testing the clinical utility of serious games for mental health problems are rather scarce, but similarly wide varied. Future research should aim to explore the benefits of serious gaming in a more compact manner: well-defined serious games, firmly rooted in standard evidence-based intervention strategies, and which address specific (mental) health problems should be tested in multiple, comparable randomized trials, to facilitate the integration of the obtained results.

For follow-up measurements, the overall effect size was non-significant. However, only 8 studies reported follow-up measurements, and the heterogeneity within them was high and significant. Future studies should explicitly aim to investigate long-term benefits of serious gaming, by employing follow-up assessments.

Moderator Effects

Moderation analysis, run exclusively for the post intervention data set, showed the effect size was moderated by participant’s age, number of sessions, duration of session, and study quality, with younger participants, fewer and shorter sessions, and lower quality of the study being associated with higher statistical effect size, without observing significant effect sizes exceeding the recommended minimum effect size representing the practically significant effect for social science data. The finding regarding the participants’ age seems to suggest that serious games work better for the youngest. However, a more probable possibility is that certain games are more appropriate for certain developmental stages. One should note that most of the games tested in the studies we included in this meta-analysis was rather less complex games, making them probably less interesting and/or engaging for older participants. The clinical gain may be dependent on the challenge raised by the game for a certain user. We believe that the biggest potential of the serious games lies specifically in their particularity of presenting soundly scientific content in a playful manner that is highly individualized. Therefore, future serious games used for (mental) health purposes should capitalize on this feature of individualization. Ideally, the game should be highly responsive and adaptable to the user's response, in such a manner that promotes continuous engagement and enjoyment. In this way, serious games could be successfully used for clinical benefits with various age categories. We believe that the actual technological development is an asset for developing such serious games. Our results seem to suggest that older youth need more challenging and complex games, but this remains to be tested in future research, using more rigorous studies.

Type of outcome variable did not moderate the results at post intervention, although ESs were significant only for behavioral and cognitive outcome measures. From the practical point of view, the effect size was significant only for behavioral outcomes. However, a main limitation of this analysis was the limited number of outcome measures included in the other two categories (i.e., emotional, and psychophysiological outcomes). Therefore, these results should be cautiously interpreted, keeping in mind the limited statistical power of the analysis. Future studies should consider the effects of serious gaming specifically in terms of emotional and psychophysiological benefits, to allow extracting more pertinent conclusions.

Similarly, the type of measurement (in terms of self-report, interview-based, parent-report or objective measures) did not moderate the ES. However, we obtained significant ESs only for self-report and interview-based measures, with significant practical effect size for self-report measures. Though, the majority of the outcome measures were self-report, while only one study included a single interview-based measure. Therefore, our results are likely to suffer from lack of statistical power. What we can remark for now is the excessive reliance on self-report measures when investigating the benefits of serious gaming. As self-report measures are more susceptible to demands effect and artefacts (Williamson 2007), future studies should include additional types of measures, especially when considering the clinical gain of serious gaming. Obviously, self-report measures are invaluable when information on user experience, engagement and entertainment is needed, but one should keep in mind that the goal of serious games transcends this. When looking for evidence regarding the benefits of serious gaming, researchers should strive to minimize possible confounders, like demand effects.

Type of the game did not moderate the ES, with educational and therapeutic games yielding similar small, but significant effects. This finding supports the possibility that serious gaming might be successfully used both with preventative and therapeutic purposes. However, future research should strive to maximize the benefits of serious gaming, by modifying the game specifications and/or gaming regimen to promote the enlargement of the ES.

Risk of Bias

Publication bias favoring positive results was small. However, the quality of the studies was low and wedged the effect sizes, with lower quality studies yielding the larger ESs. We used relatively conservative criteria in judging the quality of the studies, and some of them may did not match perfectly the specific of the studies we considered: for example, in judging the existence of the "treatment manual" (criterion 2), we took into consideration if there were clear game specifications in terms of game contents and recommended use instead of a "treatment manual". Despite this, the finding that the lower the quality, the higher the ES, is worrying: we clearly need high quality studies adequately documenting the potential benefits of serious games in children and adolescents before using serious games with a clinical purpose – either preventative or therapeutic.

As serious games represent relatively new potential interventions, no dedicated quality criteria have been developed for the assessment of studies investigating their benefits. Future research should consider the development of such criteria.

P-Curve

We conducted a p-curve analysis to investigate the p value distributions in comparison to the expected distributions and in this way to determine the presence of evidential values (Simonsohn et al. 2014). This p-curve analysis converts the p values for each study into z scores and in a second step it divides the sum of the z scores by the square root of the number of p values, and in the third step the overall z score that is compared to a null of 33% power. Results obtained from the continuous test using the Stouffer method suggested evidential value for the overall studies, $z = -3.39$, $p = .0003$ (full p-curve with $ps < .05$) and $z = -4.82$, $p < .0001$ (half p-curve with $ps < .025$). p-curve analysis estimated that the mean power was 48%. CI 90% (19%, 75%). As Fig. 3 shows, results are right-skewed, thus suggesting the existence of evidential value.

Limitations

Our meta-analysis has several limitations. First, the high heterogeneity we found within our sample of studies make difficult the meaningful interpretation of the available data. While openly admitting this serious limit in the interpretation of our results, we also want to point out it as a general limitation of the field, and emphasize the need for high quality, comparable studies investigating the potential benefits of serious gaming in children and adolescents. Second, some subgroup analyses involved a few studies, limiting the interpretation of moderating analysis. Third, most of the studies we included tested poorly specified games, which prevented some potential important moderation analysis (like game technology, game content, level of interactivity, etc.). Fourth, most of the outcome measures

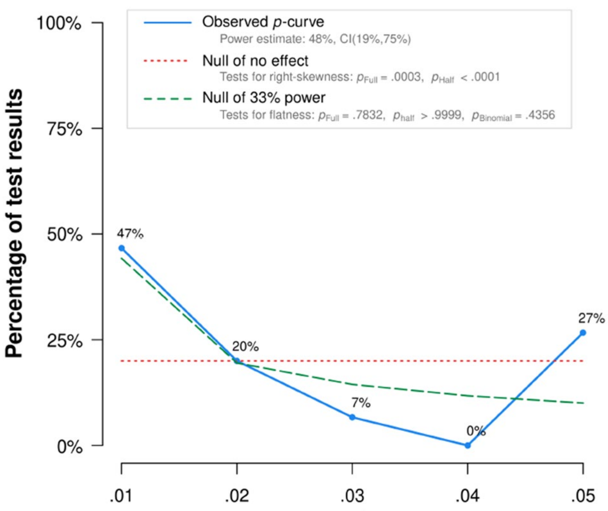

Fig. 3 Distribution of observed p values compared to the expected distribution of p values. The observed p-curve includes 15 statistically significant ($p < .05$) results, of which 10 are $p > .025$. There were additional results entered but excluded from p-curve because they ere $p > .05$

reported in the majority of the studies were self-report. Finally, the quality of studies was low and wedged the effect size.

Pragmatic Implications and Conclusions

Results of this meta-analysis suggest that, most probably, serious games are not yet ready for dissemination as a stand-alone treatment/prevention strategy or as an adjunct to treatment-as-usual interventions aimed to reduce symptoms or promote (mental) health/health-related behavioral change in children and adolescents. The ES is small, not exceeding the recommended practical level for social science data, and the quality of the existing studies is suboptimal. However, we believe these results provide reasonable steps in developing further the serious games' potential of serving as valuable clinical tools. First, additional research investigating the clinical utility of serious gaming for children and adolescents need to be programmatically conducted. We need well designed, pre-registered (Kvarven et al. 2020) well controlled, high quality studies, adequately powered to detect potential benefits of serious games, both on short and long term. Second, games tested should be clearly specified, in terms of goals, contents, and tasks. With well specified games and high quality studies, we will be in the position of maximize the clinical potential of computerized gaming. Third, research testing the clinical potential of serious games in children should go beyond the self-report measures, and explore the clinical benefits while considering also the potential drawbacks. In this respect, future research should establish the optimal „gaming dose". A practical way of avoid excessive reliance on self-report measures while trying to determine the optimal „gaming dose" could be using objective information about game usage, collected into the game. For example, recording number of game initialization per day/week (i.e., the intensity of game usage), the duration of the game session, level achieved in the game, the behavior of the game user during the game (in terms of actions, points gained/lost, etc.), and so on could prove to be a more practical and useful way of quantifying the „gaming dose" than simply prescribing a certain number of gaming sessions.

To conclude, although the field of serious games in children and adolescent seemed to have generated mainly random, poor quality research work, there is serious room for improvement. We look forward to seeing further research in this field turning promises into deeds.

Acknowledgements This work was supported by a grant awarded to David Oana Alexandra from the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project number PN-II-PT-PCCA-2013-4-1937.

Authors' Contribution We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. Contributions: The 1st author contributed to the design of the study, coding system, literature search, interpretation of data, and manuscript writing; the 2nd author brought contributions to the literature search, coding, and manuscript writing; the 3rd author brought contributions to the literature search, data coding, and manuscript writing; the 4th author brought contributions to the literature search, data coding, data analysis, interpretation of data, and manuscript writing.

Funding This work was supported by a grant awarded to Oana A. David from the Romanian National Authority for Scientific Research, CNCS – UEFISCDI [Grant Number PN-II-PT-PCCA-2013-4-1937].

Data Availability Data is not currently stored in a public repository.

Compliance with Ethical Standards

Conflicts of interest We wish to confirm that there are no known conflicts of interest associated with this publication.

Ethics Approval and Consent This is a review study. The BBU Research Ethics Committee has confirmed that no ethical approval/consent is required.

References

Arns, M., de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity—a meta-analysis. *Clinical EEG and Neuroscience, 40*(3), 180–189. https://doi.org/10.1177/155005940904000311.

Baranowski, T., Baranowski, J., Cullen, K. W., Marsh, T., Islam, N., Zakeri, I., ... & Demoor, C. (2003). Squire's Quest!: Dietary outcome evaluation of a multimedia game. *American Journal of Preventive Medicine, 24*(1), 52–61.

Baranowski, T., Baranowski, J., Thompson, D., Buday, R., Jago, R., Griffith, M. J., et al. (2011). Video game play, child diet, and physical activity behavior change: A randomized clinical trial. *American Journal of Preventive Medicine, 40*(1), 33–38. https://doi.org/10.1016/j.amepre.2010.09.029.

Baranowski, T., Blumberg, F., Buday, R., DeSmet, A., Fiellin, L. E., Green, C. S., et al. (2016). Games for health for children—current status and needed research. *Games For Health Journal, 5*(1), 1–12. https://doi.org/10.1089/g4h.2015.0026.

Baranowski, T., Buday, R., Thompson, D. I., & Baranowski, J. (2008). Playing for real: Video games and stories for health-related behavior change. *American Journal of Preventive Medicine, 34*(1), 74–82. https://doi.org/10.1016/j.amepre.2007.09.027.

Bartholomew, L. K., Gold, R. S., Parcel, G. S., Czyzewski, D. I., Sockrider, M. M., Fernandez, M., ... & Swank, P. (2000). Watch, Discover, Think, and Act: Evaluation of computer-assisted instruction to improve asthma self-management in inner-city children. *Patient Education and Counseling, 39*(2-3), 269–280.

Beale, I. L., Kato, P. M., Marin-Bowling, V. M., Guthrie, N., & Cole, S. W. (2007). Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. *Journal of Adolescent Health, 41*(3), 263–270.

Beaumont, R, Sofronoff, K. (2008). A multi-component social skills intervention for children with Asperger syndrome: The junior detective training program. *Journal of Child Psychology and Psychiatry, 49*, 743–753.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Fixed-effect versus random-effects models. In *Introduction to meta-analysis* (pp. 77–86). London: Wiley. https://doi.org/10.1002/9780470743386.ch13.

Bosworth, K., Espelage, D., DuBay, T., Daytner, G., & Karageorge, K. (2000). Preliminary evaluation of a multimedia violence prevention program for adolescents. *American Journal of Health Behavior, 24*(4), 268–280.

Brown, S. J., Lieberman, D. A., Gemeny, B. A., Fan, Y. C., Wilson, D. M., & Pasta, D. J. (1997). Educational video game for juvenile diabetes: Results of a controlled trial. *Medical Informatics, 22*(1), 77–89.

Cartreine, J. A., Ahern, D. K., & Locke, S. E. (2010). A roadmap to computer-based psychotherapy in the United States. *Harvard Review of Psychiatry, 18*(2), 80–95. https://doi.org/10.3109/10673221003707702.

Ceranoglu, T. A. (2010). Star Wars in psychotherapy: Video games in the office. *Academic Psychiatry : The Journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry, 34*(3), 233–236. https://doi.org/10.1176/appi.ap.34.3.233.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7–18.

Clarke, S., Arnab, S., Dunwell, I., & Brown, K. (2012). PR: EPARe: A game-based approach to relationship guidance for adolescents. *Procedia Computer Science, 15*, 38–44.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (p. 2). Hilsdale, NJ: Lawrence Earlbaum Associates.

Cole, S. W., Yoo, D. J., & Knutson, B. (2012). Interactivity and reward-related neural activation during a serious videogame. *PLoS One, 7*(3), e33909.
Collishaw, S., Maughan, B., Natarajan, L., & Pickles, A. (2010). Trends in adolescent emotional problems in England: A comparison of two national cohorts twenty years apart. *Journal of Child Psychology and Psychiatry, 51*(8), 885–894. https://doi.org/10.1111/j.1469-7610.2010.02252.x.
Compton, S. N., March, J. S., Brent, D., Albano, A. M., Weersing, V. R., & Curry, J. (2004). Cognitive-behavioral psychotherapy for anxiety and depressive disorders in children and adolescents: An evidence-based medicine review. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(8), 930–959. https://doi.org/10.1097/01.chi.0000127589.57468.bf.
Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation.
Coyle, D., Doherty, G., & Sharry, J. (2009). An evaluation of a solution focused computer game in adolescent interventions. *Clinical Child Psychology and Psychiatry, 14*(3), 345–360. https://doi.org/10.1177/1359104508100884.
Crenshaw, D. A. (2008). *Therapeutic engagement of children and adolescents: Play, symbol, drawing, and storytelling strategies*. Incorporated: Jason Aronson.
Cuijpers, P., Smit, F., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). Efficacy of cognitive–behavioural therapy and other psychological treatments for adult depression: Meta-analytic study of publication bias. *The British Journal of Psychiatry, 196*(3), 173–178. https://doi.org/10.1192/bjp.bp.109.066001.
DeShazo, J., Harris, L., & Pratt, W. (2010). Effective intervention or child's play? A review of video games for diabetes education. *Diabetes Technology & Therapeutics, 12*(10), 815–822. https://doi.org/10.1089/dia.2010.0030.
DeSmet, A., Shegog, R., Van Ryckeghem, D., Crombez, G., & De Bourdeaudhuij, I. (2014a). A systematic review and meta-analysis of interventions for sexual health promotion involving serious digital games. *Games for Health Journal, 4*(2), 78–90. https://doi.org/10.1089/g4h.2014.0110.
DeSmet, A., Van Ryckeghem, D., Compernolle, S., Baranowski, T., Thompson, D., Crombez, G., et al. (2014b). A meta-analysis of serious digital games for healthy lifestyle promotion. *Preventive Medicine, 69*(Supplement C), 95–107. https://doi.org/10.1016/j.ypmed.2014.08.026.
Dias, M., & Agante, L. (2011). Can advergames boost children's healthier eating habits? A comparison between healthy and non-healthy food. *Journal of Consumer Behaviour, 10*(3), 152–160.
Drechsler, R., Straub, M., Doehnert, M., Heinrich, H., Steinhausen, H. C., & Brandeis, D. (2007). Controlled evaluation of a neurofeedback training of slow cortical potentials in children with attention deficit/hyperactivity disorder (ADHD). *Behavioral and Brain Functions, 3*(1), 35.
Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.
Farrell, L. J., & Barrett, P. M. (2007). Prevention of childhood emotional disorders: Reducing the burden of suffering associated with anxiety and depression. *Child and Adolescent Mental Health, 12*(2), 58–65. https://doi.org/10.1111/j.1475-3588.2006.00430.x.
Ferguson, C. J. (2009). Is psychological research really as good as medical research? Effect size comparisons between psychology and medicine. *Review of General Psychology, 13*(2), 130–136.
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. *Behavioural and Cognitive Psychotherapy, 40*(5), 529–541.
Folkvord, F., Anschütz, D. J., Buijzen, M., & Valkenburg, P. M. (2013). The effect of playing advergames that promote energy-dense snacks or fruit on actual food intake among children. *The American Journal of Clinical Nutrition, 97*(2), 239–245.
Fuchslocher, A., Niesenhaus, J., & Krämer, N. (2011). Serious games for health: An empirical study of the game "Balance" for teenagers with diabetes mellitus. *Entertainment Computing, 2*(2), 97–101.
Gamberini, L., Barresi, G., Maier, A., & Scarpetta, F. (2008). A game a day keeps the doctor away: A short review of computer games in mental healthcare. *Journal of CyberTherapy and Rehabilitation, 1*(2), 127–145.
Gentile, D. A., & Walsh, D. A. (2002). A normative study of family media habits. *Journal of Applied Developmental Psychology, 23*(2), 157–178. https://doi.org/10.1016/S0193-3973(02)00102-8.
Gevensleben, H., Holl, B., Albrecht, B., Vogel, C., Schlamp, D., Kratz, O., ... & Heinrich, H. (2009). Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. *Journal of Child Psychology and Psychiatry, 50*(7), 780–789.

Goh, D. H., Ang, R. P., & Tan, H. C. (2008). Strategies for designing effective psychotherapeutic gaming interventions for children and adolescents. *Computers in Human Behavior, 24*(5), 2217–2235. https://doi.org/10.1016/j.chb.2007.10.007.

Goodman, A., Joyce, R., & Smith, J. P. (2011). The long shadow cast by childhood physical and mental problems on adult life. *Proceedings of the National Academy of Sciences, 108*(15), 6032–6037. https://doi.org/10.1073/pnas.1016970108.

Gould, M. S., Munfakh, J. L. H., Lubell, K., Kleinman, M., & Parker, S. (2002). Seeking help from the internet during adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(10), 1182–1189. https://doi.org/10.1097/00004583-200210000-00007.

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *The American Psychologist, 69*(1), 66–78. https://doi.org/10.1037/a0034857.

Guse, K., Levine, D., Martins, S., Lira, A., Gaarde, J., Westmorland, W., et al. (2012). Interventions using new digital media to improve adolescent sexual health: A systematic review. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 51*(6), 535–543. https://doi.org/10.1016/j.jadohealth.2012.03.014.

Guy, S., Ratzki-Leewing, A., & Gwadry-Sridhar, F. (2011). Moving beyond the stigma: Systematic review of video games and their potential to combat obesity. *International Journal of Hypertension, 2011,* 179124. https://doi.org/10.4061/2011/179124.

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128. https://doi.org/10.3102/10769986006002107.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. London: Academic Press.

Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions*. London: Wiley.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal, 327*(7414), 557–560.

Horgan, Á., & Sweeney, J. (2010). Young students' use of the Internet for mental health information and support. *Journal of Psychiatric and Mental Health Nursing, 17*(2), 117–123. https://doi.org/10.1111/j.1365-2850.2009.01497.x.

Horne-Moyer, H. L., Moyer, B. H., Messer, D. C., & Messer, E. S. (2014). The use of electronic games in therapy: A review with clinical implications. *Current Psychiatry Reports, 16*(12), 520. https://doi.org/10.1007/s11920-014-0520-6.

Huss, K., Winkelstein, M., Nanda, J., Naumann, P. L., Sloand, E. D., & Huss, R. W. (2003). Computer game for inner-city children does not improve asthma outcomes. *Journal of Pediatric Health Care, 17*(2), 72–78.

Kato, P. M., Cole, S. W., Bradlyn, A. S., & Pollock, B. H. (2008). A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. *Pediatrics, 122*(2), e305–e317. https://doi.org/10.1542/peds.2007-3134.

Kharrazi, H., Lu, A. S., Gharghabi, F., & Coleman, W. (2012). A scoping review of health game research: Past, present, and future. *Games for Health Journal, 1*(2), 153–164. https://doi.org/10.1089/g4h.2012.0011.

Knopf, D., Park, M. J., & Mulye, T. P. (2008). The mental health of adolescents: A national profile, 2008. *San Francisco, CA: National Adolescent Health Information Center*. https://pdfs.semanticscholar.org/8619/3766e17c9e5e834dca8120d5896cefe20be4.pdf

Kumar, V. S., Wentzell, K. J., Mikkelsen, T., Pentland, A., & Laffel, L. M. (2004). The DAILY (Daily Automated Intensive Log for Youth) trial: A wireless, portable system to improve adherence and glycemic control in youth with diabetes. *Diabetes Technology & Therapeutics, 6*(4), 445–453.

Kvarven, A., Strømland, E., Wollbrant, C., Andersson, D., Johannesson, M., Tinghög, G., ... & Myrseth, K. O. R. (2020). The intuitive cooperation hypothesis revisited: A meta-analytic examination of effect size and between-study heterogeneity. *Journal of the Economic Science Association*, 1–16.

Li, W. H., Chung, J. O., & Ho, E. K. (2011). The effectiveness of therapeutic play, using virtual reality computer games, in promoting the psychological well-being of children hospitalised with cancer. *Journal of Clinical Nursing, 20*(15–16), 2135–2143.

Light, R. J., & Pillemer D. B. (1984). Summing Up—The Science of Reviewing Research. Cambridge, MA,USA: Harvard University Press. http://www.hup.harvard.edu/catalog.php?isbn=9780674854314&content=reviews

Lu, A. S., Baranowski, T., Thompson, D., & Buday, R. (2012). Story immersion of videogames for youth health promotion: A review of literature. *Games for Health Journal, 1*(3), 199–204. https://doi.org/10.1089/g4h.2011.0012.

Mcfarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games: An exploration by TEEM of the contribution which games can make to the education process*. http://www.teem.org.uk/publications/teem_gamesined_full.pdf
McPherson, A. C., Glazebrook, C., Forster, D., James, C., & Smyth, A. (2006). A randomized, controlled trial of an interactive educational computer package for children with asthma. *Pediatrics, 117*(4), 1046–1054.
Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., et al. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey replication–adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017.
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. https://doi.org/10.1136/bmj.e2598.
Newman, M. G., Szkodny, L. E., Llera, S. J., & Przeworski, A. (2011). A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: Is human contact necessary for therapeutic efficacy? *Clinical Psychology Review, 31*(1), 89–103. https://doi.org/10.1016/j.cpr.2010.09.008.
NPD Group. (2011). The video game industry is adding 2-17 year-old gamers at a rate higher than that age group's population growth [press release]. https://www.npd.com/wps/portal/npd/us/news/press-releases/pr_111011/—References—Scientific Research Publish. (n.d.). Retrieved September 27, 2017, from http://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=1056382
Panic, K., Cauberghe, V., & De Pelsmacker, P. (2014). Promoting dental hygiene to children: Comparing traditional and interactive media following threat appeals. *Journal of Health Communication, 19*(5), 561–576.
Pao, M., & Bosk, A. (2011). Anxiety in medically ill children/adolescents. *Depression and Anxiety, 28*(1), 40–49. https://doi.org/10.1002/da.20727.
Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. *Computers & Education, 53*(3), 603–622. https://doi.org/10.1016/j.compedu.2009.04.001.
Paperny, D. M., & Starn, J. R. (1989). Adolescent pregnancy prevention by health education computer games: Computer-assisted instruction of knowledge and attitudes. *Pediatrics, 83*(5), 742–752.
Pempek, T. A., & Calvert, S. L. (2009). Tipping the balance: Use of advergames to promote consumption of nutritious foods and beverages by low-income African American children. *Archives of Pediatrics and Adolescent Medicine, 163*(7), 633–637. https://doi.org/10.1001/archpediatrics.2009.71.
Primack, B. A., Carroll, M. V., McNamara, M., Klem, M. L., King, B., Rich, M., et al. (2012). Role of video games in improving health-related outcomes: A systematic review. *American Journal of Preventive Medicine, 42*(6), 630–638. https://doi.org/10.1016/j.amepre.2012.02.023.
Przeworski, A., & Newman, M. G. (2006). Efficacy and utility of computer-assisted cognitive behavioural therapy for anxiety disorders. *Clinical Psychologist, 10*(2), 43–53. https://doi.org/10.1080/13284200500378779.
Rahmani, E., & Boren, S. A. (2012). Videogames and health improvement: A literature review of randomized controlled trials. *Games for Health Journal, 1*(5), 331–341. https://doi.org/10.1089/g4h.2012.0031.
Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ, 342,* d549. https://doi.org/10.1136/bmj.d549.
Shames, R. S., Sharek, P., Mayer, M., Robinson, T. N., Hoyte, E. G., Gonzalez-Hensley, F., ... & Umetsu, D. T. (2004). Effectiveness of a multicomponent self-management program in at-risk, school-aged children with asthma. *Annals of Allergy, Asthma & Immunology, 92*(6), 611–618.
Shegog, R., Bartholomew, L. K., Parcel, G. S., Sockrider, M. M., Mâsse, L., & Abramson, S. L. (2001). Impact of a computer-assisted education program on factors related to asthma selfmanagement behavior. *Journal of the American Medical Informatics Association, 8*(1), 49–61.
Sherer, M. (1995). The effect of computerized simulation games on the moral development of youth in distress. *Computers in Human Services, 11*(1-2), 81–95.
Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143,* 534–547. https://doi.org/10.1037/a0033242.
Steiner, N. J., Frenette, E. C., Rene, K. M., Brennan, R. T., & Perrin, E. C. (2014). In-school neurofeedback training for ADHD: Sustained improvements from a randomized control trial. *Pediatrics, 133*(3), 483–492. https://doi.org/10.1542/peds.2013-2059.

Steiner, N. J., Sheldrick, R. C., Gotthelf, D., & Perrin, E. C. (2011). Computer-based attention training in the schools for children with attention deficit/hyperactivity disorder: A preliminary trial. *Clinical Pediatrics, 50*(7), 615–622.

Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., ... & Schultz, R. T. (2010). Using computerized games to teach face recognition skills to children with autism spectrum disorder: The Let's Face It! program. *Journal of Child Psychology and Psychiatry, 51*(8), 944–952.

Todirita, I. R., & Lupu, V. (2013). Gambling prevention program among children. *Journal of Gambling Studies, 29*(1), 161–169.

Turnin, M. C., Tauber, M. T., Couvaras, O., Jouret, B., Bolzonella, C., Bourgeois, O., et al. (2001). Evaluation of microcomputer nutritional teaching games in 1876 children at school. *Diabetes & Metabolism, 27*(4 Pt 1), 459–464.

Verbeken, S., Braet, C., Goossens, L., & Van der Oord, S. (2013). Executive function training with game elements for obese children: A novel treatment to enhance self-regulatory abilities for weight-control. *Behaviour Research and Therapy, 51*(6), 290–299.

Wilkinson, N., Ang, R. P., & Goh, D. H. (2008). Online video game therapy for mental health concerns: A Review. *International Journal of Social Psychiatry, 54*(4), 370–382. https://doi.org/10.1177/0020764008091659.

Williamson, A. (2007). Using self-report measures in neurobehavioural toxicology: Can they be trusted? *NeuroToxicology, 28*(2), 227–234. https://doi.org/10.1016/j.neuro.2006.03.009.

Yien, J. M., Hung, C. M., Hwang, G. J., & Lin, Y. C. (2011). A game-based learning approach to improving students' learning achievements in a Nutrition course. T*urkish Online Journal of Educational Technology-TOJET, 10*(2), 1–10.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.