🅐🅐🅐 **University of Reading**

Immersive virtual reality and digital applied gaming interventions for the treatment of mental health problems in children and young people: the need for rigorous treatment development and clinical evaluation

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the

work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/jcpp.13400

Publisher: Wiley

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Annual Research Review: Immersive virtual reality and digital applied gaming interventions for the treatment of mental health problems in children and young people: the need for rigorous treatment development and clinical evaluation

Brynjar Halldorsson,[1,2,3] **Claire Hill,**[4] **Polly Waite,**[1,2,4] **Kate Partridge,**[5] **Daniel Freeman,**[2,6] **and Cathy Creswell**[1,2]

[1]Department of Experimental Psychology, University of Oxford, Oxford, UK; [2]Department of Psychiatry, University of Oxford, Oxford, UK; [3]Department of Psychology, Reykjavik University, Reykjavik, Iceland; [4]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK; [5]CAMHS Anxiety and Depression Pathway, Berkshire Healthcare Foundation Trust, University of Reading, Reading, UK; [6]Oxford Health NHS Foundation Trust, Oxford, UK

Background: Mental health problems in children and young people are common and can lead to poor long-term outcomes. Despite the availability of effective psychological interventions for mental health disorders, only a minority of affected children and young people access treatment. Digital interventions, such as applied games and virtual reality (VR), that target mental health problems in children and young people may hold a key to increasing access to, engagement with, and potentially the effectiveness of psychological treatments. To date, several applied games and VR interventions have been specifically developed for children and young people. This systematic review aims to identify and synthesize current data on the experience and effectiveness of applied games and VR for targeting mental health problems in children and young people (defined as average age of 18 years or below). **Methods:** Electronic systematic searches were conducted in Medline, PsycINFO, CINAHL, and Web of Science. **Results:** Nineteen studies were identified that examined nine applied games and two VR applications, and targeted symptoms of anxiety, depression, and phobias using both quantitative and qualitative methodologies. Existing evidence is at a very early stage and studies vary extensively in key methodological characteristics. For applied games, the most robust evidence is for adolescent depressive symptoms (medium clinical effect sizes). Insufficient research attention has been given to the efficacy of VR interventions in children and young people. **Conclusions:** The evidence to date is at a very early stage. Despite the enthusiasm for applied games and VR, existing interventions are limited in number and evidence of efficacy, and there is a clear need for further co-design, development, and evaluation of applied games and VR before they are routinely offered as treatments for children and young people with mental health problems. **Keywords:** Mental health; qualitative methods; RCT design; treatment trials.

Introduction

Mental health problems in children and young people are common, affecting approximately one in eight 5- to 19-year-olds (Vizard, Pearce, & Davis, 2018). They typically have a substantial negative impact on development and school, social, and health functioning (Green, McGinnity, Meltzer, Ford, & Goodman, 2005; Pompili et al., 2010), present a risk for ongoing mental health problems (Copeland, Angold, Shanahan, & Costello, 2014) and bring significant social costs (Fineberg et al., 2013). Despite the availability of effective psychological interventions for mental health disorders, only a minority of affected children and young people access support or treatment (Reardon, Harvey, & Creswell, 2019), with studies finding as few as 2%

receiving specialist, evidence-based interventions for some disorders (Lawrence et al., 2016; Reardon, Harvey, et al., 2019). Digital mental health interventions have been used to increase access to evidence-based treatments for mental health problems in children and young people (Hollis et al., 2017) and adults (Andersson, Cuijpers, Carlbring, Riper, & Hedman, 2014), either as fully automated intervention programs or in combination with other traditional therapies. With recent advances in computerized technologies, the range and scope of digital health interventions have evolved and changed dramatically over the last decade (Hollis et al., 2017). Emerging interventions include applied games (also known as serious games) and virtual reality (VR).

Applied games are 'digital interventions that employ games or substantial game elements in an effort to educate and/or change patterns of

experience and/or behavior' (Fleming et al., 2017). When it comes to targeting mental health problems, evidence-based interventions can be translated into computer gaming formats and use features of computer games (e.g., challenges and levels) to target symptoms (Fleming et al., 2017).

On the other hand, VR is the use of computer modeling and simulation that enables a person to interact with an artificial three-dimensional visual or other sensory environment. With VR, people can enter simulations (typically delivered via a headset) of the situations that trouble them, and so, in the case of anxiety difficulties for example, this can give them an opportunity to re-evaluate their fears, test out therapeutic strategies, and acquire new learning which transfers to the real world.

Applied games and VR have been subjected to more extensive evaluation as treatments for mental health problems within the adult literature (e.g., Freeman et al., 2017) compared to the literature about children and young people. However, promising evidence is emerging of clinical gains using applied games or VR for mental health problems in children and young people (Maskey, Rodgers, et al., 2019; Merry et al., 2012). Given that many children and young people have grown up surrounded by and using digital devices that often play an integral part in their lives (Lenhart, Purcell, Smith, & Zickuhr, 2010), modern digital interventions may have particular appeal and utility amongst this population (Bakker, Kazantzis, Rickwood, & Rickard, 2016). Furthermore, as children and young people could potentially access these technologies in their homes, applied games and VR may help to overcome geographical barriers to accessing treatment, reduce other barriers to face-to-face interventions (e.g., stigma), and promote the reach of interventions to children and young people who would not normally seek help through traditional mental health services (Fleming et al., 2017; Freeman et al., 2017; Lau, Smit, Fleming, & Riper, 2017).

Wider benefits of applied games and VR may also come from their automated capability. Typically, VR has been used by therapists as an adjunct to face-to-face intervention, but there is a new generation of automated VR cognitive treatments which bring the potential to widely expand opportunities for access to effective treatments (Freeman et al., 2018). For example, reducing reliance on therapist delivery can reduce health care costs (Lambe et al., 2020), and standardized treatment approaches can be effectively disseminated by making sure that key 'treatment ingredients' are built in and always delivered (Farrell et al., 2020; Lambe et al., 2020). While many of these advantages will apply across different forms of digital intervention, applied games, and VR also bring the potential to overcome practical challenges in, for example, exposing people to certain fears that may otherwise be costly, difficult, or even dangerous to reproduce in real situations (e.g., repeating airplane take-offs and hurricanes; Farrell et al., 2020; Freeman et al., 2017; Lambe et al., 2020). Furthermore, applied games and VR may overcome some of the challenges of adherence and sustained engagement with self-directed interventions, for example, by offer a greater degree of support (e.g., from a virtual therapist). In addition, the fact that the user has control over the frequency and intensity of exercises and that VR and gaming environments can be adjusted to each user's specific needs may also enhance treatment adherence and acceptability (Hollis et al., 2017).

With these considerations in mind, the use of applied games and VR to target mental health problems in children and young people appears to be a logical step to increase access to, engagement with, and, potentially, the effectiveness of psychological treatments. To date, several applied games and VR interventions have been specifically developed for children and young people. A recent review of applied games showed moderate effects in reducing symptoms of depression in young people (Lau et al., 2017), but less is known about their effectiveness in targeting other mental health problems and a systematic review of VR-based therapies in this population has not yet been conducted.

Aims of this review

The aim of the current review is to provide an up-to-date evaluation by means of a systematic review of studies that have assessed the effectiveness of applied games or VR in treating mental health problems in children and young people. Specifically, we aim to identify and synthesize current data on studies that include an active intervention that involves at least one (i) applied game element, or (ii) VR element, which aimed to target mental health problems in children and young people. In addition, we set out to explore children and young people's experience (e.g., acceptability, adherence, expectations, and evaluations) of using these digital interventions.

Methods

The systematic review was conducted in accordance with guidance in the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2009) statement and the protocol was preregistered with PROSPERO (ID: CRD42020163056; available from https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020163056). Four electronic databases, MEDLINE, PsycINFO, CINAHL, and Web of Science Core Collection, were searched. Database searches were conducted on 5th November 2019 and were limited to all papers from 1990, to reflect the first widespread commercial release of consumer VR headsets. No other restrictions were applied during the search phase. Additionally, we conducted backward and forward citation hand searches for all studies included in the review in March 2020. The search string is available via the review's PROSPERO record.

Eligibility criteria

The inclusion and exclusion criteria were piloted and refined by four review authors (BH, CC, PW, and CH) using a sub-sample of papers. Studies were deemed eligible for inclusion if they met the following criteria:

1. The paper was available in English, in a peer-review journal.
2. The paper reported on humans.
3. The paper reported novel findings. Papers reporting reviews, meta-analyses, biographies, clinical guidelines, dissertations, theses, commentaries, or summaries of previous reported research were not included.
4. The paper reported on children and adolescents up to (and including) age 18 years. Due to the scarcity of research in these populations, studies including participants with an upper age limit of 21 years were included if the average age of the sample was less than 18 years.
5. The paper reported on participants that were selected for inclusion on the basis of meeting diagnostic criteria for a mental health disorder or showing elevated symptoms of mental health problem/s. In line with the typical configuration of children's services (mental health vs neurodevelopmental), we excluded neurodevelopmental disorders and their symptoms (e.g., Autistic Spectrum Disorders, and Attention Deficit and Hyperactivity Disorder) where aspects of the neurodevelopmental disorder (e.g., social skills and impulsivity) were the target of the intervention, although studies in which mental health problems were targeted amongst children with neurodevelopmental disorders were included (so long as other inclusion/exclusion criteria applied).
6. The paper included an active intervention involving at least one applied game element or VR element—which aimed to target mental health problems. Applied games were defined as 'digital interventions that employ games (or substantial game elements) in an effort to educate and/or change patterns of experience and/or behavior'. Thus, interventions where the digital 'game element' was used to aid treatment (e.g., *to give participants a bgreak from treatment, or to provide a reward for participating in treatment*) as opposed to directly targetting mental health symptoms were not eligible. VR was defined as 'the use of computer modeling and simulation that enables a person to interact with an artificial three-dimensional (3D) visual or other sensory environment.
7. The paper reported outcome/s using any of the following:

   a. A recognized diagnostic tool for DSM or ICD mental health disorder (completed by child/adolescent and/or parent)
   b. A validated measure of symptoms of DSM or ICD mental health disorders (completed by child/adolescent, parent, and/or teacher)
   c. Outcomes related to children and/or young peoples' experience (i.e., adherence, expectations, evaluations, and acceptability)

Note: Because of the early stage of the development of this field, we did not restrict inclusion on the basis of study design and therefore included both randomized controlled trials and case studies/series. We also included studies exploring children and young people's experiences of VR and/or applied game-based interventions, which included qualitative approaches (e.g., data from interviews).

Papers were excluded if the study was a universal intervention and/or in a non-clinical population without elevated symptoms of mental health problems. Additionally, papers were excluded if the VR or applied game element were examined in the context of intellectual disabilities or physical health conditions. Finally, where studies included qualitative data, collected through interviews or open-ended responses to questionnaires, this was only included in the analysis if the feedback was collected in a systematic way (e.g., if quotes were given with information of how participant feedback was elicited).

Study selection

A flowchart of the study selection process is shown in Figure 1. All electronic database search results were exported to Endnote version X9 (The Endnote Team, 2013). The searcheens retrieved 11,083 records; 7,944 of which were retained after duplicate records were removed. For quality assurance of study identification, two reviewers (BH and KP) screened all titles and abstracts of identified studies. Inter-rater reliability between the two reviewers was calculated at the initial phase of title/abstract screening as 99%, kappa = 0.946. Where reviewers disagreed at the title/abstract stage, papers went through to full-text screening. Abstract screening led to the exclusion of 7,160 articles; full-text articles for remaining 784 citations were reviewed for eligibility. A paper could be excluded at any stage of the full-text screening process on the basis of a 'no' response to any of the eligibility criteria; the first criterion that was not met was recorded as the reason for rejection. Duplicates were removed at both the title/abstract and full-paper screening stages. Reference lists of retained articles were inspected for relevant studies, and we also conducted hand searches and citation chaining to identify additional studies; bibliographic databases were used again to retrieve abstracts, and, if appropriate, full-text articles. A total of 19 studies were included in the systematic review. Inter-rater reliability between the two reviewers (BH and KP) for the inclusion/exclusion of full-text papers was 93.8%, kappa = .68. For papers that were accepted via the full-text paper screening, appropriate data were extracted by two reviewers (BH and KP) and then reviewed to ensure accuracy. Disagreements among reviewers were initially discussed by the two review authors (BH and KP) and if consensus was not reached, other review authors (CC/DF/CH) were consulted to reach a final decision.

Data synthesis

Due to considerable heterogeneity among the studies included in this review, we have adopted a descriptive approach to data synthesis, whereby short summaries of included studies are presented.

Quality rating

We assessed the quality of studies using two rating checklists (one for quantitative and one for qualitative studies) developed by Kmet, Cook, and Lee (2004). This is an appraisal tool appropriate for rating studies with a variety of designs. If a study included both quantitative and qualitative methodology, they were rated on each scale. Each checklist item was rated on a 0–2 scale (0 = not met; 1 = partially met; 2 = fully met). The quantitative checklist included 13 items (maximum score of 26) and the qualitative checklist included 10 items (maximum score of 20). On the quantitative checklist, where items were not applicable to the study design (e.g., power analyses for case studies), the item was not included in the calculation of the summary score. A summary score was calculated for each paper by summing the total score obtained across relevant items and dividing by the total possible score giving a score between 0 and 1. Therefore, the scores are adjusted according to their study design, and, although there are no direct benchmarks for appraising the quality, this does allow for a direct comparison of all the studies identified in the
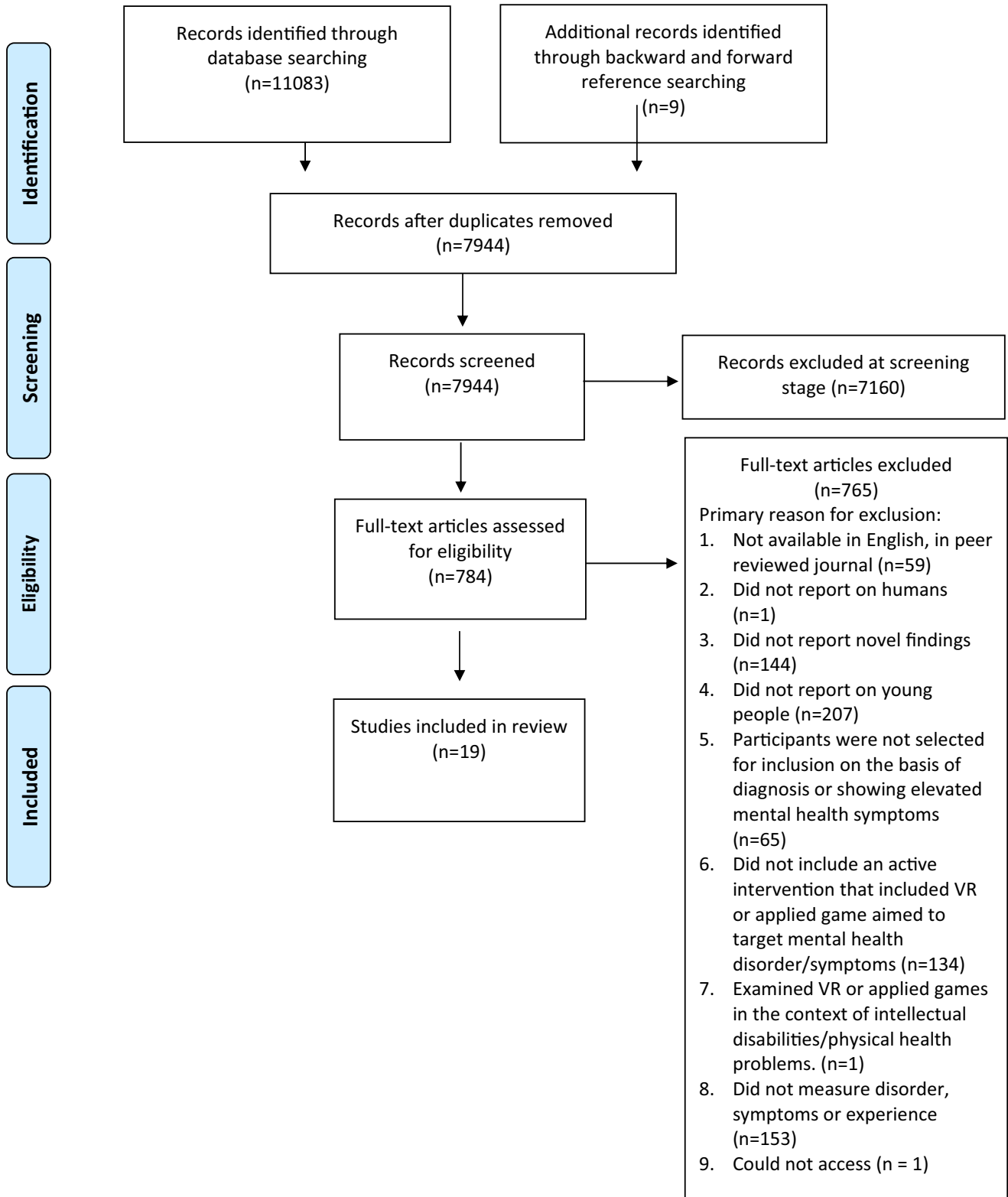
Figure 1 PRISMA flowchart of study selection process [Colour figure can be viewed at wileyonlinelibrary.com]

review. Each study was assessed and independently rated by CH and PW, who then discussed discrepancies and agreed consensus ratings. Twelve studies were rated using the checklist for quantitative studies, 2 studies were assessed using the checklist for qualitative studies, and 5 studies that used both qualitative and quantitative methods were assessed using both checklists. Regardless of quality classification, all studies were included in the review.

Estimation of effect sizes

Where possible we calculated both within (e.g., pre-post) and between group study effect sizes. It is important to note the limitations of within-group effects (e.g., they may simply reflect a regression to the mean) and priority should be given to the between group effects that are reported. However, as some studies did not include any control condition, we included the

within-group effects so that we had some way of comparing outcomes across different interventions. For continuous outcomes, effect size calculations were was based on the reported mean questionnaire score at pre- and post-intervention, and their standard deviations using the following online calculator: https://www.psychometrica.de/effect_size.html. When studies were not explicit in what their primary outcome measure was and/or used multiple measures, we selected the measure that had been standardized in children and young people and most in line with the study aims (i.e., if the intervention focused on reducing anxiety problems we chose an anxiety measure). Effect sizes were interpreted using Cohen's (1988) suggested reference values of 0.20, 0.50, and 0.80 as small, medium, and large, respectively. Several studies also reported outcomes after one or more follow-up periods, which varied from one to sixteen months. Separate effect size calculations were conducted for these studies where pre-treatment mean scores and follow-up mean scores and their standard deviations were used. Three studies did not provide relevant data to allow for effect size calculations at any time-point. For one study (Fleming, Dixon, Frampton, & Merry, 2012) Cohen's d was based on the Cohen's d reported in the paper. Effect sizes were coded as positive or negative to aid interpretation of the data. A positive effect size for within-group differences indicates an increase in symptom score. For between group comparisons, a positive effect size indicates that participants receiving the digital intervention had a lower symptom score. For change in dichotomous outcomes (i.e., diagnostic status and/or remission rates), odds ratios were transformed into Cohen's d—a positive effect size for diagnostic status or remission, indicates that a higher proportion of participants receiving the digital intervention no longer met diagnostic status or were in remission.

Results
Study characteristics

Applied games. Sixteen (84.2%) of the 19 studies explored children and young people's experience of using applied games and/or their effectiveness in targeting mental health problems (see Tables 1 and 2), and examined nine different applied games.

Two research groups authored ten of the sixteen studies. Out of the sixteen studies, eight (50%) were randomized controlled trials (RCTs; Fleming et al., 2012; Merry et al., 2012; Scholten, Malmberg, Lobel, Engels, & Granic, 2016; Schoneveld, Lichtwarck-Aschoff, & Granic, 2018; Schoneveld et al., 2016; Stasiak, Hatcher, Frampton, & Merry, 2014; Wijnhoven et al., 2020; Wols, Lichtwarck-Aschoff, Schoneveld, & Granic, 2018); six (37.5%) included qualitative data (Bobier, Stasiak, Mountford, Merry, & Moor, 2013; Carlier et al., 2019; Carrasco, 2016; Coyle, McGlade, Doherty, O'Reilly, & Acm, 2011; Fleming, Lucassen, Stasiak, Shepherd, & Merry, 2016; Schoneveld, Lichtwarck-Aschoff, & Granic, 2019); two (12.5%) were open pilot feasibility studies (Bobier et al., 2013; Lucassen, Merry, Hatcher, & Frampton, 2015); one (6.3%) was a controlled clinical trial (Knox et al., 2011); and one (6.3%) was a case study (Carlier et al., 2019). Out of the nine applied games, two (22.2%) have been evaluated more than once (i.e., MindLight and SPARX). Across these studies, participant ages ranged from 7 to

19 years. Study sample sizes ranged from 2 to 187 participants (mean = 62.06; SD = 63.48) and all studies involved children and young people of mixed gender except for Carrasco (2016) and Carlier et al. (2019) which only included females. The majority of studies were conducted in high-income countries, including the United Kingdom, Netherlands, and New Zealand.

Four (44.4%) of the nine applied games (i.e., MindLight, New Horizon, Dojo, and gNATS Island) focused on anxiety symptoms/disorders and were conducted with children and young people with elevated symptoms of anxiety in general populations (Knox et al., 2011; Scholten et al., 2016; Schoneveld et al., 2016, 2018, 2019; Wols et al., 2018), children and young people with autism spectrum disorder (ASD; Carlier et al., 2019; Wijnhoven et al., 2020), or children and young people referred for the treatment of anxiety problems (Coyle et al., 2011). Three (33.3%) applied games (i.e., SPARX, The Journey and The Quest for the Rest) targeted depression. They were conducted with children and young people with elevated symptoms of depression (Carrasco, 2016; Fleming et al., 2012, 2016; Merry et al., 2012; Stasiak et al., 2014) or admitted for severe psychiatric disorder including depression (Bobier et al., 2013). One (11.1%) program (Rainbow SPARX) specifically targeted elevated levels of depression in 'sexual minority' youth (Lucassen et al., 2015) and one (11.1%) program (which included two games: The Journey to the Wild Divine and Freeze Framer) targeted both symptoms of anxiety and depression and was conducted in children and young people with sub-clinical or clinical levels of anxiety (Knox et al., 2011).

Descriptions of each game are listed in Table 3. The applied games mainly implemented CBT through a variety of approaches, some using a limited number of treatment components (e.g focusing mainly on relaxation), and some using a range of interventions delivered over a longer time.

Virtual reality. Three (15.8%) of the 19 studies investigated children and young people' experience of VR and/or its effectiveness in targeting mental health symptoms (see Tables 1 and 2), and examined two VR applications. Out of the three studies, one (33.3%) was a RCT (Maskey, Rodgers, et al., 2019); one study (33.3%) was a pilot study without a control condition (Maskey, Lowry, Rodgers, McConachie, & Parr, 2014); and one study (33.3%) reported qualitative data (Parrish, Oxhandler, Duron, Swank, & Bordnick, 2016). Participant age ranged from 7 to 18 years. Study sample sizes ranged from 9 to 41 participants (mean = 27.33; SD = 16.50) and all studies involved children and young people of both genders, except for Maskey et al. (2014) which only included males. Studies were all conducted in high-income countries, including the United Kingdom and United States. The two VR applications were

Table 1 Studies that include information on children and young people's experience

Study	Country	Name of intervention	Target population	Age range (years)	Setting	Total sample size (% female)	How experience was measured (reported by)
Applied games focusing on anxiety							
Wijnhoven et al. (2020)	Netherlands	MindLight	Children and adolescents with ASD and elevated symptoms of anxiety	8 to 16	Clinic	109 (22.9)	Game adherence (child/adolescent)
Schoneveld et al. (2018)	Netherlands	MindLight	Children with elevated symptoms of anxiety	7 to 12	School	174 (59.20)	Game expectations (child); Game appeal (child)
Schoneveld et al. (2016)	Netherlands	MindLight	Children and adolescents with elevated symptoms of anxiety	8 to 13	School	136 (54.8)	Game expectations (child/adolescent); Game appeal (child/adolescent)
Schoneveld et al. (2019)	Netherlands	MindLight	Children and adolescents with elevated symptoms of anxiety	8 to 13	School	13 (46.15)	Game acceptability (child/adolescent)
Coyle et al. (2011)	UK	gNATS Island	Children and adolescents with anxiety difficulties (referred for treatment)	11 to 16	Clinic	6 (33.33)	Game acceptability (child/adolescent)
Applied games focusing on depression							
Bobier et al. (2013)	New Zealand	SPARX	Adolescents admitted for severe psychiatric disorder (including depressive disorders)	16 to 18	Inpatient	20 (40)	Game uptake (adolescent); Game adherence (adolescent); Game acceptability (adolescent)
Fleming et al. (2016)	New Zealand	SPARX	Adolescents with elevated symptoms of depression	13 to 16	School	39 (38.5)	Game acceptability (adolescent); Game adherence (adolescent) Game experience (adolescent)
Fleming et al. (2012)	New Zealand	SPARX	Adolescents with elevated symptoms of depression	13 to 16	School	32 (44)	Game adherence (adolescent)
Merry et al. (2012)	New Zealand	SPARX	Children and adolescents with elevated symptoms of depression	12 to 19	School	187 (62.8)	Game adherence (child/adolescent); Game acceptability (child/adolescent)
Lucassen et al. (2015)	New Zealand	Rainbow SPARX	Adolescents with elevated symptoms of depression who are also sexually attracted to the same sex, both sexes, or who are questioning their sexuality (i.e., sexual minority youth).	13 to 19	School; Home; Research center	21 (47.6)	Game adherence (adolescent); Game acceptability (adolescent)
Stasiak et al. (2014)	New Zealand	The Journey	Adolescents with elevated symptoms of depression	13 to 18	School	34 (41)	Game acceptability (adolescent)
Carrasco (2016)	Chile	The Quest for the Rest	Children and adolescents with elevated symptoms of depression	12 to 18	Clinic	15 (100)	Game acceptability (child/adolescent)
Virtual Reality focusing on phobia, social anxiety, and/or trauma							
Maskey, Rodgers, et al. (2019)	UK	Blue Room	Children and adolescents with ASD and specific phobia	7 to 15	Clinic	32 (21.9)	VR adherence (child/adolescent)
Parrish et al. (2016)	USA	Series of social-related VR environments	Adolescents with elevated symptoms of social anxiety	13 to 18	School	41 (65.9)	VR realism/presence/immersion (adolescent); VR acceptability (adolescent)

developed for targeting specific phobia (i.e., 'Blue Room'; Maskey et al., 2014; Maskey, Rodgers, et al., 2019) or social anxiety (Parrish et al., 2016). Studies were conducted with children and young people with autism spectrum disorder and specific phobia (Maskey et al., 2014; Maskey, Rodgers, et al., 2019) or general populations with symptoms of social anxiety. The VR components of included studies are listed in Table 3.

Outcomes

Experience. Children and young people's experience (i.e., adherence, expectations, evaluations, and acceptability) of applied games or VR are summarized in Table 4.

Twelve (75%) of the sixteen studies included children and young people's evaluations of applied games (Bobier et al., 2013; Carrasco, 2016; Coyle et al., 2011; Fleming et al., 2012, 2016; Lucassen et al., 2015; Merry et al., 2012; Schoneveld et al., 2016, 2018, 2019; Stasiak et al., 2014; Wijnhoven et al., 2020). Overall, the majority of children and young people completed the required treatment modules (e.g., Fleming et al., 2016), expected the applied games to be helpful in targeting their mental health symptoms (prior to use; e.g., Schoneveld et al., 2018), found them relevant (e.g., appropriately anxiety inducing; e.g., Schoneveld et al., 2019) and acceptable (e.g., enjoyable and useful; e.g., Bobier et al., 2013). The two studies that also evaluated experience of another computerized intervention, including non-therapeutic commercial games, reported no significant group differences in treatment adherence (Wijnhoven et al., 2020) and expectations for lowering fear (prior to use; Schoneveld et al., 2016). However, one study also compared the experience to face-to-face treatment (i.e., group CBT) and found that children and young people rated group CBT as significantly more relevant to their daily lives than the applied game post-treatment and at a 3-month follow-up (with medium to large between-group effects, respectively), but not at a 6-month follow-up (small between-group effect; Schoneveld et al., 2018). However, both groups rated their intervention as equally appealing to themselves and others and no group differences were found on reported difficulty or the extent to which the interventions induced anxiety.

For VR, two of the three studies included children and young people's experience of the VR application, showing that participants completed all sessions (Maskey, Rodgers, et al., 2019) and found the VR environments appropriately anxiety provoking (Parrish et al., 2016; see Table 3). Neither study compared the experience of VR to a different intervention.

Mental health symptoms. Effect sizes for the self-, parent- and, where reported, clinician-rated

outcomes for applied games and VR are shown in Table 5.

For applied games that focused on anxiety, four out of six studies used MindLight and found small to large within-group effects post-treatment and at follow-up on self- and parent-reported symptoms of anxiety within ASD (Wijnhoven et al., 2020) and general populations with elevated levels of anxiety (Schoneveld et al., 2016, 2018; Wols et al., 2018). However, when compared to another intervention, children and young people with autism who used MindLight did not self-report significantly greater symptom reduction at any time-point compared to children and young people that played a non-therapeutic commercial game (small between-group effects; Schoneveld et al., 2016; Wijnhoven et al., 2020). This is consistent with Scholten et al. (2016), who also found non-significant and small between-group effects between Dojo and a commercial game. These non-significant group differences were corroborated by parent report, except for in Wijnhoven et al. (2020) where parents rated children and young people who played MindLight as significantly less anxious (with small between-group effects) at follow-up (but not post-treatment) compared to parents of children who played a commercial game. Notably, however, in the only study to compare an applied game to a face-to-face intervention, Schoneveld et al. (2018) found that MindLight was as effective as face-to-face CBT in reducing anxiety symptoms in a sample of pre-adolescent children (7- to12-year-olds). However, notably, the within-group effect sizes for both interventions were similar in size to that found for the commercial game (not aimed at reducing anxiety symptoms) in Wijnhoven et al. (2020).

In the four studies targeting children and young people with depression or elevated symptoms of depression, three included the applied game SPARX and found medium to large within-group effects on depressive symptoms at post-treatment and follow-up (Fleming et al., 2012; Lucassen et al., 2015; Merry et al., 2012). Furthermore, at post-treatment, adolescents receiving SPARX were significantly more likely to be in remission (medium effect size) and have lower depressive symptoms (effect size not reported) than adolescents on a waiting list (Fleming et al., 2012). When compared to treatment as usual (face-to-face counseling), playing SPARX was associated with similar improvements in remission rates and depressive symptoms at all time-points (small effect size; Merry et al., 2012). Only one study has compared an applied game for depression to another computerized intervention (psychoeducation). Both interventions had large within-group effects on clinician-rated depression severity. Whilst adolescents who played The Journey were rated as significantly less depressed at post-treatment (medium effect size), changes in remission rates did not differ significantly between groups (although there was a large between-group effect size; Stasiak et al., 2014).

Table 2 Studies that have information on mental health symptom/disorder outcomes from applied games/VR

Study	Country	Name of intervention	Target population	Age range (years)	Setting	Total sample size (% female)	Outcome variable (reported by)
Applied games focusing on anxiety							
Wijnhoven et al. (2020)	Netherlands	MindLight	Children and adolescents with ASD and elevated symptoms of anxiety	8 to 16	Clinic	109 (22.9)	Symptoms severity (child/adolescent/parent); Remission rates (clinician)
Schoneveld et al. (2018)	Netherlands	MindLight	Children with elevated symptoms of anxiety	7 to 12	School	174 (59.20)	Symptoms severity (child/parent)
Wols et al. (2018)	Netherlands	MindLight	Children with elevated symptoms of anxiety	8 to 12	School	43 (53.48)	Symptoms severity (child); In-game play behaviors (research assistants)
Schoneveld et al. (2016)	Netherlands	MindLight	Children and adolescents with elevated symptoms of anxiety	8 to 13	School	136 (54.8)	Symptoms severity (child/adolescent/parent)
Carlier et al. (2019)	Belgium	New Horizon	Children with ASD and elevated symptoms of anxiety	8 to 10	Clinic	2 (100)	Symptoms severity (child/parent)
Scholten et al. (2016)	Netherlands	Dojo	Children and adolescents with elevated symptoms of anxiety	11 to 15	School	138 (65.0)	Symptoms severity (child/adolescent)
Applied games focusing on depression							
Fleming et al. (2012)	New Zealand	SPARX	Adolescents with elevated symptoms of depression	13 to 16	School	32 (44)	Symptoms severity (adolescent/clinician); Remission rates (clinician); Clinically significant change (clinician)
Merry et al. (2012)	New Zealand	SPARX	Children and adolescents with elevated symptoms of depression	12 to 19	School	187 (62.8)	Symptoms severity (child/adolescent/clinician); Remission rates (clinician); Treatment response (clinician)
Lucassen et al. (2015)	New Zealand	Rainbow SPARX	Adolescents with elevated symptoms of depression who are also sexually attracted to the same sex, both sexes, or who are questioning their sexuality (i.e., sexual minority youth).	13 to 19	School; Home; Research center	21 (47.6)	Symptoms severity (adolescent);
Stasiak et al. (2014)	New Zealand	The Journey	Adolescents with elevated symptoms of depression	13 to 18	School	34 (41)	Symptoms severity (adolescent/clinician); Remission rates (clinician); Treatment response (clinician)

(continued)

Table 2 (continued)

Study	Country	Name of intervention	Target population	Age range (years)	Setting	Total sample size (% female)	Outcome variable (reported by)
Applied games focusing on anxiety and depression							
Knox et al. (2011)	USA	The Journey to the Wild Divine; Freeze Framer	Children and adolescents with sub-clinical or clinical anxiety problems	9 to 17	Clinic	24 (37.5)	Symptoms severity (child/adolescent)
Virtual Reality focusing on phobia, social anxiety, and/or trauma							
Maskey, Rodgers, et al. (2019)	UK	Blue Room	Children and adolescents with ASD and specific phobia	7 to 15	Clinic	32 (21.9)	Symptoms severity (child/adolescent/parent); Phobic behavior rating (research assistant); Treatment response (research assistant)
Maskey et al. (2014)	UK	Blue Room	Children and adolescents with ASD and specific phobia	7 to 13	Clinic	9 (0)	Symptoms severity; Phobic behavior rating; Confidence ratings

In the only trial of an applied game that targeted symptoms of both anxiety and depression in a treatment-seeking sample, the game was associated with a significant advantage in terms of symptoms of anxiety (large effect) and depression (small effect) compared with a waiting list (Knox et al., 2011).

For VR, two out of three studies focused on specific phobias. Both used 'Blue Room' and when evaluated in children and young people with ASD, they found small to medium within-group effects on child- and parent-reported anxiety/phobia symptoms (Maskey et al., 2014; Maskey, Rodgers, et al., 2019). However, no significant group differences were found when Blue Room was compared to a waiting list control (medium effect size; Maskey, Rodgers, et al., 2019). The third study (Parrish et al., 2016) focused on social anxiety in adolescents, but no data were provided for symptom outcomes (only qualitative information was provided).

Quality ratings

Quality ratings ranged from 0.25 to 0.96 (out of a possible range from 0 to 1) with an average quality rating of 0.80 for quantitative studies and from 0.2 to 0.85 (out of a possible range of 0 to 1) with an average quality rating of 0.43 for qualitative studies. For quantitative studies, higher quality studies generally scored highly for the research question being sufficiently described, participant selection, and sample size. Areas where studies tended to receive lower scores were describing the sample's characteristics, randomizing participants to intervention groups, reporting well defined and robust outcome measures, and reporting of the blinding of investigators.

For qualitative studies, often the research questions, study context, and overall study design were reasonably well described; however, most studies were limited in terms of the connection to theory or wider knowledge, explanation of how the data were collected, and methods to verify the findings. Notably, none of the qualitative studies demonstrated any evidence of reflexivity.

Discussion

This systematic review identified 19 studies that have examined children and young people's experience of and the effectiveness of using applied games or VR for mental health problems. Despite the enthusiasm and promise of this line of intervention, it is important to highlight that the evidence to date is at a very early stage with studies being limited to interventions for anxiety, depression, and phobias only. For applied games, overall, there is evidence to suggest that children and young people find them helpful, enjoyable, and engage with them. However, children and young people may not necessarily find them relevant for addressing their mental health problems. Nonetheless, when it comes to treatment

Table 3 Description of applied games and VR

Applied Games	
Mindlight	A horror-themed survival video game for 8- to 12-year-old children based on CBT techniques: relaxation through neurofeedback (the player wears an EEG headset), exposure training, and attention-bias modification. Using these techniques, the game aims to teach children how to cope with anxiety.
New Horizon	A 2D mobile Android exploration game made up of four mini-games, two of which contain CBT techniques. The 'Senses' and 'Breathing' mini-games incorporate visualization as a relaxation technique and focused breathing.
Journey to the Wild Divine / Freeze Framer 2	Journey to the Wild Divine involves an assortment of experiences in a fantasy land, for example, the user has a goal of building a bridge across a valley. Imagery and sound are used to aid relaxation. As the user's breathing slows and tension decreases (measured using heart rate variability and skin conductance), the bridge is built. If the user experience frustration or anxiety, the bridge disappears.
	Freeze Framer 2 allows the player to engage in activities such as coloring a meadow, making a rainbow, or floating in a hot air balloon on the computer screen. Imagery and sound are used to aid relaxation. Heart rate variability and skin conductance level are measured.
Dojo	A motion management video game designed to reduce anxiety in adolescents. Dojo incorporates two evidence-based strategies: emotion regulation training and heart rate variability biofeedback. Emotion regulation strategies are practiced in challenges that become increasingly difficult if the player's heart rate increases.
gNATS island	A computer game designed to support face-to-face CBT interventions for adolescents aged 10–15. While navigating through a 3D tropical island, players encounter little creatures called gNATS which represent automatic negative thoughts (of which there are nine). Through conversations with game characters, players are introduced to strategies for identifying and challenging negative thoughts.
SPARX	An interactive fantasy game designed to deliver CBT for the treatment of adolescent mild to moderate depression. The player undertakes a series of challenges to restore balance in the fantasy world dominated by GNATS (Gloomy Negative Automatic Thoughts). The content in the seven levels include psychoeducation, relaxation skills, interpersonal skills, activity scheduling, problem-solving, SPARX (Smart Positive Active Realistic X-factor thoughts), cognitive restructuring, distress tolerance, and relapse prevention.
The Journey	A computerized CBT program for depressed adolescents. The player follows a quest through a fantasy environment. The game comprises seven modules, each with a different topic; introduction to CBT model, behavioral activation, problem-solving, cognitive restructuring (identifying and challenging unhelpful thoughts), relaxation techniques, relapse prevention. Information is presented through interactive exercises, animations and illustrative video clips.
The Quest for the Rest	A video game that follows the story of a teenager called Maya who is feeling sad. The game incorporates and scores game behavior in the areas of recognition and modification of negative cognitive bias, interpersonal skills and interpersonal problem-solving and behavioral activation and a healthy lifestyle. Feedback is given to reinforce positive behavior.
Virtual Reality	
Blue Room	A fully immersive virtual reality environment (VRE) that uses interactive computer-generated audio-visual images projected onto the walls and ceilings of a 360 degree screened room (no need for a headset or goggles). The Blue Room is suitable for phobias that can be visually represented and addressed. A therapist delivers CBT techniques whilst in the room with the participant. Scenes are individualized, incorporating an exposure hierarchy related to the feared stimulus.
Series of social-related VR environments	VR public speaking environment: Participants were asked to give a speech in front of a virtual audience using a head-mounted display that tracks movement. The audiences was prompted to look sleepy, distracted, as though they disagreed, and were puzzled in a consistent manner during the speech for each participant. At the end of the scenario, the audience was prompted to clap politely. VR part environment: Participants start on the walkway outside a house party where party music was playing (in the head set) and individuals can be seen visiting inside an open front door at the party. Participants were encouraged to interact naturally with others. This environment runs on an automatic timer, moving the participant through various parts of the home where the participant is exposed to several social interactions with individuals.

of depression, there is some cause for optimism about the potential for applied games as studies reported significant group differences with medium effect sizes for both changes in symptoms and remission rates and with more robust support for SPARX than any other applied game. Specifically, remission was significantly greater among adolescents that played SPARX compared to a waiting list (Fleming et al., 2012). Furthermore, SPARX achieved similar outcomes to an alternative, face-to-face counseling treatment (Merry et al., 2012), with both interventions achieving medium to large within-group effects, which are similar to the effects found for other face-to-face psychological interventions for adolescent depression (Goodyer et al., 2017). When it comes to anxiety, however, there is greater need for caution. Here, pre-post effect sizes were typically in the small to medium range (whereas these are typically large for face-to-face CBT; James, James, Cowdrey, Soler, & Choke, 2015), and comparisons to

Table 4 Platform evaluations

Study	Name of intervention	Comparator Computerized	Face-to-face	W/L	None	Platform evaluation measure	Results
Applied games focusing on anxiety							
Wijnhoven et al. (2020)	MindLight	√				**Game adherence** % of participants that completed 6/6 sessions of MindLight or commercial game	**Game adherence** No significant group differences (73.6% for MindLight, 80.4% for commercial game).
Schoneveld et al. (2018)	MindLight		√			**Game/group CBT expectations** Participants read a short description of both interventions at baseline and rated how helpful they felt they would be for lowering their fear. **Game/Group CBT acceptability** Participants rated post-treatment five statements that related to the intervention they were assigned to: 'I found it fun to participate in MindLight/CBT'; 'I think __ is fun for other children'; 'I can use what I learned from __ in my daily life well'; 'I found some exercises in __ stressful'; 'I found some exercises in __ difficult'.	**Game/Group CBT expectation** No significant group differences. **Game/Group CBT acceptability** Group CBT rated as significantly more relevant to participant's daily life at post-treatment and 3-month follow-up. No other significant group differences emerged.
Schoneveld et al. (2016)	MindLight	√				**Game expectations** Participants read a short description of both games at baseline and rated whether they believed their 'real-life' behavior could be improved by playing them. **Game acceptability** Ratings of six game evaluation items, that is, Game appeal; Appeal to others; Relevance; Flow; Anxiety inducing; Difficulty	**Game expectations** No significant group differences emerged. **Game acceptability** MindLight rated significantly more anxiety-inducing; commercial computer game rated significantly more appealing to 'myself' and more likely to induce feelings of flow; No significant group differences on reported difficulty, relevance, and the extent to which children believed the games would appeal to other children.
Schoneveld et al. (2019)	MindLight				√	**Game play experience** Semi-structured interviews exploring participant's views on the game's motivational characteristics, including what in-game activities participants liked more and less, and what they would add or remove.	**Game play experience** Participants liked the overall look and feel of MindLight; game successful in evoking anxiety; variation in how challenging participants found the game.
Coyle et al. (2011)	gNats island				√	**Game acceptability** Participants rated how enjoyable and helpful they found the game; whether they would recommend the game to a friend; and whether they have changed in any way since completing the game.	**Game acceptability** Enjoyment: 83% found gNats Island to be 'very' or 'extremely' enjoyable / 16% found it not to be enjoyable at all; Helpfulness: 50% found gNats Island to be 'very' or 'extremely' helpful / 50% found it 'kind of helpful'; Recommend to a friend: 83% would recommend to a friend / 16.67% would not recommend to a friend; Change in self: 100% reported they had changed since completed the game (i.e., feeling less worried, more positive, more confident, and feeling better).

(continued)

Table 4 (continued)

Study	Name of intervention	Comparator				Platform evaluation measure	Results
		Computerized	Face-to-face	W/L	None		

Applied games focusing on depression

Study	Name of intervention	Computerized	Face-to-face	W/L	None	Platform evaluation measure	Results
Bobier et al. (2013)	SPARX				√	**Game Adherence** The number (%) of young people who chose to access SPARX whilst in an inpatient unit. % of participants that completed 7/7 modules. **Game acceptability** Satisfaction questionnaire.	**Game adherence** Over 9-months, 42 inpatients, 24 (57%) were eligible and invited to use SPARX; 22 (92%) accepted the invitation; 20 (91%) patients trialed SPARX; 2 (9%) dropped out. 10% finished 7/7 modules (33% completed at least 4/7). **Game acceptability** Usefulness: 86% felt that SPARX was useful or very useful; Appeal to others: 93% felt that SPARX would appeal to other young people; Recommend to friends: 79% said they would recommend SPARX to their friends; Other: One patient (7%) said they preferred individual talking therapy to using the computer whilst another felt that using SPARX made her feel worse about her illness; Two participants said they did not like SPARX in general.
Fleming et al. (2016)	SPARX				√	**Game adherence** % of participants that completed 7/7 modules **Game acceptability** Satisfaction questionnaire **Game experience** Semi-structured interview exploring: (i) Personal experiences of SPARX and (ii) Views about how SPARX might best be used to support others	**Game adherence** 72% finished 7/7 modules. **Game acceptability** 86% found SPARX useful; 86% would recommend SPARX to a friend **Game experience** Personal experiences of SPARX (themes): *Overall feedback*: 'fun and helpful'; 'entertaining'; 'it feels caring'; 'limited in terms of gaming'; *Personal impact of cCBT*: 'it's more about anger than depression'; 'transforming anger'; 'Transforming thinking'; 'It wasn't helpful'; *Completion*: 'Narrative'; 'Schooling environment' Views about how SPARX might best be used to support others (themes): *Impact of cCBT on help seeking*: 'It smoothes the way'; 'I don't want/need therapy now'; *Role for computerized therapy*: cCBT is freeing ('counseling is too much'; 'cCBT is empowering and easy to learn from'); *Therapeutically limited*; *Universal or individual use*: 'Everyone has down times'; It's too hard to target; 'It's boring, if you do not need it'; 'Need for varied approaches'.
Fleming et al. (2012)	SPARX			√		**Game adherence** % of participants that completed 7/7 modules	**Game adherence** 69% completing 7/7 modules (81% completed at least 4/7).

(continued)

Table 4 (continued)

Study	Name of intervention	Comparator				Platform evaluation measure	Results
		Computerized	Face-to-face	W/L	None		
Merry et al. (2012)	SPARX		√			**Game adherence** % of participants that completed 7/7 modules **Game acceptability** Satisfaction questionnaire.	**Game adherence** 60% completed 7/7 modules (86% completed at least 4/7). **Game acceptability** 95% of participants in the SPARX group and 98.6% in the treatment as usual group (*p* = 0.37) believed that the type of support they received would appeal to other teenagers; 80.5% of participants in the SPARX group and 95.8% in the treatment as usual group (*p* = 0.01) would recommend the treatment to their friends. Of those who completed SPARX, 53.2% would have liked the sessions to stay the length they were (most reported taking 20 to 40 minutes to complete each module); 44.3% wanted the sessions to be longer; 61.5% reported that they completed all or most of the set challenges ('homework').
Lucassen et al. (2015)	Rainbow SPARX				√	**Game adherence** % of participants that completed 7/7 modules **Game acceptability** Satisfaction questionnaire.	**Game adherence** 81% Completed 7/7 modules (90% completed at least 4/7). **Game acceptability** 80% participants indicated that they would recommend Rainbow SPARX to friends; 85% thought that the intervention would appeal to other young people.
Stasiak et al. (2014)	The Journey	√				**Game acceptability** Satisfaction questionnaire.	**Game acceptability** Did they like it?: 56% Liked it / 33% ok / 11% Did not like it How did they rate it?: 57% Excellent; 33% ok; 11% Poor Did they find it easy to use?: 67% Very easy;33% Mostly easy Did they find it useful? 56% Very/fairly useful; 44% OK, but could be improved; Would they recommend it to other adolescents?: 22% Would recommend once improved; 11% Would not recommend.
Carrasco (2016)	The Quest for the Rest				√	**Game acceptability** A satisfaction questionnaire - each item is a phrase that expresses an opinion about the value and benefit derived from playing the game. Participants were asked to express their level of agreement for each phrase by choosing one of the five possible answers (Nothing; Slightly; Moderately; A lot; Extremely)	**Game acceptability** The group acceptability mean score was 1.88 (SD = 1.09; 27% reported acceptability mean scores values below 1; 33% reported mean values between 1.5 and 2.5; 33% reported mean acceptability values above 2.5.

Virtual Reality focusing on phobia, anxiety, and/or trauma

(continued)

Table 4 (continued)

| Study | Name of intervention | Comparator | | | | Platform evaluation measure | Results |
		Computerized	Face-to-face	W/L	None		
Maskey, Rodgers, et al. (2019)	Blue Room			√		**VR adherence** % of participants that completed 4/4 VR sessions.	**VR adherence** 100% of participants completed 4/4 VR treatment sessions.
Parrish et al. (2016)					√	**VR Experience (realism/presence/immersion)** PQ: The extent of realism, interactions, involvement and naturalness of the VR experience IQ: The extent of presence and immersion in the virtual context **VR acceptability** Debriefing after VR session.	**VR Experience (realism/presence/immersion)** PQ: The non-socially anxious participants reported significantly higher levels of realism. IQ: No significant group differences **VR acceptability** Presentation environment: 98% rated the presentation environment as more anxiety provoking than the party environment; 61% described their reaction to the presentation environment using one of the following words: 'scared', 'nervous', 'worries', 'embarrassing', or 'afraid' 34% used words such as: 'fun', 'unprepared', 'genuine', 'weird', and 'cool'; 7% described the environment as 'unrealistic' 2% suggested he would be more nervous if the avatars were his own age; Party environment 20% used the word 'real' or 'normal' to describe their experience.

Table 5 Effect sizes and quality ratings across studies

| | | Quality Rating | | Comparator | | | | Effect size | | | | | |
		Quant	Qual	Computerized	Face-to-face	W/L	None	Within-group Platform Post-treatment	Within-group Platform FU	Comparator Post-treatment	Comparator FU	Between group Post-treatment	Between group FU
Study	Name of Intervention												
Applied games focusing on anxiety													
Wijnhoven et al. (2020)	MindLight	0.81		✓				SCAS-C = −.83$^{\wedge\circ}$•, SCAS-P = −.81°•	–	SCAS-C = −.64°, SCAS-P = −.81°	SCAS-C (3m) = −.78°, SCAS-P (3m) = −.81°	SCAS-C =.14, SCAS-P =.06	ADIS-IV-P any anxiety disorder (3m) =.10, SCAS-C (3m) =.00, SCAS-P (3m) =.46**
Schoneveld et al. (2018)	MindLight	0.88			✓			SCAS-C = −.60$^{\wedge\circ}$, SCAS-M= −.39°, SCAS-F = −.32°	SCAS-C (3m/6m) = −.75/−1.07°, SCAS-M (3m/6m) = −.47/−.60°, SCAS-F (3/6m) = −.36/−.62°	SCAS-C = −.63°, SCAS-M = −.44°, SCAS-F =−.42°	SCAS-C (3m/6m) = −.84°/−.88°, SCAS-M (3m/6m) = −.74°/.94°, SCAS-F (3m/6m) = −.61°/−.81°	SCAS-C =.03, SCAS-M =.00, SCAS-F =.10	SCAS-C (3m/6m) = −.05/.17, SCAS-M (3m/6m) = −.16/.16, SCAS-F (3m/6m) = −.21/−.17
Wols et al. (2018)	MindLight	0.88					✓	SCAC-C = °$^{\wedge}$	SCAS-C = −.79°	–	–	–	–
Schoneveld et al. (2016)	MindLight	0.88		✓				SCAS-C= −.28, SCAS-M = −.26, SCAS-F = −.30	SCAS-C (3m) = −.51*, SCAS-M (3m) = −.44*, SCAS-F (3m) = −.56*	SCAS-C = −.13, SCAS-M = −.27, SCAS-F = −.16	SCAS-C (3m) = −.52*, SCAS-M (3m) = −.26*, SCAS-F (3m) = −.16*	SCAS-C = .43, SCAS-M = .00, SCAS-F = .23	SCAS-C (3m) =.23, SCAS-M (3m) =.18, SCAS-F (3m) =.44
Carlier et al. (2019)	New Horizon	0.79	0.20				✓	SCAS-C = °$^{\wedge}$	–	–	–	–	–
Scholten et al. (2016)	Dojo	0.85		✓				SCAS-C = −.27	SCAS-C (3m) = −.35*	SCAS-C = −.23	SCAS-C (3m) = −.45*	SCAS-C =.11	SCAS-C (3m) = −.03
Applied games focusing on depression													
Fleming et al. (2012)	SPARX	0.88				✓		CDRS-R = −1.61°•∞	CDRS-R = °•	CDRS-R = °•	CDRS-R = °•	CDRS-R (remission) =.79**, CDRS-R = **•	–
Merry et al. (2012)	SPARX	0.92					✓	CDRS-R = −.82°•	CDRS-R (3m) = −1.38°	CDRS-R = −.70°	CDRS-R (3m) = −1.26°	CDRS-R (remission) =.21, CDRS-R =.11	CDRS-R (3m) =.13, CDRS-R (3m) =.04
Lucassen et al. (2015)	Rainbow SPARX	0.91					✓	CDRS-R = −.72*•	CDRS-R = −.72°	–	–	–	–

(continued)

Table 5 (continued)

Study	Name of Intervention	Quality Rating Quant/Qual	Comparator Computerized	Comparator Face-to-face	Comparator W/L	Comparator None	Effect size Within-group Platform Post-treatment	Within-group Platform FU	Within-group Comparator Post-treatment	Within-group Comparator FU	Between group Post-treatment	Between group FU
Stasiak et al. (2014)	The Journey	0.92	✓				CDRS-R = −2.57°•	CDRS-R = −2.17°	CDRS-R = −.78°	CDRS-R = −1.02°	CDRS-R (remission) = 1.18 CDRS-R = .53**	CDRS-R (remission) = .59 CDRS-R = .18°°
Applied games focusing on anxiety and depression												
Knox et al. (2011)	The Journey to the Wild Divine; Freeze Framer	0.81		✓			MASC = −.66° CDI = −.75°	-	MASC = −.25° CDI = −.05°	-	MASC = 1.31** CDI = .44**	-
Virtual Reality focusing on phobia, social anxiety, and trauma												
Maskey, Rodgers, et al. (2019)	Blue Room	0.86			✓		-^	FSSC-R-P (6m) = −.31	-	FSSC-R-P (6m) = −.27	-	FSSC-R-P (6m) = −.50
Maskey et al. (2014)	Blue Room	0.79				✓	-^	SCAS-C (6w/6m/12–16m) = −.42/ −.78/−.87 / SCAS-P (6w/6m/12–16 m) = −.40/ −.60/−.63	-	-	-	-

W/L = Waiting list; FU = Follow-up' 6w = 6 weeks; 3m = 3 months; 6m = 6 months; 12–16m = 12–16 months; 12m = 12 months; SCAS-C = Spence Children's Anxiety Scale Child report; SCAS-P = Spence Children's Anxiety Scale Parent report; ADIS-IV-P = Anxiety Disorders Interview Schedule for DSM-IV: Parent version; FSSC-R = Fear Survey Schedule for Children-Revised; MASC = Multidimensional Anxiety Scale for Children; CDRS-R = Children's Depression Scale – Revised; Questionnaires marked in bold = the study defined this as their primary outcome measure. Symbols: * = significant within-group difference; ** = significant between group differences; ° = statistical analysis examining within-group differences were not reported; °° = statistical analysis examining between group differences were not reported; ' = effect sizes not reported in paper and no information provided to calculate them; ^ = completers only analyses; • = main outcome measure as defined in study; ∞ = effect size as reported in paper

non-therapeutic (e.g., commercial) games failed to identify significant differences (e.g., Wijnhoven et al., 2020).

For VR, the limited literature on children and young people's experience suggests that they adhere to VR interventions and the VR environments evoke feelings of anxiety. From the three studies that we identified, only two reported symptom outcomes showing small to medium within-group effects for changes in fears and anxiety, but, VR had no statistically significant advantage over waiting list, albeit in a small trial (Maskey, Rodgers, et al., 2019).

Limitations of the current literature

In addition to the general lack of studies to examine the experience and effectiveness of applied games or VR to treat mental health problems in children and young people, interpretation of the existing evidence-base needs to take into account several important limitations, specifically the lack of concept clarity, the wide variation in intervention approaches, a failure to take in to account potential developmental differences in terms of what works for whom, a reliance on self- and/or parent report, a lack of consistency in methods used, and an absence of or lack of reporting on the co-design process (i.e., the active involvement of stakeholders in the development of the technology). Each of these limitations and associated implications for future research are now discussed in turn.

Lack of concept clarity. An important source of variation across studies is inconsistency in how 'applied games' and 'VR' have been defined. Applied games comprise both 'serious games' and 'gamification', which, as highlighted by Fleming et al. (2017), have both been defined in various ways in the literature. There is also wide variation in terms of the type of games that are used to deliver interventions (from coloring tasks to engaging with characters in fantasy game world environment). Similarly, the term 'VR' is often applied to rather different hardware and seldom elaborated on in reports which can make it unclear what exactly is being delivered. Furthermore, the term VR is sometimes applied to non-interactive and non-immersive technologies. For example, we excluded several studies that described using virtual reality (e.g., Dewis et al., 2001; Falconer, Davies, Grist, & Stallard, 2019; Gutierrez-Maldonado, Magallon-Neri, Rus-Calafell, & Penaloza-Salazar, 2009; Maskey, McConachie, et al., 2019; St-Jacques, Bouchard, & Bélanger, 2010) however the intervention did not rely on VR hardware and was, for example, delivered via a two-dimensional computer screen with, therefore, limited immersive capability.

Variable intervention approaches. There is also wide variation in the treatment mechanisms that

have been targeted by game mechanics – that is, the 'vehicles' by which therapeutic change is delivered, particularly for applied games that target anxiety problems. Notably, exposure is considered a key treatment ingredient in CBT for anxiety problems in children and young people (Kendall et al., 2006; Peris et al., 2017) and recent research has highlighted the limited (and potential detrimental) impact of relaxation exercises (Peris et al., 2015; Whiteside et al., 2020), yet the majority of applied games for anxiety problems focused only on training relaxation, and only one (i.e., MindLight; Wijnhoven et al., 2020) included exposure. In contrast, the two VR interventions applied exposure as their main and only treatment component. When it comes to applied games targeting depression, the treatment content in SPARX, particularly, was more extensive and aligned with the mechanisms typically targeted in face-to-face interventions (i.e., included psychoeducation, relaxation, interpersonal skills, activity scheduling, problem-solving, cognitive restructuring, distress tolerance, and relapse prevention). But still here it remains unclear to what extent the core mechanisms are effectively changed by the intervention and the extent to which that relates to treatment outcomes.

Lack of consideration of developmental factors. Existing applied games and VR interventions for children and young people have been pioneering but have so far failed to take into account possible developmental differences in the presentation of and what is likely to maintain mental health problems in children and young people, particularly among the studies targeting anxiety problems. For example, with a few notable exceptions (e.g., Carlier et al., 2019; Fleming et al., 2012, 2016; Wols et al., 2018), studies have typically included children and young people from broad age ranges (e.g., 9–17 years; Knox et al., 2011), despite there being developmental differences in both the clinical characteristics (e.g., baseline severity and comorbid psychopathology; Kendall et al., 2010; Waite & Creswell, 2014) and maintenance mechanisms (e.g., role of threat and safety cues; Waters, Theresiana, Neumann, & Craske, 2017; attribution and interpretation biases; Creswell, Murray, & Cooper, 2014) of mental health problems from childhood to adolescence. To date, too little attention has been given to identifying evidence-based and developmentally appropriate treatment components to inform the development of applied games.

Limited outcome measurements. The failure to take into account developmental differences also has implications for outcome measurement. The papers included in this review largely relied on self- and/or parent report measures to assess outcomes, with only two studies (Maskey, Rodgers, et al., 2019; Wijnhoven et al., 2020) including gold-standard clinician assessments. While questionnaire

measures bring advantages in terms of time and costs, the appropriateness of relying on different reporters is likely to vary for children and young people at different ages. For example, child report questionnaires were commonly used to identify research participants with elevated anxiety symptoms and/or measure the effectiveness of applied games; however, the specificity and sensitivity of child self-report questionnaires are low among preadolescent children (Evans, Thirlwall, Cooper, & Creswell, 2016; Reardon, Creswell, et al., 2019), leading to recent recommendations to prioritize parent/carers report for younger children (Creswell et al., 2020). Wider issues with outcome measurement included the common inclusion of a range of different questionnaires (including unstandardized ones) at variable time-points without defining the primary outcome, leading to a greater risk of overemphasizing one, potentially spurious, significant result.

Notably, given the potential for applied games and VR to increase the efficiency of treatment, there was a lack of consideration of health economic outcomes across studies. Attention should also be given in clinical outcome studies to the potential occurrence of adverse effects. Commercial VR head-mounted displays are often not recommended for children, principally it seems due to caution about having screens close to the eyes, although this is less of a concern for the limited time spent in therapeutic treatments. Furthermore, the role of parents in successful implementation of applied games and VR is unclear. It will be essential that these issues are addressed going forward if we want to have a sufficiently robust evidence-base for applied games and VR to consider integrating these approaches in practice.

Variability in methods. In addition to the variability in intervention approaches and assessments, included studies also varied extensively in key methodological characteristics. Four studies used a computerized control group condition; three used a waitlist control, ten had no comparison group, and only one program (i.e., SPARX) has been evaluated within a real-world setting. The index intervention was compared to a face-to-face intervention in two studies, but, again, there was variability in the nature of the face-to-face interventions which included, for example, a shortened version of school-based group CBT (Schoneveld et al., 2018) and treatment as usual (mainly school-based counseling; Merry et al., 2012). Only two trials were set up as non-inferiority trials – one finding evidence of non-inferiority to group CBT (Schoneveld et al., 2018) and the other finding non-inferiority in comparison to counseling (Merry et al., 2012). It is important to note that the extent of change in both conditions in the former trial was small, so we cannot feel confident concluding that either the

game or the group CBT was particularly effective. Given the wide range of effect sizes and comparison conditions found in existing studies, this research needs to be underpinned by a priori standards for the necessary level of evidence required in order to claim that offering applied games/VR interventions will make a clinically important contribution to the settings in which they are delivered.

It was encouraging to see that researchers have started to examine the experience of using applied games and VR via qualitative methods to investigate issues related to acceptability and satisfaction; however, future research would benefit from using a rigorous qualitative methodology in terms of the method of selecting participants, methods of data collection and analysis (providing a conceptual and critical analysis of the data), and use of reflexivity (Braun & Clarke, 2013).

Limited co-design. It is a considerable limitation of the studies reviewed that little reference is made to whether the interventions presented were co-designed, that is, actively involved key stakeholders (e.g., service users, clinicians, service providers, and researchers) in the design process to ensure interventions meet their needs and are engaging and usable. Recent guidance for digital mental health innovations emphasize the importance of co-design (Bevan Jones et al., 2020; Hill et al., 2018; Richards et al., 2016) due to the benefits it brings in terms of: (a) design quality (e.g., Yardley, Morrison, Bradbury, & Muller, 2015), (b) adherence (e.g., Howe, Batchelor, Coates, & Cashman, 2014), (c) usability (e.g., Maguire, 2001), and (d) stakeholder acceptance and adoption (e.g., Wölbling et al., 2012). Experiencing applied games/VR as effective and enjoyable is key for ensuring adherence and ultimately successful dissemination (Read & Shortell, 2011). Only the papers reporting on SPARX and Mindlight make any reference to the involvement of young people in the design of the games, although details are limited and so it is unclear to what extent a co-design process was undertaken. Interestingly, the premise of one Mindlight paper (Schoneveld et al., 2019) was to gain feedback from children in order to inform the redevelopment of the game due to issues with acceptability (Schoneveld et al., 2018). This highlights the importance of involving the intended users in the design and development process *from the start*. One paper (Carlier et al. (2019) included some information about clinicians having input to inform how to make the game appropriate for children with ASD, but again details on the process were limited. We would strongly recommend that future game/VR innovations for mental health are not only co-designed, but that the development process is published in order to allow transparency. This principle extends to the adaptation of games for different contexts, where we would recommend

adapting the original game for the new intended user group through a co-design process, as was done for the adaptation of SPARX for 'sexual minority' young people (Rainbow SPARX; Lucassen et al., 2015).

Strengths and limitations

This systematic review has several strengths, including its consideration of both children and young people's experience of and the effectiveness of using applied games or VR for mental health problems and quantification of the size of the effect. Furthermore, the systematic nature of the review ensured a rigorous approach, and the use of a quality assessment tool enhanced the critical evaluation of the findings. Nevertheless, a number of limitations must be considered. First, our effect size calculations may have been overinflated as we assumed statistical independence between pre- and post-intervention/follow-up scores. Second, conclusions cannot be drawn about the effectiveness of applied games or VR on more discrete aspects/symptoms of a condition (e.g., social skills deficits and attentional factors) as we focused on treatment studies targeting and measuring mental health problems. Third, we made the decision to only include studies where the applied game or VR was considered the active part of the treatment. This meant that, for example, studies that described using a game as a part of computerized CBT (Khanna & Kendall, 2010) were not included. Fourth, although we used a standardized quality rating assessment tool, the quality ratings should be interpreted with some caution as many of the studies were case studies and so several of the quality rating criteria were not relevant and were therefore excluded from the summary score calculations in line with Kmet et al. (2004). Subsequently, although the average quality ratings suggest that the papers reviewed are of reasonable to good quality, this should not be interpreted as indicating a rigorous methodological quality per se but rather that the studies are of a reasonable quality for the type of study that they are. Finally, the limited reporting in many studies regarding the applied game or VR elements that were applied means we may have missed studies altogether, further limiting the generalizability of the findings.

Conclusions

The potential for applied games and VR interventions to effectively treat mental health disorders in CYP makes them an appealing avenue for development and implementation. However, despite enthusiasm for these technologies, this review highlights the need for further robust (developmentally informed) theory and user-driven design, and evidence of acceptability and clinical- and cost-effectiveness before they can be made widely available as treatments for children and young people with mental health problems. Going forwards, the field must also demonstrate the ability to scale and implement effective applied games and VR within or alongside clinical service provision.

Correspondence

Cathy Creswell, Department of Experimental Psychology, University of Oxford, The Observatory Quarter, Oxford OX2 6GG, UK; Email: cathy.creswell@psych.ox.ac.uk

---

Key points

- Research examining children and young people's experience of and the effectiveness of using applied games or VR for mental health problems is in its infancy.
- Although children and young people enjoy using applied games, overall, their effectiveness in targeting mental health problems appears limited, except for depressive symptoms.
- Very little research attention has been given to VR interventions for mental health problems in children and young people making it difficult to draw conclusions about their effectiveness.
- Greater consideration of developmental factors, inclusion of evidence-based treatment components, and involvement of children and young people in the development of applied games and VR is required.

References

Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., & Hedman, E. (2014). Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: A systematic review and meta-analysis. *World Psychiatry*, *13*, 288–295.

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2016). Mental health smartphone apps: Review and evidence-based recommendations for future developments. *JMIR Mental Health*, *3*, e7.

Bevan Jones, R., Stallard, P., Agha, S.S., Rice, S., Werner-Seidler, A., Stasiak, K., … & Merry, S. (2020). Practitioner review: Co-design of digital mental health technologies with children and young people. *Journal of Child Psychology and Psychiatry*, *61*, 928–940.

Bobier, C., Stasiak, K., Mountford, H., Merry, S., & Moor, S. (2013). When 'e'therapy enters the hospital: Examination of the feasibility and acceptability of SPARX (a cCBT programme) in an adolescent inpatient unit. *Advances in Mental Health*, *11*, 286–292.

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. London: Sage.

Carlier, S., Van der Paelt, S., Ongenae, F., De Backere, F., De Turck, F., & Acm. (2019). Using a serious game to reduce stress and anxiety in children with autism spectrum disorder. In *Proceedings of the 13th Eai International Conference on Pervasive Computing Technologies for Healthcare* (pp. 452–461).

Carrasco, A.E. (2016). Acceptability of an adventure video game in the treatment of female adolescents with symptoms of depression. *Research in Psychotherapy: Psychopathology, Process and Outcome*, *19*, 10–18.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edn). New York: Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Copeland, W.E., Angold, A., Shanahan, L., & Costello, E.J. (2014). Longitudinal patterns of anxiety from childhood to adulthood: The Great Smoky Mountains study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*, 21–33.

Coyle, D., McGlade, N., Doherty, G., O'Reilly, G., & Acm. (2011). Exploratory evaluations of a computer game supporting cognitive behavioural therapy for adolescents. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2937–2946). https://doi.org/10.1145/1978942.1979378

Creswell, C., Murray, L., & Cooper, P. (2014). Interpretation and expectation in childhood anxiety disorders: age effects and social specificity. *Journal of Abnormal Child Psychology*, *42*, 453–465.

Creswell, C., Nauta, M.H., Hudson, J.L., March, S., Reardon, T., Arendt, K., … Kendall, P.C. (2020). Research Review: Recommendations for reporting on treatment trials for child and adolescent anxiety disorders–An international consensus statement. *Journal of Child Psychology and Psychiatry*. https://doi.org/10.1111/jcpp.13283

Dewis, L.M., Kirkby, K.C., Martin, F., Daniels, B.A., Gilroy, L.J., & Menzies, R.G. (2001). Computer-aided vicarious exposure versus live graded exposure for spider phobia in children. *Journal of Behavior Therapy and Experimental Psychiatry*, *32*, 17–27.

Evans, R., Thirlwall, K., Cooper, P., & Creswell, C. (2016). Using symptom and interference questionnaires to identify recovery among children with anxiety disorders. *Psychological Assessment*, *29*(7), 835–843. http://dx.doi.org/10.1037/pas0000375

Falconer, C.J., Davies, E.B., Grist, R., & Stallard, P. (2019). Innovations in practice: Avatar-based virtual reality in CAMHS talking therapy: Two exploratory case studies. *Child and Adolescent Mental Health*, *24*, 283–287.

Farrell, L.J., Miyamoto, T., Donovan, C.L., Waters, A.M., Krisch, K.A., & Ollendick, T.H. (2020). Virtual reality one-session treatment of child-specific phobia of dogs: A controlled, multiple baseline case series. *Behavior Therapy*. (in press). https://doi.org/10.1016/j.beth.2020.06.003

Fineberg, N.A., Haddad, P.M., Carpenter, L., Gannon, B., Sharpe, R., Young, A.H., … & Sahakian, B.J. (2013). The size, burden and cost of disorders of the brain in the UK. *Journal of Psychopharmacology*, *27*, 761–770.

Fleming, T.M., Bavin, L., Stasiak, K., Hermansson-Webb, E., Merry, S.N., Cheek, C., … & Hetrick, S. (2017). Serious games and gamification for mental health: current status and promising directions. *Frontiers in Psychiatry*, *7*, 215.

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*, 529–541.

Fleming, T., Lucassen, M., Stasiak, K., Shepherd, M., & Merry, S. (2016). The impact and utility of computerised therapy for educationally alienated teenagers: The views of adolescents who participated in an alternative education-based trial. *Clinical Psychologist*, *20*, 94–102.

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., … & Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*, 625–632.

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, *47*, 2393–2400.

Goodyer, I.M., Reynolds, S., Barrett, B., Byford, S., Dubicka, B., Hill, J., … & Fonagy, P. (2017). Cognitive behavioural therapy and short-term psychoanalytical psychotherapy versus a brief psychosocial intervention in adolescents with unipolar major depressive disorder (IMPACT): A multicentre, pragmatic, observer-blind, randomised controlled superiority trial. *The Lancet Psychiatry*, *4*, 109–119.

Green, H., McGinnity, Á., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of children and young people in Great Britain, 2004*. Basingstoke, UK: Palgrave Macmillan. http://no-pa.uk/wp-content/uploads/2015/02/Mental-health-of-children.pdf

Gutierrez-Maldonado, J., Magallon-Neri, E., Rus-Calafell, M., & Penaloza-Salazar, C. (2009). Virtual reality exposure therapy for school phobia. *Anuario de Psicologia*, *40*, 223–236.

Hill, C., Creswell, C., Vigerland, S., Nauta, M.H., March, S., Donovan, C., … & Kendall, P.C. (2018). Navigating the development and dissemination of internet cognitive behavioral therapy (iCBT) for anxiety disorders in children and young people: A consensus statement with recommendations from the# iCBTLorentz Workshop Group. *Internet Interventions*, *12*, 1–10.

Hollis, C., Falconer, C.J., Martin, J.L., Whittington, C., Stockton, S., Glazebrook, C., & Davies, E.B. (2017). Annual Research Review: Digital health interventions for children and young people with mental health problems–A systematic and meta-review. *Journal of Child Psychology and Psychiatry*, *58*, 474–503.

Howe, D., Batchelor, S., Coates, D., & Cashman, E. (2014). Nine key principles to guide youth mental health: Development of service models in New South Wales. *Early Intervention in Psychiatry*, *8*, 190–197.

James, A.C., James, G., Cowdrey, F.A., Soler, A., & Choke, A. (2015). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *The Cochrane Library*.

Kendall, P.C., Compton, S.N., Walkup, J.T., Birmaher, B., Albano, A.M., Sherrill, J., ... & Piacentini, J. (2010). Clinical characteristics of anxiety disordered youth. *Journal of Anxiety Disorders, 24*, 360–365.

Kendall, P.C., Robin, J.A., Hedtke, K.A., Suveg, C., Flannery-Schroeder, E., & Gosch, E. (2006). Considering CBT with anxious youth? Think exposures. *Cognitive and Behavioral Practice, 12*, 136–148.

Khanna, M., & Kendall, P.C. (2010). Computer-assisted cognitive behavioral therapy for child anxiety: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 78*, 737–745.

Kmet, L.M., Cook, L.S., & Lee, R.C. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields.* https://doi.org/10.7939/R37M04F16

Knox, M., Lentini, J., Cummings, T., McGrady, A., Whearty, K., & Sancrant, L. (2011). Game-based biofeedback for paediatric anxiety and depression. *Mental Health in Family Medicine, 8*, 195–203.

Lambe, S., Knight, I., Kabir, T., West, J., Patel, R., Lister, R., ... & Freeman, D. (2020). Developing an automated VR cognitive treatment for psychosis: gameChange VR therapy. *Journal of Behavioral and Cognitive Therapy, 30*, 33–40.

Lau, H.M., Smit, J.H., Fleming, T., & 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

Lawrence, D., Hafekost, J., Johnson, S.E., Saw, S., Buckingham, W.J., Sawyer, M.G., ... & Zubrick, S.R. (2016). Key findings from the second Australian child and adolescent survey of mental health and wellbeing. *Australian & New Zealand Journal of Psychiatry, 50*, 876–886.

Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010). *Social media & mobile internet use among teens and young adults. Millennials.* Pew internet & American Life Project.

Lucassen, M.F., Merry, S.N., Hatcher, S., & Frampton, C.M. (2015). Rainbow SPARX: A novel approach to addressing depression in sexual minority youth. *Cognitive and Behavioral Practice, 22*, 203–216.

Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies, 55*, 587–634.

Maskey, M., Lowry, J., Rodgers, J., McConachie, H., & Parr, J.R. (2014). Reducing specific phobia/fear in young people with autism spectrum disorders (ASDs) through a virtual reality environment intervention. *PLoS One, 9*, e100374.

Maskey, M., McConachie, H., Rodgers, J., Grahame, V., Maxwell, J., Tavernor, L., & Parr, J.R. (2019). An intervention for fears and phobias in young people with autism spectrum disorders using flat screen computer-delivered virtual reality and cognitive behaviour therapy. *Research in Autism Spectrum Disorders, 59*, 58–67.

Maskey, M., Rodgers, J., Grahame, V., Glod, M., Honey, E., Kinnear, J., ... & Parr, J.R. (2019). A randomised controlled feasibility trial of immersive virtual reality treatment with cognitive behaviour therapy for specific phobias in young people with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*, 1912–1927.

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: British Medical Journal, 344*, e2598.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*, e1000097.

Parrish, D.E., Oxhandler, H.K., Duron, J.F., Swank, P., & Bordnick, P. (2016). Feasibility of virtual reality environments for adolescent social anxiety disorder. *Research on Social Work Practice, 26*, 825–835.

Peris, T.S., Compton, S.N., Kendall, P.C., Birmaher, B., Sherrill, J., March, J., ... & Piacentini, J. (2015). Trajectories of change in youth anxiety during cognitive—behavior therapy. *Journal of Consulting and Clinical Psychology, 83*, 239.

Peris, T.S., Piacentini, J., Bergman, R.L., Caporino, N.E., O'Rourke, S., Kendall, P.C., ... & Sakolsky, D. (2017). Therapist-reported features of exposure tasks that predict differential treatment outcome for youth with anxiety. *Journal of Consulting and Clinical Psychology, 83*(2), 239–252.

Pompili, M., Serafini, G., Innamorati, M., Dominici, G., Ferracuti, S., Kotzalidis, G.D., ... & Lester, D. (2010). Suicidal behavior and alcohol abuse. *International Journal of Environmental Research and Public Health, 7*, 1392–1431.

Read, J.L., & Shortell, S.M. (2011). Interactive games to promote behavior change in prevention and treatment. *JAMA, 305*, 1704–1705.

Reardon, T., Creswell, C., Lester, K.J., Arendt, K., Blatter-Meunier, J., Bögels, S.M., ... & Eley, T.C. (2019). The utility of the SCAS-C/P to detect specific anxiety disorders among clinically anxious children. *Psychological Assessment, 31*, 1006–1018.

Reardon, T., Harvey, K., & Creswell, C. (2019). Seeking and accessing professional support for child anxiety in a community sample. *European Child & Adolescent Psychiatry, 29*, 649–664.

Richards, D., Vigano, N., O'Callaghan, D.D., O'Brien, E., Mooney, J., & Bonner, C. (2016). Towards a gold standard for internet-delivered programs in behavioral and mental health. *European Psychiatry, 33*, S762.

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*, e0147763.

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? *Prevention Science, 19*, 220–232.

Schoneveld, E.A., Lichtwarck-Aschoff, A., & Granic, I. (2019). What keeps them motivated? Children's views on an applied game for anxiety. *Entertainment Computing, 29*, 69–74.

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, 63*, 321–333.

Stasiak, K., Hatcher, S., Frampton, C., & Merry, S.N. (2014). A pilot double blind randomized placebo controlled trial of a prototype computer-based cognitive behavioural therapy program for adolescents with symptoms of depression. *Behavioural and Cognitive Psychotherapy, 42*, 385–401.

St-Jacques, J., Bouchard, S., & Bélanger, C. (2010). Is virtual reality effective to motivate and raise interest in phobic children toward therapy? A clinical trial study of in vivo with in virtuo versus in vivo only treatment exposure. *The Journal of Clinical Psychiatry, 71*, 924–931.

The Endnote Team. (2013). *Endnote (Version EndNote X9) [64 bit].* Philadelphia: Clarivate.

Vizard, T., Sadler, K., Ford, T., Merad, S., Brodie, E., Forbes, N., & Pearce, N. (2018). *Mental health of children and young people in England, 2017.* Leeds, UK: Health and Social Care Information Centre. https://dera.ioe.ac.uk/32622/11/MHCYP%202017%20Survey%20Design%20and%20Methods.pdf

Waite, P., & Creswell, C. (2014). Children and adolescents referred for treatment of anxiety disorders: Differences in clinical characteristics. *Journal of Affective Disorders, 167*, 326–332.

Waters, A.M., Theresiana, C., Neumann, D.L., & Craske, M.G. (2017). Developmental differences in aversive conditioning, extinction, and reinstatement: A study with children, adolescents, and adults. *Journal of Experimental Child Psychology, 159*, 263–278.

Whiteside, S.P., Sim, L.A., Morrow, A.S., Farah, W.H., Hilliker, D.R., Murad, M.H., & Wang, Z. (2020). A meta-analysis to guide the enhancement of CBT for childhood anxiety: exposure over anxiety management. *Clinical Child and Family Psychology Review, 23*, 102–121.

Wijnhoven, L.A.M.W., Creemers, D.H.M., Vermulst, A.A., Lindauer, R.J.L., Otten, R., Engels, R.C.M.E., & Granic, I. (2020). Effects of the video game 'Mindlight' on anxiety of children with an autism spectrum disorder: A randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry, 68*, 101548.

Wölbling, A., Krämer, K., Buss, C.N., Dribbisch, K., LoBue, P., & Taherivand, A. (2012). Design thinking: An innovative concept for developing user-centered software. In A. Maedche, A. Botzenhardt, & L. Neer (Eds.), *Software for people* (pp. 121–136). Berlin, Heidelberg: Springer.

Wols, A., Lichtwarck-Aschoff, A., Schoneveld, E.A., & Granic, I. (2018). In-game play behaviours during an applied video game for anxiety prevention predict successful intervention outcomes. *Journal of Psychopathology and Behavioral Assessment, 40*, 655–668.

Yardley, L., Morrison, L., Bradbury, K., & Muller, I. (2015). The person-based approach to intervention development: Application to digital health-related behavior change interventions. *Journal of Medical Internet Research, 17*, e30.