Dragons of Afterlands: Evaluating the Effectiveness of an Augmented Reality Board Game on Adolescent Wellbeing

Idil Kilinc

June, 2022

Research submitted in partial fulfilment of the requirements for the degree of Doctor in Clinical Psychology (DClinPsy), Royal Holloway, University of London

Table of Contents

List of Tables

List of Figures

Part I:

Lay Summary

Wellbeing is a broad construct that includes having a good quality of life, being happy, thriving in our day to day lives and having good relationships with others. Wellbeing in adolescence is important because it can have lasting impact on the rest of our lives. Previous studies have found that wellbeing influences many areas of our lives including our mental health, physical health, school success and employment. Therefore, it is important to support adolescents and help them develop themselves and learn skills so that they can take care of their mental health.

Projects to support adolescents and teach them wellbeing skills exist and these often take place in schools. However, these can sometimes be difficult for staff to teach, as these topics may be unfamiliar to them. Some students who may be more at risk of having poor wellbeing or having mental health problems may also struggle to learn through typical classroom-based teaching.

As children nowadays grow up with technology, they may find it more fun and interesting to learn through technology, and may actually learn better this way as they pay more attention. One of the new technologies that has been developed is extended reality, which includes virtual reality and augmented reality.

- Virtual reality: Users wear headsets that show them an entirely computer generated world, that they can interact with. (e.g. Oculus Rift)
- Augmented reality: Users use smartphones to see computer generated elements in their present environment. (e.g. Pokemon Go)

Previous researchers have used these technologies to improve teaching in schools and have found that students learn well this way. Researchers have also used these technologies in

therapy and have found that they can help treat mental health problems, like depression or anxiety. In my study, I wanted to test whether these technologies can be helpful to improve wellbeing knowledge and skills in adolescents.

My study had two parts. Firstly, I searched the literature to see if any other researchers had used extended reality technologies to improve wellbeing in adolescents, to understand how they used it and whether these attempts were successful in improving wellbeing. My search showed that there were only ten studies looking at this topic. Although more than half of these studies showed that participating in these extended reality projects can be good for wellbeing, it was difficult to be certain about these results due to research limitations. Most studies got participants to play regular games or watch videos and did not focus specifically on teaching wellbeing skills. Each study was very different from another and therefore, it was not possible to say whether we would find the same results if we were to apply the same project elsewhere with different adolescents. Also, most studies only measured wellbeing immediately after the project and did not look at what happens to wellbeing in the long term, which would be more helpful. There were only two studies that used augmented reality, meaning that there was not enough research on this topic to comment on its use with adolescents and benefits to wellbeing.

For the second part of my study, I tested a newly developed augmented reality project (called Dragons of Afterlands) in schools to see if it is successful in improving wellbeing. Dragons of Afterlands is a board game, using augmented reality, that adolescents play together in class. While playing the game, adolescents complete challenges that are designed to teach them wellbeing skills, such as recognizing emotions or handling conflict with others. Adolescents played the game once a week for five to six weeks and I assessed their wellbeing using questionnaires at the beginning and end of the study and one month later. I then compared these

wellbeing scores to those of adolescents who did not play the game to see if there were any differences.

The results showed that there were no differences in wellbeing of adolescents who played the game and those who didn't, showing that the game was not successful in improving wellbeing in this study. I also asked adolescents to complete another questionnaire on how motivated they were to attend these game sessions and found that those who played the game were less motivated than those who attended regular classes at school, which was unexpected. Some possible explanations for these findings could be that there may have been a mismatch between what adolescents already knew about wellbeing and what the game was aiming to teach (the content may have been too easy), the features of the game may not have been as engaging (e.g. there were technical glitches), the adolescents may have had other priorities (e.g. schoolwork).

Although the current version of the game was not successful, the learning points from this study will be helpful in guiding future versions of the game, as well as the field of extended reality wellbeing research overall. Some learning points and future directions include working more collaboratively with adolescents and schools to better meet their needs, adjusting the game content to skills and knowledge of players, and the advantages and challenges of testing extended reality projects early on in schools.

Part II: The Systematic Review

The Use of Extended Reality Interventions to Improve Wellbeing in Adolescents and Young People

Abstract

Wellbeing plays an important role in mental health and quality of life for individuals, particularly during developmental periods like adolescence. Many interventions have been developed to promote wellbeing in adolescents, with the literature shifting towards interventions that benefit from technological enhancements. The aim of the present study was to review effective extended reality (XR) (immersive technologies, including virtual reality and augmented reality) interventions that have been used to improve wellbeing in adolescents. American Psychological Association (APA) PsycINFO, Institute of Electrical and Electronics Engineers (IEEE) Xplore and Web of Science were searched for articles using XR interventions with individuals between the ages of 10-24 and assessing pre-post changes in wellbeing outcomes. Findings were evaluated based on participant characteristics, study characteristics, and study outcomes and synthesized narratively. Risk of bias was assessed using the Cochrane Risk of Bias Tool and of the ten papers assessed, three were found to have low risk of bias, three had some risk of bias and four had high risk of bias. The results of the review were mixed with the majority of studies (60%) reporting at least some improvements in wellbeing post intervention. The review highlighted gaps in the literature, such as the lack of longitudinal interventions, lack of studies with younger populations, lack of AR studies which indicated that further research is needed for conclusive and generalizable outcomes.

Introduction

Mental wellbeing is conceptualized as more than just the lack of mental illness and focuses on individuals feeling good and functioning well in their lives (World Health Organisation [WHO], 2004b; Department of Health, 2014). Wellbeing has been a topic of interest in psychology for decades and multiple definitions have been developed for it over time, with each definition slightly altering our understanding of the construct.

A comprehensive definition of wellbeing is one developed by Ryan and Deci (2000), which defines wellbeing as a construct that comprises of both hedonic and eudaimonic aspects. Hedonic or subjective wellbeing refers to the influence of positive and negative emotions on individuals' mental state and focuses on happiness, satisfaction and pleasure attainment; whereas eudaimonic or psychological wellbeing refers to positive functioning in daily life, positive relationships with others, meaning-making and self-realization (Keyes et al., 2002; Ryan & Deci, 2001; Bolier et al., 2013; Diener, 1984). Optimal wellbeing, then, indicates high levels of hedonic and eudaimonic wellbeing (Keyes et al., 2002) and includes positive affect and cognitions towards life, a global sense of satisfaction, positive mental health and a good quality of life (Ben-Zur, 2003).

Wellbeing plays an important role in individuals' lives as it is linked to physical health and longevity (Diener et al., 2017), employment, social relationships and mental health (Department of Health, 2014). Wellbeing in childhood and adolescence is particularly important, as wellbeing in early years can have lasting impact into adulthood (Department of Health, 2013c). However, wellbeing often declines during adolescence, reaching its lowest point at around age 14-15 (Department of Health, 2013b). According to the World Health Organization, an important component of prevention of mental health problems is the promotion of

individuals' strengths and skillset to reduce their susceptibility to future mental health difficulties (WHO, 2004a). Promotion of wellbeing and mental health is most effective when it takes place in youth (WHO, 2013), highlighting the importance of developing interventions for this particular age group.

Wellbeing Interventions for Children and Adolescents

Considering the need for mental health and wellbeing support in adolescence, several interventions have been developed to target these issues. Wellbeing interventions for children and adolescents have often targeted skills development in relation to mental health disorders and have been reactive to problems in school rather than taking a preventative approach (Vostanis et al., 2013). One meta-analysis, taking a symptom focused approach, found that resilience focused interventions, building skills and resources based on CBT principles for depressive symptoms, anxiety symptoms and general psychological distress have shown the most promise in improving mental health outcomes (which were assessed based on internalizing and externalizing symptoms) (Dray et al., 2017). Mindfulness based interventions and interventions using Acceptance and Commitment Therapy (ACT) principles were also shown to lead to positive mental health outcomes (van Agteren et al., 2021).

More recently, there has been a shift towards positive psychology interventions that focus on building on strengths and capacities as preventative factors rather than solely taking a symptom-focused approach. A range of positive psychology content was used in previous interventions, with most of these interventions focusing on social and emotional learning (Kuosmanen et al., 2019). One systematic review found promising results for life skills training interventions targeting social and emotional competences, attitudes about self, others and about going to school (Sancassiani et al., 2015). An example of such a positive psychology

intervention targeting these areas was a large scale study where students participated in activities and discussions, read stories and watched videos on themes, such as gratitude, optimism, self-concordant goal setting and developing flexible mindsets, which led to increases in self-esteem, self-efficacy, optimism and interpersonal resilience (Shoshani & Steinmetz, 2013). Although reviews on this topic have generally found positive effects overall for these interventions, challenges exist in the literature around low quality of evidence (van Agteren et al., 2021), heterogeneity in characteristics of interventions, recruitment of student populations across developmental stages, use of outcome measures preventing direct comparisons (Sancassiani et al., 2015) and lack of long term interventions (Adi et al., 2007).

The majority of existing wellbeing interventions offered to adolescents are based in schools, with digital or community based interventions lacking (Kuosmanen et al., 2019). Interventions commonly target adolescents directly (rather than taking a systemic approach around training for school staff or parents) (Vostanis et al., 2013). Whilst most of the available interventions have utilized face to face delivery methods, digital interventions taking a school-based positive psychology approach may be more equitable and accessible for students to help them learn about wellbeing (Francis et al., 2021). Administering face to face wellbeing interventions in schools can be challenging, as adequate staff training and supervision needs to be provided to ensure fidelity in the delivery of interventions (Gee et al., 2021). Most interventions in the literature are not evidence based and open to adaptations by staff depending on changing circumstances in schools (Vostanis et al., 2013), which makes it difficult to evaluate them and build an evidence base. Technologically driven interventions have the advantage of being more standardized, as they often do not require as much staff input as face to face teaching based interventions. Online computer and smartphone based positive psychology interventions

for adolescents, focusing on flourishing of individuals, have been shown to decrease anxiety and depression and increase wellbeing, however there is a need for more technologically enhanced, controlled, longitudinal studies that are developed specifically for target populations and settings (Baños et al., 2017).

Extended Reality Technology

One rapidly developing area of technological innovation is the field of extended reality (XR). Extended reality (XR) encompasses augmented reality (AR), virtual reality (VR) and mixed reality (MR) technologies. As described by Ventura et al. (2018), VR creates a fully immersive virtual environment, where users feel as if they are inside the virtual world and experiencing it in first person point of view, usually through a headset, such as the Oculus Rift, that provide interactive three dimensional visuals and audio. AR, on the other hand, enhances user experiences by building on the existing real world environment in which the users find themselves, by adding computer generated virtual elements to it. A prime example of AR is the mobile app game, Pokémon Go, where users continue to experience real world (e.g. walking in the park) with the addition of superimposed app generated three dimensional visuals (Pokémons) that appear as if they exist in the real world. Mixed reality is a developing field that combines both AR and VR elements, such as the Microsoft HoloLens; however there is very limited research on mixed reality at present. As the field of extended reality rapidly grows, new definitions and categorizations emerge.

XR is a rapidly growing field with applications across many industries (Parekh et al., 2020). The advantage of XR technologies, such as VR, over non-XR technologies is the immersion and sense of flow (Csikszentmihalyi & Csikzentmihaly, 1990) it can create for its users (Pallavicini & Pepe, 2019). This immersion and sense of flow allows for higher levels of

concentration in tasks and absorption in the experience (Michailidis et al., 2018), which can increase intrinsic motivation for the activity (Csikszentmihalyi & Csikzentmihaly, 1990). XR can have use in mental health through the imagery it creates, which can capture attention and amplify emotional effects of interventions.

Use of XR with Children and Adolescents

XR technology has been used to deliver several mental health interventions across populations, settings and conditions since its development. With adult populations, therapeutic techniques have been successfully applied to VR and these interventions have been found to have clinically significant effects in the treatment of mental health problems such as stress, phobias, anxiety, eating disorders, depression and PTSD and have been used to aid neurocognitive assessments, pain management interventions and rehabilitation (Carl et al., 2019; Kothgassner et al., 2019; Parsons et al., 2017; Spiegel et al., 2019; Tong et al., 2015; Opris et al., 2012). Research on efficacy of these interventions for child and adolescent populations; particularly using controlled trials, are lacking compared to adult literature (Kothgassner & Felnhofer, 2021); however, there is a growing body of evidence to suggest that similar positive effects can be observed in younger people around the treatment of anxiety, depression, phobias, interventions to alleviate pain (Eijlers et al., 2019; Halldorsson et al., 2021; Bouchard, 2011). XR interventions have also been used in hospital settings and were deemed to be safe, engaging and accessible for reducing pain and anxiety of adolescent inpatients (Ridout et al., 2021). XR interventions have been developed to teach social skills to children with autism by helping them visualize concepts (Lorenzo et al., 2019; Chung & Chen., 2017) and have been found to be effective in improving social interactions, communication skills, attention skills and functional skills in children and adolescents with autism (Berenguer et al., 2020). AR and VR technology

have also been used within the exergaming framework of promoting exercise and improving physical health in adolescents and was found to be appealing and acceptable (Farič et al., 2021) and lead to increases in physical activity (Ni et al., 2019; Benzing & Schmidt, 2018).

Use of XR in the Context of Wellbeing Promotion

There is emerging evidence with non-clinical populations that digital technologies can have a positive impact on positive mental health and wellbeing overall (Collins et al., 2019). XR has mostly been used in the context of treatment of mental health disorders; however, there is a growing body of evidence around its use in adult and older adult populations in relation to wellbeing (Carroll et al., 2021; Montana et al., 2020).

Use of VR to promote wellbeing has been a focus on investigation within the workplace context. Studies have found that delivering VR interventions at work improves affect (increased happiness, decreased anxiety, anger and sadness), promotes relaxation and reduces stress overall using a range of intervention contents (Naylor et al., 2019; Adhyaru & Kemp, 2022; Naylor et al., 2020; Riches et al., 2021). These interventions were shown to be acceptable, with participants reporting that there could be significant benefit to making these wellbeing interventions more widely available (Naylor et al., 2019; Riches et al., 2021). VR wellbeing interventions were also delivered to adults to promote the development of their emotion regulation skills, and these interventions were found to promote wellbeing as well (Montana et al., 2020). Interventions in these studies included relaxation techniques, using biofeedback to regulate physiological arousal, behavioral activation and social skills development. The wellbeing outcome measures were taken broadly and included scales for anxiety, mindfulness, depression and coping. Although the diversity in the intervention contents and the outcome measures posed methodological concerns in relation to building a robust evidence base in the

literature, it also demonstrated the flexibility of XR as a delivery tool that is adaptable to various intervention approaches.

The use of XR technology to promote wellbeing in clinical and non-clinical older adults also found some promising results for psychological and social wellbeing, particularly around environmental mastery, social interactions (Lee et al., 2019), mood and apathy (D'Cunha et al., 2019). Similar methodological issues as the adult studies were present in the literature regarding heterogeneity in range of interventions and designs, small sample sizes and technological challenges (such as difficulties using technologies independently, system errors and these leading to frustration).

Existing literature on the use of XR interventions in the context of wellbeing with adult populations shows that XR interventions can be beneficial as wellbeing interventions. It is important to note that research on XR is predominantly on VR technology, with very limited evidence around AR interventions. VR interventions provide a sense of enjoyment and engagement, presence in the moment and an activation of affective motivational states which contribute to wellbeing (Hatta et al., 2022). Evidence also suggests that VR interventions can help increase self-compassion, leading to an increase in self-care behaviors and meditation by facilitating the construction of positive mental images in participants (Cebolla et al., 2019). VR interventions can potentially modify emotional responses through reappraisals of emotional stimuli and give sense of personal growth and autonomy (Montana et al., 2020). XR technology can also enhance the effectiveness of the interventions being delivered by engaging participants to attended more sessions (Hadley et al., 2019; Bosworth, 2016; Marsch & Borodovsky, 2016; Bruijniks et al., 2020). Considering the promising findings on the use of XR in adult literature around wellbeing, and the fact that XR has already been successfully applied within treatment

context with adolescent samples, the application of XR technology to promotion of wellbeing in adolescents holds potential.

Current Review Aims

Although there is a growing body of literature on the use of XR interventions in mental health, to our knowledge, there have been no systematic reviews looking at the use of XR interventions for adolescent wellbeing. The present review aims to outline the types of extended reality interventions that have been offered to adolescents and young people to improve their wellbeing and to evaluate the effectiveness of these interventions. Specifically, the review aims to answer the following question: What are effective XR interventions that have been used to improve wellbeing in adolescents and young people?

Methods

The present study is a systematic review using narrative synthesis. It was conducted in line with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021).

Eligibility Criteria

The eligibility criteria for the present review are outlined using the PICOS framework in Table 1. The search focused on identifying research looking at the impact of XR technologies on the wellbeing of adolescents and young people. The age range for adolescence was selected as 10-24 years of age in line with Sawyer et al., (2018)'s conceptualization of adolescence, which takes into account earlier onset of puberty and delayed role transitions into adulthood (completion of education, marriage, parenthood, financial independence etc.) in modern society, as well as an improved understanding of continued neurological developments into early

twenties. This age range was used in other similar reviews and literature of adolescent mental health (Best et al., 2014; Orben et al., 2020; Morrish et al., 2018; Blakemore, 2019). The search was not limited by publication date as the availability of the technology created a natural limit for this.

Table 1

Eligibility Criteria

	Inclusion Criteria	Exclusion Criteria
Population	Participants should be between the ages of 10 and 24. • Ages 10 and 24 were taken as absolute cut-off points. In cases where mean ages were reported, age range of participants were calculated using standard deviations and samples with participants that fell outside of the 10-24 years range were excluded from the review. Participants can be males and/or females. Participants can be from any ethnicity, country and socioeconomic background. There are no language restrictions, as long as the intervention materials and outcome measures are accessible to the sample population and the findings are reported in English.	Studies that do not report age of participants. Studies that include any participants who fall outside of this age range. Studies that are aimed at participants with learning disabilities.
Intervention	Interventions must use immersive virtual reality, augmented reality, 3D, visuohaptic or extended reality technologies. Interventions can be delivered in group or individual formats.	Interventions only focusing on remote healthcare provision or health education. Interventions that only use serious gaming, gamification or apps without immersive extended reality qualities (e.g. computer based video games, mobile game apps).

	Interventions can use any content aimed at improving mental health and wellbeing.	Papers that do not include interventions tested on participants (e.g. papers describing new tools without delivering the intervention).
Comparators	There are no restrictions on comparators. Studies are not required to have comparison or control groups as long as they meet the outcomes requirements and follow a pre-post intervention design that allows for within groups comparisons.	
Outcomes	Studies must measure at least one wellbeing outcome. Studies must have baseline measures and at least one post intervention measure related to the wellbeing outcome measure (if quantitative). Outcome measures can be qualitative or quantitative. Outcome measures are not required to be standardized, although this will be taken into consideration as a strength or limitation in the interpretation of the findings.	Studies only reporting on mental health symptoms and diagnoses (e.g. anxiety, depression). Studies that only collect post intervention measures of wellbeing. Studies only assessing acceptability or usability of interventions.
Study Design and Types of Study	There are no restrictions on settings. Studies can include samples obtained from the community, schools, hospitals or inpatient settings. There are no restrictions on type of study. Studies can include observational or experimental studies, cross-sectional or longitudinal studies. There are no restrictions around grey literature and theses. These were accessed via the databases (outlined below) and not searched for separately. Studies must be available in English. There are no restrictions on publication date.	Studies where full text is not available in English. Studies where only protocols or abstracts are available, book chapters, reviews (narrative reviews, systematic reviews, literature reviews, meta-analyses), commentaries.

Information Sources

Scoping searches were conducted to determine suitability of databases and to select appropriate search terms. The scoping searches included testing the search terms on various databases to determine their suitability, as well as identifying similar reviews on PROSPERO to inform decision making. The review was registered on PROSPERO (CRD42021261196).

The topic of the present review is interdisciplinary, relating to psychology (mental health and wellbeing, child development, psychological interventions), computer science (extended reality technologies) and potentially education or healthcare depending on the setting of research projects. The following three databases, American Psychological Association (APA) PsycINFO, Institute of Electrical and Electronics Engineers (IEEE) Xplore and Web of Science, were selected to provide a comprehensive cover of the fields of interest. PsycINFO was chosen as a database of publications in psychology, while IEEE Xplore was selected as a database of publications in computer science, electrical engineering/ electronics and allied fields. Web of Science provides access to databases from a range of academic disciplines and was included to capture any papers that may have been published outside of the remit of the other two databases. Other databases like PubMed and Education Resources Information Center (ERIC) were considered; however scoping searches did not yield useful results from these and the decision was made to exclude these databases. The final selection of databases was supported by similar reviews that were previously published in the field (Kavanagh et al., 2017; Halldorsson et al., 2021). Following scoping searches and the selection of databases, the full search was conducted in November 2021.

Search Strategy

The search terms that were used in this review are outlined below. A librarian within the School of Life Sciences at Royal Holloway University of London was consulted to obtain feedback on the search strategy and the selection and application of search terms to relevant databases. Boolean operators were used to combine search terms and truncation was applied and adjusted depending on the algorithms of the databases used. Reference lists of relevant reviews and research papers identified were searched to identify any other articles of relevance.

Search terms:

("augmented reality") OR ("virtual reality") OR (immersive) OR ("alternate reality") OR ("extended reality") OR ("3D") OR (visuohaptic)) AND ((well-being) OR (wellbeing) OR ("mental health")) AND ((teen*) OR ("young people") OR (child*) OR (adolescen*) OR (student*))

Selection Process

Databases were searched and the results were uploaded to Zotero, where duplicates were removed. The remaining papers were screened based on their title and abstracts. A grid of eligibility criteria was created to assist in the decision making process regarding inclusion and exclusion of papers. A second reviewer independently screened 10% of papers for eligibility based on their title and abstracts. There was 92% agreement among the papers screened by both reviewers. Full text articles of papers fitting screening criteria were then obtained and these were reviewed for their eligibility. The second reviewer reviewed 15% of the full text articles and there was 85% agreement at this stage. Disagreements were resolved via discussion among the reviewers and both thesis supervisors were consulted where resolution could not be reached.

Seven of the papers identified at the initial screening stage were research protocols. The authors of these papers were contacted to ascertain whether there were any publications resulting from these studies at the time of the review that may have been missed by the search. Authors were also asked whether they had any unpublished data they would be interested in contributing to the present review. This did not result in any papers or data that could be included in this review and these protocols were removed from the search list.

Data Items and Collection Process

The following information was extracted from the selected papers: 1) author and year of publication, 2) type of study (RCT, case studies etc.), 3) intervention delivery method (headsets, phones etc.), 4) intervention content, 5) duration of intervention, 6) comparators, 7) follow up, 8) wellbeing related outcome measures, 9) demographics (population of study, age, gender, ethnicity), 10) sample size, 11) drop out, 12) location and setting of the study and 13) results. The content of the collected data focused on wellbeing related measures, where applicable, as these were the topic of interest for the present review (e.g. outcome measures, results and follow up information was only extracted and reported in relation to the wellbeing measures if the studies explored multiple domains). Missing or unclear information was extracted and labelled as such and results were interpreted accordingly. Data extraction was conducted independently by the researcher by hand.

Study Risk of Bias Assessment

The original quality assessment strategy devised at the start of the study was the following:

a) The Cochrane Revised Risk Bias Tool (Sterne et al., 2019) was planned to be used for randomized controlled trials, to assess methods of randomization, recruitment of participants, deviations from interventions, outcome measures, missing data and reported results;
b) The Integrated Quality Criteria for the Review of Multiple Study Designs (Zingg et al., 2016) was planned to be used to assess other types of studies (e.g. case studies) for bias in recruitment and allocation, outcome measures and blinding, follow up, analytical rigor and reporting and ethical considerations.

As the result of the study selection only yielded randomized controlled trials, only the Cochrane Revised Risk Bias Tool was used for quality assessment. Papers were rated as Low Concern, Some Concern or High Concern using the appropriate version of the assessment tool based on the study design (individually randomized parallel group trials, cluster randomized trials or crossover trials). Quality assessment was conducted by the main researcher for all papers. 25% of the papers were also assessed by the second reviewer. There was one minor disagreement regarding one of the papers that did not affect the overall rating of the paper. This disagreement was resolved through a discussion and the paper was retained in the review. Judgements on risk of bias were presented in tables and taken into consideration in the analysis.

Effect Measures and Synthesis Methods

Analysis of selected papers included a consideration into the effectiveness of the interventions in improving participant wellbeing, types of technology used, content of interventions, participant characteristics and study design. To achieve this, included studies were grouped based on participant characteristics (population, age, gender, ethnicity, location), study characteristics (study design, intervention content, comparators and delivery method), and study outcomes (outcome measures used and results). The findings were presented in tables and

synthesized narratively. These tables of PICOS factors were also used to assess the heterogeneity of studies included.

Due to heterogeneity in included studies, a similar approach to effect measurements was taken as a recent systematic review by Haldorsson et al. (2021). Similar interventions were grouped and compared directly where possible. In cases where heterogeneity prevented direct comparisons, effect sizes for different interventions on the wellbeing of participants were considered. Within and between group effect sizes were calculated, where possible and meaningful, using the mean scores of the wellbeing outcome measures at baseline and post intervention and the standard deviations. Effect sizes that were reported by authors were extracted directly. Where multiple outcome measures were used, measures most in line with the review aims (i.e. wellbeing related outcome measures) were selected and reported. Frequencies, means and standard deviations were reported, where appropriate and possible depending on the data made available in the selected studies. An online tool was used to assist with effect size calculations (Lenhard & Lenhard, 2016).

A meta-analysis was not conducted due to the high heterogeneity between selected studies (in intervention content, design, outcome measures) and the small number of studies included that were assessed to have low risk of bias. Instead, data was synthesized narratively. Studies that were rated as having lower concerns for quality of bias were considered to hold more weight in certainty of evidence and strengths and limitations of the review were taken into consideration when drawing conclusions.

Results

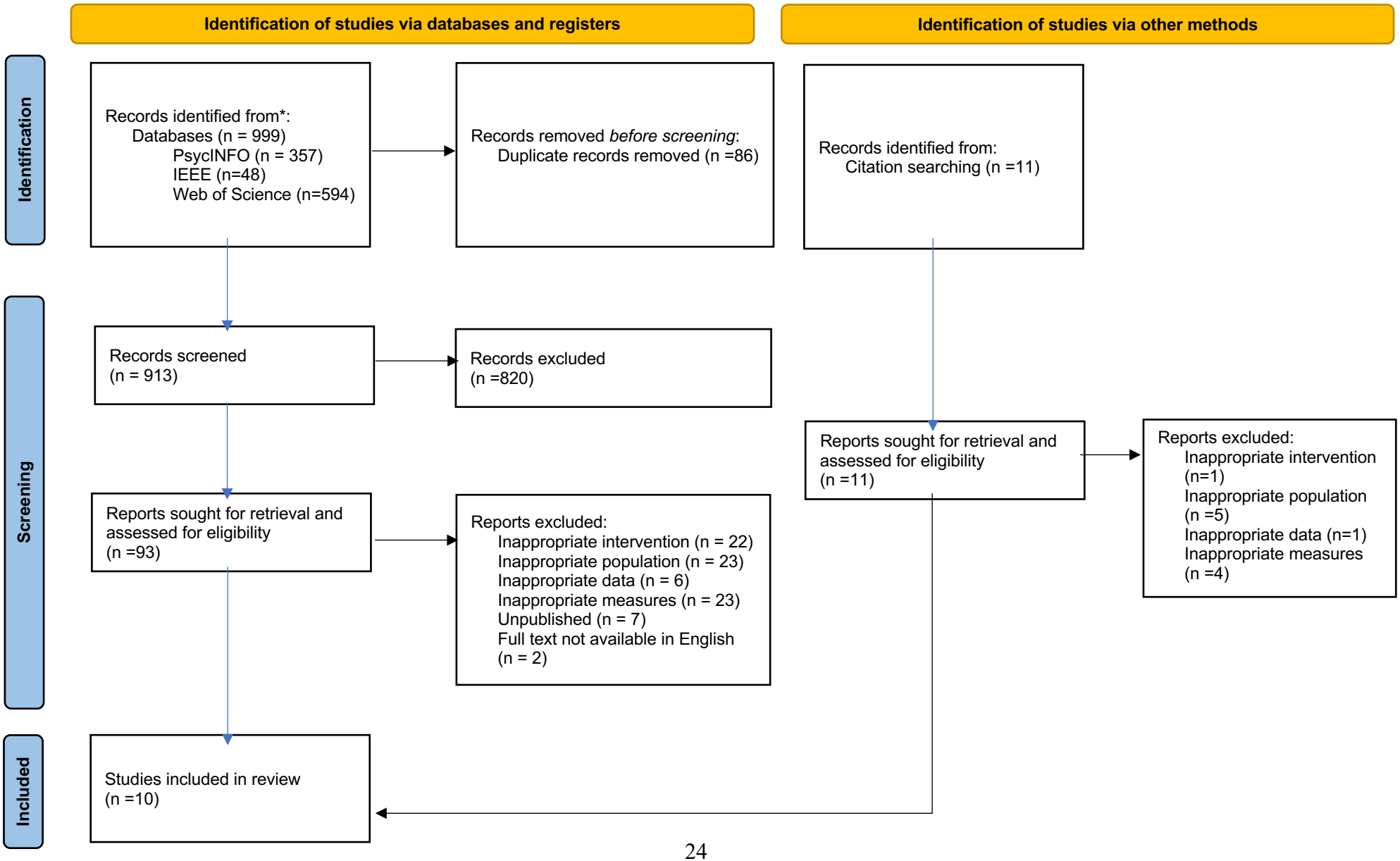
Study Screening and Selection

Selection process was outlined using the updated PRISMA Flow Diagram (Page et al., 2020) in Figure 1. Initially 999 studies were identified through the databases PsycINFO, IEEE and Web of Science. Following the removal of duplicates, 913 papers remained, which were screened jointly with the second reviewer as described above. 93 papers were selected for full text review and all were able to be retrieved. Following full text review, 83 papers were excluded due to reasons detailed in the PRISMA diagram and the remaining 10 papers were included in the review.

Citation searching was conducted in the reviews excluded at the initial screening stage to identify relevant papers and eleven potential papers were identified through this process. These papers were screened against the inclusion criteria and none were deemed to be suitable for inclusion in the review.

Figure 1.

PRISMA Flow Diagram



Study Characteristics

The selected 10 studies were summarized according to their participant characteristics and study design and intervention details in Tables 2 and 3.

A total of 1049 participants were included in the studies, with participant sizes ranging from five to 203. Most participants were students recruited from universities (651 students, 62% of overall participants and seven out of 10 studies). Two studies recruited participants from schools and one study recruited from a summer school (total of 398 students, 38% of overall participants). Participant ages ranged from 11 to 24, with a majority of the sample representing the older end of the range due to university samples used (seven out of 10 studies with mean age over 18). All studies had mixed gender participants with seven of the studies including predominantly female participants (51.2% to 80% of the sample). Most studies took place in the USA (seven of the 10 studies) and the remaining studies took place in the China, Taiwan and Spain. Ethnicity was not reported for five of the studies and of the remaining studies, two had samples consisting of predominantly white individuals (75.4% and 70.4%) while the other three had relatively mixed ethnicity samples. Study drop out was not reported in four of the papers and was less relevant for these studies as they tested one off interventions without follow up. Of the studies that reported dropout rates, these ranged from 4-40%.

Table 2.

Participant Characteristics

Author and Year	Population	Number of Participants (included in analyses)	Drop Out	Age	Gender (%Female)	Ethnicity	Location
Browning et al., 2020	University students	82	16 (16%)	Mean = 20 SD = 1.2	47	Non-Hispanic White (35%), Asian/ Asian American (54%), African American/ Hispanic/ Mixed (11%)	USA
Crosswell & Yun., 2020	University students at a midsize northwestern university recruited through subject pool	5	1 (20%) (partial data)	Age range 18-20 (means not reported)	80	Caucasian (n = 2), Hispanic or Latino (n = 1), African American (n = 1), Other (n = 1)	USA
Hadley et al., 2019	Student recruited from urban public middle schools	85	4 (4%)	Mean = 13 SD ranges from 0.91- 0.82 in both arms	55	African American (33%), Mixed race (30%), Other (29-18%), Caucasian (10-18%) Hispanic (42-34%)	USA
Hsieh & Chen, 2019	Fifth and sixth grade students recruited through cluster sampling	123	21 (15%)	Mean = 12.53 SD = 1.03	51.2	Not reported	Taiwan
Liu et al., 2020	University students at a public	60	Not reported	Mean = 20.38 SD = 2.16	50	Not reported	China

	university recruited through WeChat						
Rupp et al., 2019	University students from a large university recruited through participant pool	136	Not reported	Mean = 19.8 SD = 2.44	48.5	Not reported	USA
Ruiz-Ariza et al., 2018	Summer school students	190	63 (25%)	Mean = 13.32 SD = 1.07	49.5	Not reported	Spain
Richesin et al., 2021	Undergraduate psychology students	44	Not reported	Mean = 21 SD = 2.26	81.8	White (70.4%), Hispanic (4.1%), Asian (8.2%), Black (4.1%)	USA
Plante et al., 2003	Undergraduate introduction to psychology students	121	Not reported	Mean = 18.58 SD = 1.12	59.5	Not reported	USA
Pratscher, 2021	University students recruited through participant pool	203	162 (40%)	Mean = 18.88 SD = 3	59.6	White (75.4%)	USA

All included studies were randomized controlled trials (RCT) and took place at the location of recruitment. Six of the studies only looked at the effects of single session interventions, measuring effects of intervention on wellbeing immediately after the intervention. In these single session studies, the intervention durations lasted between five to 30 minutes, with a majority of the interventions (four out of six) lasting under 10 minutes. Studies reporting on longer interventions (four out of 10 included studies) lasted between four to 10 weeks (two four week long studies, one eight week long study and one 10 week long study). Most studies did not have follow ups (eight out of 10) with only one four week long study including a three month follow up and one five minute intervention reporting on one week follow up.

The extended reality element of the included studies heavily focused on VR technology (eight out of 10) with only two studies using augmented reality. Both studies that used augmented reality were similar in design and content, looking at the effects of playing the augmented reality mobile app game Pokemon Go on wellbeing compared to no treatment control groups of students attending regular school classes (Ruiz Ariza et al., 2018; Hsieh & Chen, 2019). The remaining eight interventions using VR technology relied on a range of VR headsets and devices of varying degrees of immersiveness. The intervention contents and comparators of these studies showed large amounts of variation, which can be grouped into the following three broad categories,

a) participating in activities in real life compared to participating in VR: Five studies compared the effects of engaging in real world activities vs immersive virtual reality activities. These included experiencing forest scenery (Browning et al., 2020), participating in meditation (Crosswell & Yun, 2020), practicing didactically learned safety and emotion regulation skills

(Hadley et al., 2019), engaging in drawing (Richesin et al., 2021), engaging in mountain biking (Plante et al., 2003).

b) watching videos in non-immersive vs immersive conditions: Two studies (Rupp et al., 2019; Liu et al., 2020) examined whether watching the same video (scenic sights of China or an astronaut floating in space) on smartphones compared to varying degrees of immersive VR technologies had an impact on wellbeing immediately afterwards.

c) impact of the content of immersive videos: One study looked at the impact of the content of the immersive videos watched and compared whether watching an amusement inducing video has a differential effect on wellbeing than watching an awe inducing video on VR (Pratscher, 2021).

Due to the large amount of variation in the included studies, this grouping was based broadly on similarities in aims and intervention contents of the studies.

Table 3.

Study Design and Intervention Details

Author and Year	Setting	Design	Intervention Content	Intervention Comparison	Delivery Method	Intervention Duration	Follow Up
Browning et al., 2020	Educational centre in the forest	RCT	Being exposed to nature via audio and visuals of the forest.	Group 1: Exposure to nature outdoors in the forest. Participants were asked to sit outside in the forest and observe the nature. Group 2: Watching a 360 degree video of the same forest as the outdoors condition on a VR headset Group 3: Sitting indoors in front of a white wall (no treatment control)	Group 2: 2015 Samsung Gear VR headset with Galaxy Note 5 smartphone inserted	6 Minutes (single session)	Single session intervention. No follow up
Crosswell & Yun, 2020	University	RCT	Watching and/or listening to mindfulness videos filmed in scenic and tranquil settings with university counsellor guiding through mindfulness techniques designed to enhance stress and	Group 1: Unguided self-meditation (Participants did not receive any guidance or mindfulness content (video or audio) and were asked to meditate on their own.) Group 2: audio guided meditation (Participants listened to a	Group 1: No device Group 2: Mobile phone Group 3: Google Daydream VR headset	4 Weeks (once a day for 12 min)	No follow up

			anxiety management skills	specific YouTube video sent by the researchers) <u>Group 3:</u> VR guided meditation (Participants watched a specific YouTube video sent by the researchers on VR. This is the same video as Group 2 but with VR components).			
Hadley et al., 2019	School	RCT	All participants attended a group intervention aimed to enhance emotion regulation skills and reduce poor decision making. This group intervention lasted 4 weeks, once a week and included didactical teaching followed by role play (in the form of Group 1 or 2) to teach sexual and substance use risk information and emotion regulation skills. Topics included relationship between emotions and behaviours,	<u>Group 1:</u> Participants used role plays in an immersive virtual reality environment following the didactic workshop to practice skills covered in the teaching (e.g. assertive communication). <u>Group 2:</u> Participants role played the same scenarios as Group 1 following didactic workshops, however these were done in person in groups of their peers.	<u>Group 1:</u> VR headset (details not reported)	4 weeks (once a week for 2 hours including 1.5hrs of didactic teaching and 0.5hr of virtual reality or role play)	3 Months

			identifying emotional arousal and triggers, coping skills, influence of peer relationships and risk taking.				
Hsieh & Chen, 2019	Primary school	RCT	Playing Pokemon Go (a popular AR mobile app game where players catch and interact with "Pokemons" by completing various tasks and going to different locations. Players can interact with each other during gameplay and encouraged to walk more to level up.)	Group 1: No treatment control. (Participants did not play Pokemon Go and were not asked to participate in any particular activities) Group 2: Participants played Pokemon Go (leisurely, unsupervised, monitored through in-app data)	Group 2: Pokemon Go app and smartphone	10 Weeks (daily, average of 40 min per day, ranging from 13-68 min per day, dependent on participant preferences)	No follow up
Liu et al., 2020	University	RCT	Watching a 360 degree video called "VR China", produced by National Geographic, containing sceneries from Chinese natural and cultural attractions	Group 1: Participants viewed the uplifting 360 degree video on iPhone once Group 2: Participants viewed the uplifting 360 degree video on VR once	Group 1: iPhone 8 with UtoVR app Group 2: Pico 4K G2 VR goggles	8 Minutes (single session)	Single session intervention. No follow up

Rupp et al., 2019	University	RCT	Watching a 360 degree video of an astronaut floating along the International Space Station and listening to an audio track describing different modules and history of the International Space Station	Group 1: Participants watched 360 degree video on smartphone in 3D format Group 2: Participants watched 360 degree video on Google Cardboard Group 3: Participants watched 360 degree video on Oculus Rift DK2 Group 4: Participants watched 360 degree video on Oculus Rift CV1	Group 1: Smartphone Group 2: Google Cardboard Group 3: Oculus Rift DK2 Group 4: Oculus Rift CV1	6.25 Minutes (single session)	Single session intervention. No follow up
Ruiz-Ariza et al., 2018	Summer school	RCT	Playing Pokemon Go (a popular AR mobile app game where players catch and interact with "Pokemons" by completing various tasks and going to different locations. Players can interact with each other during gameplay and	Group 1: No treatment control (Participants did not play Pokemon Go and were not asked to participate in any particular activities) Group 2: Participants played Pokemon Go (leisurely, unsupervised, monitored through in-app data)	Group 2: Pokemon Go app and smartphone	8 Weeks (daily, average of 40 min per day)	No follow up

			encouraged to walk more to level up.)				
Richesin et al., 2021	University	RCT	VR controls interacted with an office room in the simulation, 2D and 3D groups were instructed to draw freely.	Group 1: VR active control. Participants were told to move freely in the VR simulation and interact with objects such as stapler, drawers etc. Group 2: Participants engaged in 2D regular drawing using crayons, markers and pens on paper. Group 3: Participants engaged in 3D VR drawing.	Group 1: HTC Vive headset for the VR control group Group 3: Oculus Quest headset and Google Tilt brush app	15 Minutes (single session)	Single session intervention. No follow up
Plante et al., 2003	University	RCT	Engaging in "Trek Extreme Mountain Biking" VR video game or VR biking experience using 2000 Cycle FX	Group 1: Watching a non-XR video about biking (active control) Group 2: Playing a mountain biking VR video game Group 3: Riding a stationary exercise bike only	Not reported	30 Minutes (single session)	Single session intervention. Follow up same day at bedtime

				Group 4: Riding a stationary bike while engaging in a VR biking experience			
Pratscher, 2021	University	RCT	Awe condition: 5 min video of a drone flying over the Alps in 4k definition (landscapes, panoramic views of nature) Amusement condition: 5 min video of animals acting like humans (e.g. walrus playing a saxophone)	Group 1: Participants watched an awe inducing video on VR once Group 2: Participants watched an amusement inducing video on VR once	Groups 1 & 2: Oculus Go VR headset	5 Minutes (single session)	Single session intervention. Follow up at 1 Week

Quality Assessment

The quality of bias assessments for included studies using the appropriate version of the Cochrane Revised Risk Bias Tool were summarized in Tables 4, 5 and 6. Eight studies were assessed using the Cochrane Risk of Bias Tool for Individually Randomized Parallel Group Trials, one study was assessed using the Cochrane Risk of Bias Tool for Crossover Randomized Controlled Trials and one study was assessed using the Cochrane Risk of Bias Tool for Cluster Randomized Controlled Trials.

Of the eight studies assessed using the Cochrane Risk of Bias Tool for Individually Randomized Parallel Group Trials, three were found to have low risk of bias, two had some risk of bias and three had high risk of bias. All three of the studies deemed to have high risk of bias were assessed to be so mainly due to issues with the selection of reported results, particularly related to results being likely to have been selected from multiple eligible analyses of the data and due to results being likely to have been selected on the basis of multiple eligible outcome measurements. For one of these studies, issues were also highlighted around baseline differences between intervention groups suggesting a problem with the randomization process and large amounts of missing data. Other difficulties identified with the studies rated as having some concerns for risk of bias were related to lack of clarity around the randomization process, significant amount of missing data and its impact on the findings and issues with selection of reported results.

Table 4.

Risk of Bias Assessment for Individually Randomized Parallel Group Trials

Author and Publication Date	Randomization Process	Deviations from Intended Interventions	Missing Outcome Data	Measurement of Outcome	Selection of Reported Results	Overall Risk of Bias
Browning et al., 2020	Low	Low	Low	Low	Some	Some
Hadley et al., 2019	Low	Low	Low	Low	High	High
Liu et al., 2020	Low	Low	Low	Low	Low	Low
Plante et al., 2003	Low	Low	Low	Low	High	High
Pratscher, 2021	Some	Low	Some	Low	High	High
Richesin et al., 2021	Low	Low	Low	Low	Low	Low
Ruiz Ariza et al., 2018	Some	Low	Some	Low	Low	Some
Rupp et al., 2019	Low	Low	Low	Low	Low	Low

One study was assessed using the Cochrane Risk of Bias Tool for Crossover Randomized Controlled Trials and was found to show high concern for risk of bias. This was mainly due to carryover effects, with the participants alternating between treatment conditions randomly throughout the intervention. This did not allow sufficient time to pass to eliminate any carryover effects of learned skills and raised queries in the data analysis. There were also some concerns identified in relation to measurement of outcomes as individuals rating their subjective outcomes could have been influenced with their knowledge of the treatment conditions and their beliefs regarding their effectiveness.

Table 5.

Risk of Bias Assessment for Crossover Randomized Controlled Trials

Author and Publication Date	Randomization Process	Period and Carryover Effects	Deviations from Intended Intervention	Missing Outcome Data	Measurement of Outcome	Selection of Reported Result	Overall Risk of Bias
Crosswell & Yun, 2020	Low	High	Low	Low	Some	Some	High

The remaining study was assessed using the Cochrane Risk of Bias Tool for Cluster Randomized Controlled Trials and was deemed to show some concerns for risk of bias. This was due to lack of information on whether the allocation sequence was concealed until clusters were assigned to interventions (whether schools were told they would be control vs intervention) and baseline differences in groups (evidently more boys in experimental group and more girls in control group) that may suggest differential recruitment of participants as the authors may have expected boys to react more positively to the AR game based on previous literature.

Table 6.

Risk of Bias Assessment for Cluster Randomized Controlled Trials

Author and Publication Date	Randomization Process	Timing of Identification or Recruitment of Participants	Deviations from Intended Interventions	Missing Outcome Data	Measurement of Outcome	Selection of Reported Results	Overall Risk of Bias
Hsieh & Chen., 2019	Some	Some	Low	Low	Low	Low	Some

Study Outcomes

The outcome measures used to assess wellbeing and the findings of the studies were outlined in Table 7. Results of the studies were analyzed based on the presence of within groups effects (whether XR interventions are effective in improving wellbeing) and between groups effects (whether they are more effective in improving wellbeing compared to non-XR controls). Whilst improvements in mean participant wellbeing scores were reported in most of the studies, these changes were often not statistically significant. It should be noted that only two of the included studies reported their a-priori power analyses (Browning et al., 2020 and Hadley et al., 2019) and the lack of significant findings could be attributed to small sample sizes and lack of sufficient power to detect changes as statistically significant, as effect sizes associated with these changes were often small.

Outcome Measures

A range of outcome measures were used to assess wellbeing in the included studies. Most measures used in the selected studies focused on emotional/ hedonic aspects of wellbeing, such as PANAS and Affect Dysregulation Scale (six out of 10 studies). Cohen's Perceived Stress Scale was used by two studies and covers an explicit element of positive functioning (eudaimonic wellbeing). One study (Pratscher, 2021) used a non-standardized measure of wellbeing that was composed of different items from various questionnaires assessing subjective wellbeing, autonomy, environmental mastery, personal growth, positive relations, purpose in life, self-acceptance and self-actualization. The lack of standardization was a limitation of this study.

Effect of Participating in Activities in Real Life Compared to Participating in VR

Five studies explored the effect of participating in activities in real life compared to in XR (Browning et al., 2020; Crosswell & Yun, 2020; Hadley et al., 2019; Richesin et al., 2021; Plante et al., 2003) and assessed the impact of participating in these activities on wellbeing.

Four of the five studies (Browning et al., 2020; Hadley et al., 2019; Richesin et al., 2021; Plante et al., 2003) found improvements in wellbeing in the immersive condition but these improvements were not better than the improvements observed in the real life activities conditions.

Browning et al. (2020) found significant reductions in negative affect in all XR and non-XR experimental conditions but did not find any differences between conditions. Results indicated that being outdoors had a positive impact on wellbeing, with participants in this group experiencing a significant improvement in their positive and negative affect. Results for VR group were less robust as there was no change in positive affect which is a stronger indicator of wellbeing and a significant but small reduction in negative affect. The results showed that the VR and the outdoors conditions lead to more positive affect compared to the indoor control condition; however, the VR condition did not have an added benefit over experiencing nature in the real world. The VR condition had better wellbeing scores compared to the indoors no-treatment control group, but this was also due to reduction in positive affect in the control group. As the study used an adapted version of the PANAS scale, it was not possible to comment on the clinical significance of these changes. The a-priori power analysis indicated that a sample size of 90 was required for the analyses to be sufficiently powered. This was not achieved (N = 82) suggesting that the study is slightly underpowered.

Hadley et al., (2019) reported that participants in both experimental conditions became more emotionally aware after the interventions, where they received a didactic information on emotion regulation followed by practice sessions in VR or via role plays. The groups did not significantly differ in relation to affect dysregulation, emotional self-efficacy and emotional awareness. The a-priori power analysis conducted by the authors suggested a sample size of 85 to detect moderate effect sizes, which was met by the recruitment and retained by the follow up. Many analyses in the study yielded small effect sizes (e.g. emotional self-efficacy and affect dysregulation), where the sample may not have been sufficiently powered to detect differences. Nevertheless, the clinical significance of such changes is unclear. The main difference between groups was due to a deterioration in the role play condition in their ability to access emotion regulation strategies, whereas the VR group remained stable. Overall, the findings did not suggest an advantage of using VR above and beyond the role play method. This study had a younger age group (mean age = 13) compared to the other studies in this category (mean ages ranging from 18.5 to 21). There were no outstanding differences with regards to wellbeing effects between Hadley et al (2019) and others in this category, although comparisons were challenging due to differences in the outcome measures used (Hadley et al (2019) was the only study that used emotion regulation related measures).

Plante et al. (2003) observed an improvement in energy levels and tiredness in the experimental conditions (VR only, exercise with VR and exercise only), whereas the control group experienced a deterioration on those measures at the end of the intervention. There were no significant differences between the VR and the non-VR exercise intervention groups. The authors concluded that while exercise seemed to have a positive effect on wellbeing, there didn't seem to be an added benefit of VR. There was a positive effect for females in their energy

and tiredness levels in the exercise with VR condition at bedtime, which the authors reported could be due to the intervention or Type 1 error. Whilst this group had more positive scores compared to the other conditions at the bedtime follow up assessment, their results indicated more tiredness and less energy levels compared to their baseline assessment. The study also found that individuals in the experimental conditions experienced less calmness and more tension after the intervention, whereas the control group had the opposite effect. The findings on these variables were unexpected to the authors and suggested a mixed intervention effect on wellbeing.

Richesin et al. (2021) found that all experimental groups experienced improvements in their wellbeing scores; however, there were no significant differences between or within groups. All participant groups experienced improvement in mean scores in relation to their mood and perceived stress, showing that as a single session intervention, participating in VR, drawing by hand or drawing in VR does not differentially influence wellbeing. The power analysis was not reported, however authors commented on the likelihood of the small sample size leading to low power and higher risk of Type 2 error.

Crosswell & Yun (2020) was the only study in this category that did not find improvements in wellbeing for the immersive condition. The study reported that there was no improvement in participants' stress levels overall at the end of the trial. The mean changes reported in the study were in the positive direction, suggesting an increase in perceived stress in all groups throughout the trial, as higher PSS scores indicate higher levels of perceived stress. However, this was not clarified in the report. The methodological issues in this study caused challenges in the interpretation of its findings, as there were significant carryover effects that need to be taken into consideration. The mindfulness and mediation knowledge and skills gained

over the course of the trial from various experimental conditions could hinder any meaningful comparisons between experimental conditions. A-priori power analyses were not reported but the study only had four participants and is likely to be underpowered. The authors reported the small sample size and lack of power as limitations and reported that more sensitive outcome measures could have been used. This study was included in the review to illustrate different approaches that have been used applying XR technology to wellbeing in adolescents and young people to date; however, due to significant methodological issues it is not possible to draw conclusions with regards to intervention effectiveness.

Overall, studies did not find evidence to suggest that participating in activities on VR lead to better wellbeing outcomes than doing activities in real life. VR intervention groups lead to better wellbeing outcomes compared to no-treatment control groups and non-XR active treatment control groups, but did not outperform "real life" activities intervention groups or active VR controls. Participants in the VR intervention groups reported better wellbeing after the interventions compared to baseline in all studies except for Crosswell & Yun (2020), which was deemed to have a high risk of bias due to crossover effects.

The Effect of Watching Videos under Non-immersive vs Immersive Conditions

Two studies compared the effects of watching videos on immersive and non-immersive devices (Rupp, 2019; Liu, 2020) and assessed impacts on affect. Liu et al (2020) found that both the immersive and the non-immersive conditions significantly increased positive emotions and decreased negative emotions within groups. The improvement in positive affect was more prominent in the immersive (VR) group, with the analyses indicating a medium size effect within the VR group pre to post intervention ($d = 0.56$, compared to $d = 0.35$ in smart phone group) and a medium size effect difference between both experimental groups at post intervention ($d =$

0.35). In the Liu et al (2020) study, watching videos using immersive devices improved wellbeing significantly more than watching videos on non-immersive devices.

Similarly, Rupp et al (2019) also found an increase in positive affect and a decrease in negative affect, but only for the XR conditions, showing that watching videos on immersive devices led to improvement in wellbeing whereas doing so on non-immersive devices did not change self-reported wellbeing. The increase in positive affect in the XR conditions showed a small effect size and was significant ($\beta = 0.19$, $p = 0.01$). Both studies did not find any significant differences in negative emotions between groups. Power analyses were not reported for either study; however, Liu et al (2020) listed their small sample size as a limitation that may have affected the randomization process. Both studies were RCTs with low risk of bias and used the same outcome measure (PANAS) under different cultural contexts, suggesting that these results may be more robust for this research design.

Overall, these results showed that watching the same videos on immersive devices lead to improvements in wellbeing (as measured by affect) immediately after the viewing, whereas same wellbeing effects were not observed when participants watched videos on non-immersive devices.

The Impact of the Content of Immersive Videos

Pratscher (2021) was the only included study looking at the impact of the content of immersive videos on wellbeing, comparing the effects of watching awe-inducing or amusement-inducing videos using VR. The results showed some statistically significant differences within groups on some domains of wellbeing, such as a decrease in personal growth and an increase in self-actualization and purpose in life domains; however, these differences were small (effect sizes ranging from 0.13 to 0.29). The author concluded that these VR interventions did not lead

to any sustainable and meaningful changes in hedonic and eudemonic wellbeing of participants in either experimental group. The power calculations were not reported in this study, however, there was significant attrition in the sample, which could have altered the findings. The findings of the study were reported for the different domains of wellbeing rather than for a composite wellbeing measure overall. This, as well as a lack of standardized measures in this study posed challenges to the interpretation of the effects on wellbeing.

The Effect of Playing Immersive Games

Two studies (Ruiz Ariza et al., 2018 and Hsieh & Chen, 2019) examined the longitudinal effect of playing immersive games, specifically the mobile AR game Pokemon Go, on wellbeing of school students. There were no significant changes on the TEIQue-SF wellbeing scale within the AR groups or between AR groups and control groups of students for both studies, suggesting that solely playing AR games for a period of time did not lead to improvement or deterioration of wellbeing in adolescents and young people. A limitation of these studies was that both trials used the same intervention and the same outcome measure, limiting the generalizability of these conclusions to other AR games.

Key Findings

- There is evidence to suggest that doing activities in VR leads to improvement in wellbeing compared to not engaging in a particular activity; however doing activities using VR does not lead to better wellbeing than doing activities in real life.
- There is evidence to suggest that watching immersive videos leads to improvement in wellbeing.
- Playing regular AR games (that were not purposefully developed as wellbeing interventions) does not lead to changes in wellbeing.

Table 7.

Study Measures and Outcomes

Author and Year	Wellbeing Measures	Wellbeing Results
Browning et al., 2020	Positive and Negative Affect Scale (PANAS) (adapted version)	**Within Group Effects** Group 1 (outdoors condition) showed slightly statistically significant improvement in positive affect (mean increased from 3.37 (SD = 0.49) to 3.54 (SD = 0.66), t(21) = 2.14, p = 0.04, d = 0.29) and decrease in negative affect (mean decreased from 1.23 (SD = 0.20) to 1.17 (SD = 0.23), p = 0.034, d = -0.28). Group 2 (VR condition) did not show changes in positive affect (mean not reported, t(29) = 0.28, p = 0.78) and showed statistically significant reduction in negative affect (mean decreased from 1.32 (SD = 0.37) to 1.24 (SD = 0.46), p = 0.03, d = -0.19). Group 3 (indoors control) showed statistically significant reduction in positive affect (mean decreased from 3.00 (SD = 0.68) to 2.42 (SD = 0.71), t(29) = -4.94, p < 0.001, d = -0.83) and reduction in negative affect (mean decreased from 1.38 (SD = 0.27) to 1.20 (SD = 0.57), p < 0.001, d = -0.40). **Between Groups Effects** Stepwise regression models were used to assess differences in affect between groups. These showed that VR and outdoors conditions resulted in higher positive affect compared to indoors control condition. There were no differences between groups in changes in negative affect.
Crosswell & Yun, 2020	Cohen's Perceived Stress Scale (PSS)	**Within Group Effects** There seem to be increases in perceived stress within groups, but this is unclear in the report. **Between Groups Effects** There were non-significant differences between the VR (ΔMean = 3.80, SD = 0.86), sound only (ΔMean = 4.00, SD = 0.66) and self-guided (ΔMean = 3.92, SD = 0.64) groups (F(not reported) = 0.33, p = 0.72).

Hadley et al., 2019	a) Affect Dysregulation Scale	**Within Group Effects** There were small non-significant increases in affect dysregulation in the VR group (d = 0.21, p = 0.35) and moderate non-significant increases in the role play group (d = 0.36, p = 0.10). **Between Groups Effects** The difference in affect dysregulation between groups was non-significant (d = -0.13, p = 0.59).
	b) Difficulties in Accessing Emotion Regulation Strategies Scale	**Within Group Effects** The VR group did not show any significant changes in their difficulty in accessing emotion regulation strategies (d = 0.06, p = 0.74). The role play group showed moderate increases (d = 0.71, p < 0.01). **Between Groups Effects** There were moderate size differences between VR and role play groups on accessing emotion regulation strategies (d = -0.46, p < 0.05) due to a deterioration in the role play group.
	c) Emotional Self Efficacy Scale	**Within Group Effects** There was a small non-significant improvement in emotional self-efficacy in the VR group (d = 0.26, p = 0.23). **Between Groups Effects** There were small non-significant differences between groups in emotional self-efficacy (d = 0.20, p = 0.36).
	d) Difficulty in Emotional Awareness Scale	**Within Group Effects** Both the VR group (d = -0.50, p < 0.05) and role play conditions (d = -0.61, p < 0.01) became more emotionally aware. These changes were moderate in size. **Between Groups Effects** Differences between groups in emotional awareness was non-significant (d = 0.09, p = 0.67).
Hsieh & Chen, 2019	Trait and Emotional Intelligence Scale (TEIQue) (wellbeing subscale)	**Within Group Effects** Both the AR group (p = 0.901) and the control group (p = 0.12) did not show any significant changes in their wellbeing. **Between Groups Effects**

		The differences between AR and control group on the wellbeing measure were not significant ($p > 0.05$).
Liu et al., 2020	Positive and Negative Affect Scale (PANAS)	**Within Group Effects** The VR group experienced an increase in positive emotions (M(baseline) = 30.2 (SD = 7.14) to M(end) = 34 (SD = 6.55), $F(1, 29) = 21.93$, $p < 0.001$, $d = 0.56$) and a decrease in negative emotions (M(baseline) = 17.93 (SD = 7.90) to M(end) = 13.80 (SD = 5.15), $F(1, 29) = 22.96$, $p < 0.001$, $d = -0.62$). The smart phone group also experienced an increase in positive emotions (M(baseline) = 29.27 (SD = 5.22) to M(end) = 31.53 (SD = 7.49), $F(1. 29) = 4.746$, $p = 0.038$, $d = 0.35$) and a decrease in negative emotions (M(baseline) = 18.73 (SD = 8.16) to M(end) = 14.10 (SD =5.01), $F(1, 29) = 25.07$, $p <0.001$, $d = -0.68$). **Between Groups Effects** VR group reported more positive emotions compared to the smart phone group ($d = 0.35$, $p = 0.037$) at the end of the intervention, but there were no significant differences between groups with regards to negative emotions.
Rupp et al., 2019	Positive and Negative Affect Scale (PANAS)	**Within Group Effects** There were small increases in positive affect and decreases in negative affect for all three of the immersive conditions (Cardboard condition: ΔM(positive affect) = 0.22, ΔM(negative affect) = -0.03; DK2: ΔM(positive affect) = 0.11, ΔM(negative affect) = -0.13; CV1: ΔM(positive affect) = 0.40, ΔM(negative affect) = -0.08) but not for the phone condition (ΔM(positive affect) = -0.07, ΔM(negative affect) = 0.01). **Between Groups Effects** There was a main effect of immersion (XR interventions) on positive affect showing significant pre to post change ($t = 2.24$, $p = 0.01$), which was reported to indicate a small effect size ($\beta = 0.19$). There were no significant results for pre to post change for negative affect ($p = 0.33$). Immersion led to greater positive affect, but there was no linear increase in positive affect as level of immersion increased.

Ruiz-Ariza et al., 2018	Trait and Emotional Intelligence Scale-Short Form (TEIQue-SF) (wellbeing subscale)	**Within Groups Effects** Wellbeing scores did not change over time ($p > 0.05$). **Between Groups Effects** There were no significant differences in wellbeing scores between control and experimental condition at post intervention ($p= 0.90$).
Richesin et al., 2021	a) Positive and Negative Affect Scale (PANAS)	**Within Group Effects** Changes in wellbeing scores were uniform across groups (positive affect: $F(2, 41) = 0.21$, $p = 0.82$, $d = 0.20$; negative affect: $F(2, 41) = 1.03$, $p = 0.37$, $d = -0.45$) and all groups showed improvement in wellbeing (control group: positive affect ($\Delta M = 0.94$, $SD = 4.33$), negative affect ($\Delta M = -2.44$, $SD = 2.90$); 2D group: positive affect ($\Delta M = 0.77$, $SD = 3.24$), negative affect ($\Delta M = -4.39$, $SD = 5.25$); 3D group: positive affect ($\Delta M = 1.60$, $SD = 3.31$), negative affect ($\Delta M = -3.73$, $SD = 2.92$). **Between Groups Effects** There were no significant differences between experimental groups for positive affect ($F(2, 41) = 0.12$, $p = 0.88$, $d = 0.16$) or negative affect ($F(2, 41) = 0.76$, $p = 0.47$, $d = 0.39$).
	b) Cohen's Perceived Stress Scale (PSS)	**Within Group Effects** Changes in wellbeing scores were uniform across groups for perceived stress ($F(2, 41) = 1.01$, $p = 0.38$, $d = -0.44$) and all groups showed improvement in perceived stress (control group: $\Delta M = -1.31$, $SD = 2.41$; 2D group: $\Delta M = -1.46$, $SD = 3.05$; 3D group: $\Delta M = -0.13$, $SD = 2.90$). **Between Groups Effects** There were no significant differences between experimental groups for perceived stress ($F(2, 41) = 0.38$, $p = 0.69$, $d = 0.27$).
Plante et al., 2003	Activation Deactivation Adjective Checklist	**Within Group Effects** Participants in the experimental conditions (VR exercise, non-VR exercise and VR only) experienced more energy ($F(6, 13) = 8.97$, $p < 0.05$) and less tiredness ($F(3, 113) = 4.21$, $p < 0.05$) after the intervention compared to baseline than the control group (video only).

		Between Groups Effects There was a main effect of experimental group for energy scores ($F(3, 113) = 4.68$, $p < 0.05$), with participants in the two exercise conditions (exercise with VR and exercise without VR) reporting higher energy levels compared to the control and the VR alone conditions ($p < 0.05$). A main effect was also found for tiredness ($F(3, 112) = 6.01$, $p <0.05$), with all three experimental conditions (exercise with VR, exercise without VR and VR only) reporting feeling less tired compared to the control group ($p < 0.05$). Participants in the experimental conditions reported less calmness and more tension, or no difference, after the intervention. The control group reported more calmness after the intervention. Further analyses were not conducted on these variables as the authors reported that these variables failed to pass the manipulation check.
Pratscher, 2021	Hedonic and eudaimonic wellbeing measured by composite measures of different items from non-standardized questionnaires	**Within Group Effects** The awe group experienced a decrease in personal growth (M(baseline) = 6.25 (SD = 0.92) to M(FU) = 5.96 (SD = 1.07), $p < 0.05$, $d = -0.29$). The amusement group experienced a decrease in personal growth (M(baseline) = 6.21 (SD = 0.88) to M(FU) = 5.99 (SD = 1.18), $p < 0.05$, $d = -0.21$) and an increase in purpose in life (M(baseline) = 5.24 (SD = 1.18) to M(FU) = 5.39 (SD = 1.15), $p < 0.05$, $d = 0.13$) and self-actualization (M(baseline) = 3.62 (SD = 0.50) to M(FU) = 3.75 (SD = 0.60), $p < 0.05$, $d = 0.24$). **Between Groups Effects** There were no significant differences between groups (repeated measures ANOVA results not reported).

Note. P values were reported as they were reported in the original articles. Exact values were reported where these were made available. Effect sizes were either taken from the original articles or calculated, where sufficient data was available.

Discussion

Effectiveness of XR Interventions for Adolescent and Young People's Wellbeing

The present review investigated the effectiveness of XR interventions on improving wellbeing of adolescents and young people. Participants of the included studies were between the ages of 11 and 24, with most of the studies (70%) recruiting from university samples. The majority of the interventions (80%) used VR technology, which was similar in reviews with older populations and may be due to the AR literature on wellbeing and mental health being in its early stages (Carroll et al. 2021). The intervention contents focused on participating in activities in real life compared to in VR (50%), watching videos in immersive or non-immersive conditions (20%), playing AR game (Pokemon Go) (20%) and comparing content of immersive videos (10%). Results were mixed; ten controlled studies were identified, of which two found statistically significant improvements in wellbeing (Liu et al., 2020; Rupp et al., 2019), four found minor improvements (Richesin et al., 2021; Browning et al., 2020; Plante et al., 2003; Hadley et al., 2019) and four did not find any improvements in wellbeing (Crosswell & Yun, 2020; Hsieh & Chen, 2019; Ruiz Ariza et al., 2018; Pratscher, 2021) for their participants following the XR intervention.

Studies that compared participating in wellbeing activities in the real world as opposed to on VR (Browning et al., 2020; Hadley et al., 2019; Richesin et al., 2021; Plante et al., 2003) found some trends to suggest minor improvements in wellbeing of participants in the VR groups; however, these interventions did not outperform engaging in activities in real life. Except for one study (Crosswell & Yun, 2020) that had significant methodological issues, none of these VR interventions lead to decreases in self-reported wellbeing scores of participants. Considering that participating in leisure activities improves wellbeing (Kuykendal et al., 2015), these VR

interventions mimicking real world leisure activities, can be used as short term strategies to improve mood when needed. Individuals, who lack the opportunity or ability to participate in leisure activities in the real world, such as patients in inpatient settings or individuals who have disabilities or health conditions can use VR alternatives to experience wellbeing benefits of these activities.

The most consistent findings were observed for the two studies that compared the effects of watching videos on smartphones compared to watching the same video on immersive VR headsets (Liu et al., 2020; Rupp et al., 2019). Both studies found that watching videos on VR lead to significant improvements in mood and hedonic wellbeing, particularly in improving positive affect, immediately following the intervention, indicating that using VR can be an enjoyable experience that may potentially be more engaging and impactful. The short duration of these interventions, the lack of follow up assessments and the focus of the selected outcome measure (PANAS (Crawford et al,, 2004)) on purely immediate effects on affect makes it difficult to conclude whether these interventions can be used as wellbeing interventions in clinical or non-clinical settings. Nevertheless, the results indicate that these interventions can provide ad hoc emotional benefit to its users and that VR as a mode of delivery is compatible with wellbeing content and can potentially play an enhancing role.

Ruiz Ariza et al. (2018) and Hsieh & Chen (2019) were the only studies using AR technology and were similar in their study design with regards to the intervention used, duration of the study, sample population and outcome measures. There were no changes in wellbeing of participants in either study. This replication of findings in different cultural contexts (Spain and Taiwan), where relationship of adolescents and young people to technology and its availability may differ, indicated that the Pokemon Go AR intervention was unlikely to lead to lasting

changes in wellbeing, with follow up over eight to 10 weeks. As sense of presence and immersiveness have been found to be linked to the positive effects observed in VR interventions (Rupp et al., 2019), the limited opportunities for these in AR interventions could potentially explain the lack of effect in the AR studies in this review. However, the small number of AR studies and the lack of diversity in their content limit conclusions being drawn on the effectiveness of AR-based interventions at this stage. Further research using a wider range of interventions and outcome measures is needed to draw conclusions on the use and effectiveness of AR based interventions in wellbeing overall.

The content of the interventions varied significantly in the included studies and included interventions that were specifically developed for the trials as well as those that were simply tested or adapted. Of the studies included, only two studies used psychologically informed content in their interventions, one for practicing mindfulness (Crosswell & Yun, 2020) and the other for applying emotion regulation strategies (Hadley et al., 2019). Hadley et al (2019) found some positive effects on emotion awareness and emotional self-efficacy but the results were mixed overall.

Limitations of the Literature

The questionnaires used by the included studies mostly focused on hedonic aspects of wellbeing, measuring immediate effects on mood (e.g. PANAS). Such outcome measures, measuring current affect, run the risk of emphasizing valence and intensity of emotional experiences in the moment and measuring experiences too narrowly to capture an overview of general subjective wellbeing (Rice & Shorey-Fennell, 2020). Wellbeing outcome measures, assessing wellbeing in a broader and more comprehensive way were not utilized by the studies included in the review. One example from the adolescent wellbeing literature to illustrate the

difference in comprehensiveness of outcome measures is the Leventhal et al (2015) study, that used a well validated psychological wellbeing scale and a social wellbeing scale (KIDSCREEN-52 Psychological Wellbeing Subscale and Social Support and Peers Subscale) in conjunction with two resilience scales (Connor-Davidson Resilience Scale-10 and Child and Youth Resilience Measure-28) (alongside self-efficacy, depression and anxiety measures) to assess the impact of a resilience based intervention on wellbeing. Instead, questionnaires in the included studies focused on the immediate effects on affect, which was appropriate for the study designs of the majority of the studies in this review, as they were short interventions without follow up assessments. This choice of outcome measures, however, limited conclusions that could be drawn on wellbeing overall as a concept, consisting of both hedonic and eudaimonic wellbeing beyond transient mood effects. Further information is needed on any changes in thoughts, behaviors and attitudes related to wellbeing to assess the clinical meaningfulness of these XR tools as wellbeing interventions.

Clinical significance of results was often challenging to comment on due to the nature of wellbeing outcome measures. The outcome measures included, often used gradients to describe wellbeing of participants (e.g. higher scores showing better wellbeing) rather than utilizing cut off scores or categories, as they are not clinical measures. The review focused on statistical significance; however, this does not always manifest clinically meaningful changes for the participants. Clinical meaningfulness of findings were often not commented on by authors.

Most of the studies in the present review lacked follow up measurements. Thus, it is unclear whether the positive effects found in some of the studies would lead to any lasting impacts. Considering that XR technology is not widely consumed by the general public at this time, the research designs of several of the included studies are insufficient to distinguish true

effects of interventions compared to novelty effects of having gained a new experience. The novelty effect, noted as a potential factor contributing to positive wellbeing changes in included studies (e.g. Richesin et al. 2021), is likely to wear off as individuals become more used to the XR technology, whether through repeated exposures in trial settings or through exposures to the technology as it becomes more widely available in the society (Merchant et al., 2014). Longitudinal studies are needed to determine if XR interventions can lead to changes in wellbeing above and beyond the initial enjoyment of participating in a new experience.

The novelty effect can also impact sampling and recruitment. Findings of research in this field should be interpreted with caution as the study designs are prone to self-selection biases. The majority of the studies employed an opportunistic sampling strategy by advertising their study on university student platforms (e.g. social media platforms (Liu et al., 2020); subject pool (Crosswell & Yun, 2020)). This sampling strategy is more likely to lead to self-selection bias compared to random sampling (Tyrer & Heyman, 2016), as students interested in the topic of research would be the ones volunteering to participate. Studies investigating immediate mood and wellbeing effects after experiencing VR technology (e.g. Liu et al., 2020 and Rupp et al., 2019) are likely to attract individuals who are keen to participate in these interventions, leading to the self-selection bias. These individuals may have an affinity for technology or similar experiences and may be more likely to enjoy these interventions, or may have never experienced XR and may have a particularly strong novelty effect reaction. Similarly, individuals who volunteer for these studies and end up in the control groups, may feel disappointed, which may explain some of the decreases in wellbeing in these groups. Blinding and concealment of allocations has been highlighted as a requirement to achieve good quality research in VR (Greenleaf et al., 2019).

In addition to affinity, familiarity and competence in relation to the technology could potentially influence outcomes, with individuals who are more familiar with navigating these devices finding it easier to immerse themselves in the content or show higher motivation to continue engaging with the interventions. These factors were not assessed or commented on in the included studies. Although gaps in competence with XR technology are likely to decrease as technology becomes more widely available, particularly in younger populations, a better understanding of how these factors affect engagement and intervention outcomes at present, could inform further adaptations (e.g. inclusion of training sessions) and would therefore be valuable to examine in future studies.

Exclusion criteria of the included studies around suitability for XR (e.g. motion sickness (Hadley et al., 2019)) can reduce their generalizability and skew findings, as individuals who get excluded for these reasons are unlikely to experience positive effects. The present review did not identify any studies involving participants with physical disabilities that fit the inclusion criteria of this review. Trials using XR with individuals with physical disabilities (Singh et al., 2017; Wille et al., 2009) as well as intellectual and developmental disabilities (Lotan et al., 2010; Butti et al., 2020) have shown that these technologies can be adapted for individuals with different needs; however, it is possible that individuals with certain conditions or disabilities (such as visual impairment) may struggle to or may not be able to use these technologies. Further research is needed to understand which populations could benefit from these XR interventions and whether advances in technology could help overcome any existing barriers.

The majority of included studies consisted of participants recruited from universities. University student samples are likely to be similar overall with regards to familiarity with technology, academic achievements/cognitive capacity and health, which are likely to be

different from the general population. University students also represent the older end of the adolescence period and are likely to respond differently to XR interventions compared to younger children. The small number of studies on younger age groups (only three out of 10 studies with mean age under 18 years) in the present review prevented meaningful comparisons across developmental stages to determine differences in the effectiveness of XR interventions. There were no notable differences in the findings that could be attributed to age. Two of the three studies with younger samples (Hsieh & Chen, 2019 and Ruiz Ariza et al., 2018) did not report improvements in wellbeing, although both studies were very similar in their design and used the same intervention (AR Pokemon Go game), which meant that factors other than age could have contributed to these findings (design, content etc.). Therefore, to develop a comprehensive understanding of the impact of XR interventions on wellbeing of adolescents overall as a group, further studies are needed for younger age groups.

The majority of the research included in this review originated from the USA (seven out of the 10 studies), meaning that further research is needed to establish the generalizability of these findings in a global context. There is some evidence to suggest that cultural norms impact on how individuals react to and interact with virtual scenarios (e.g. preferred distance to avatars or reluctance to engage with certain scenarios) (Galina et al., 2018; Almog et al., 2009), meaning that cultural validity of future XR wellbeing interventions should be taken into consideration when administering to users. As the included studies were mostly focused on efficacy rather than effectiveness in real world settings, possible implications of applying these interventions in the community, where barriers around accessibility and socioeconomic considerations are unknown. As investment in the field of XR grows, more affordable versions of the technologies have started to become available. XR interventions have been regarded as “low cost” and accessible

alternatives to other available interventions (e.g. reducing the need for in person delivery of physical rehabilitation with children) (Demers et al., 2020) and there is a move towards using this technology to reduce societal disparities by making treatments available to those living in rural areas or coming from disadvantaged backgrounds (through affordability, portability, flexibility and ease of use) (Zirbel et al., 2018). It is important that future research into XR wellbeing interventions maintain this focus on preventing inequalities to access and explore and report on any challenges with regards to this.

The selection of control interventions in the studies was highlighted as a limitation by several of the authors of included studies. In some instances true controls were not used, while other studies did not have control groups at all. These study designs were vulnerable to the novelty effect not being accounted for and raised queries around true effects in between groups analyses. Future studies can account for novelty effects by including XR control groups and exploring affinity to similar technologies as potential confounders.

The limitations of the literature identified by the present review, such as small sample sizes, lack of diversity in samples, issues with selection of control groups or lack thereof, short length of interventions, lack of theoretical support and psychological underpinnings in study/intervention designs, were issues echoed in previous similar systematic reviews with older populations on XR use in wellbeing (Carroll et al., 2021). These reviews jointly highlight the need for better quality of research designs as the evidence base continues to grow.

Limitations of the Review Process

The present review examined wellbeing as its outcome variable, which is a broad construct with varying definitions in the literature, as outlined previously in the Introduction. Due to the breadth of the construct, broad search terms (“well-being", “wellbeing” and “mental

health") were used when identifying relevant articles in the databases. Identified articles were then screened to determine if they included a wellbeing outcome measure in line with the definition of wellbeing selected for the present review (comprising of hedonic and eudaimonic aspects, as stated in the Introduction). A challenge of this process was deciding which outcome measures would fall under the construct of wellbeing, as the lack of agreement in the literature on the definition of this construct also prevented from the development of clear comprehensive guidelines around recommended use of outcome measures. The approach taken in the present review was examining each outcome variable used by the potential studies and matching them with the concepts mentioned in the definition adopted at the start (e.g. positive affect, positive functioning), with disagreements discussed among the research team. Although the definition of wellbeing selected for the present review was aimed to be a comprehensive one, it is possible that variations in selected studies exist in other similar reviews depending on wellbeing definitions used (for instance, the present review excluded studies with only mental health symptom questionnaires). A more systematic selection of outcome measures could have included using a previous systematic review on young people's wellbeing outcome measures as a guide (such as Croudace et al., 2014); however, this would have also been limiting due to differences in definitions of the construct and the significant variety in the identified/ suggested outcome measures in reviews of this topic (e.g. compared to the review by Kwan & Rickwood, 2015). As there was a significant scarcity of research on the topic of interest for the present review (only 10 studies selected) a limiting approach would not have been appropriate. The present review only had three wellbeing related search terms and these were similar to some other systematic reviews looking at similar topics (e.g. Cheng et al. (2019) and Johnson et al., (2016) with the addition of mental illness terms). The aim for this was to keep the search broad, however, this strategy could

have missed out some other studies using adjacent terms without mentioning “wellbeing” or “mental health”. Based on other reviews on wellbeing, additional search terms in line with our definition of wellbeing could have included “quality of life”, “happiness”, “life satisfaction”, “wellness”, “emotional health”, “psychological health”, “functioning” and “flourishing” (Daykin et al., 2018; Pollard & Lee, 2003; Houlden et al., 2018; Zhang et al., 2020; Rose et al., 2017) and the inclusion of these terms may have led to a more comprehensive search.

The aim of the review was to evaluate effectiveness. As the use of XR in wellbeing is a relatively new field, the majority of the included studies were RCTs that took place under controlled laboratory settings. This provided limited information on real world use and effectiveness, and rather provided more information on potential efficacy. The quality of the included studies was assessed using a risk of bias tool that was appropriate for the evaluation of RCTs; however, as lab-based studies occurred under more controlled circumstances than other studies and often lacked follow up assessments, they were more likely to obtain better quality scores compared to studies that took place in more naturalistic settings. Whilst methodological issues in these less controlled studies affect their analyses and conclusions, these studies may be more similar to real world applications of these interventions and a closer representation of effectiveness.

The present review used the Cochrane Risk of Bias Tool to assess the quality of the included RCTs. The appropriate version of the tool was applied depending on the design of the study being assessed (individually randomized parallel groups, crossover randomized or cluster randomized) in line with tool usage guidelines. As the different versions of the tool slightly differ in the domains they assess, the overall quality assessment for the review had to be divided into different sections based on study designs rather than be presented in one table. Various other

quality assessment tools exist in the literature and tools such as the Quality Assessment Tool for Quantitative Studies (Thomas et al., 2004) or the Jadad Scale (Jadad et al., 1996) could have allowed for all included studies to be evaluated using the exact same criteria, potentially making direct comparisons across studies easier. Nevertheless, the decision was made to proceed with the Cochrane Risk of Bias Tool for the following reasons; 1) The possibility of using different versions of the tool specifically designed for the study type being assessed was an advantage, as it allowed for the assessment criteria to form a perfect fit with the trial designs, rendering the assessment questions more applicable and relevant. 2) Different versions of the tool were able to capture important nuances about study designs and therefore, were more informative than using a general broader scale, where details such as carryover effects (which significantly contributed to the high risk rating for the Crosswell and Yun (2020) study) could have been lost in the assessment or unclear to the reader. 3) Although different versions of the Cochrane Risk of Bias tool were designed to be used for different study designs, these versions only differed on one marking criteria from another (addition of Period and Carryover Effects for the crossover trials or Timings of Identification or Recruitment of Participants for cluster trials), meaning that direct comparisons could be made on other individual marking criteria across all included studies. Furthermore, all versions of the tool used the same quality labels (low, some or high), making interpretation of results across all versions easy to follow. 4) The Cochrane Risk of Bias Tool is significantly more commonly used compared to other tools in systematic reviews of RCTs, possibly due to it being recommended in institutional guidelines, like the Cochrane Handbook (Farrah et al., 2019). Therefore, using the same tool that other researchers may already be familiar with, reduces heterogeneity in the field and improves accessibility of the findings.

The heterogeneity in the included studies contributed to several limitations of the present review. There were only two studies using AR technology and the rest of the studies used VR interventions, meaning that the findings of this review are skewed towards the effects of VR interventions. Within the VR interventions, a range of devices were used to create virtual or immersive environments. These are likely to differ in their usability, create different user experiences and lead to varying levels of immersiveness (Bird, 2020). These technological differences between devices were not explored or taken into consideration in the analyses, as these were often not commented on in the included studies and there was significant variation within the small number of studies included that prevented meaningful groupings. It is possible that technological differences could have affected the delivery of the interventions and influenced their effectiveness. Similarly, heterogeneity was observed in the number of outcome measures used and the differences in the content covered within those measures, as well as in the content of the interventions themselves, which made comparisons between studies challenging.

Effect sizes of the present review were calculated using variables available to the researcher. For within groups effect sizes, Morris and DeShon (2002) recommended correcting for dependence among means, so that comparisons to between groups effect sizes could be made. Since correlations between means in within groups samples were not available to the researcher, Cohen's (1988) method was used to calculate effect sizes instead, and direct comparisons of within and between groups effect sizes were avoided.

Implications for Future Research

XR technology is a relatively new and developing field, and devices using XR are likely to become more widely available to the general public in the upcoming years with costs decreasing and technologies becoming more user-friendly (Davies and Bergin, 2021). XR

interventions can be applicable in schools, outpatient and inpatient care settings as well as be used by individuals as self-help tools to aid in the treatment or prevention of mental health difficulties. This can be particularly helpful for individuals who may not have access to or may struggle to engage with traditional psychological support (talking therapies, self-help resources). The small number of studies that fit the inclusion criteria for this review is indicative of the need for more extensive research in this field.

XR interventions that have been developed according to therapeutic principles to target psychological disorders have been reliably proven to have effectiveness across a range of psychiatric conditions (Park et al., 2019). This review has highlighted that existing research in the field of XR use in wellbeing for adolescents and young people has mostly utilized non-psychological content in the interventions (e.g. exercising, drawing, watching scenic videos). Further research in this field can investigate whether therapeutic interventions specifically designed to improve wellbeing have an added benefit on mental health of participants in clinical and general populations. This line of research could focus on experimenting with different psychologically based intervention content, as well as practical aspects around its use to improve its effectiveness as treatment or prevention with regards to when, where and how these interventions are best placed to be used.

The present review did not identify any qualitative studies on the effectiveness of XR interventions on adolescent and young people's wellbeing. Considering the subjective nature of wellbeing, qualitative research can improve validity, reliability and generalizability of findings in this field (Ma et al., 2015). Specifically, this can include aiding in the selection of appropriate wellbeing outcome measures for XR interventions, as well as capture more nuanced individual

experiences to shed light on why some studies did not yield positive results and the changes that can be made to address these issues.

Future research should focus on methodological improvements, such as increasing the number of participants in studies to achieve sufficient power, lengthen intervention duration and trial duration to examine long term effects and include more diverse demographic samples. Non-XR wellbeing interventions for adolescents have been delivered in studies that have addressed these methodological concerns (e.g. N = 2308, resilience-based intervention duration = five months (Leventhal et al., 2015); N = 508, physical activity-based intervention duration = 10 weeks (Smith et al., 2018a)); however, similar methodological rigour is lacking in the field of XR wellbeing interventions for adolescents and young people at present. Application of interventions in more diverse settings in clinical and non-clinical contexts can provide information on real world performances of these interventions, highlight any barriers to their use and effectiveness and inform clinical decision making.

Conclusion

The present review found that there are some positive effects of XR interventions on wellbeing of adolescents and young people; however, the findings are mixed overall and methodological issues exist in the literature, such as a lack of comprehensive assessments of wellbeing over extended periods of time. Whether XR interventions lead to clinically meaningful changes or have lasting impacts on adolescent and young people's wellbeing is unknown and further research is needed, with one potential area of research focusing on theoretically driven and psychologically informed interventions.

Part III: The Empirical Study

Non-Randomized Controlled Trial of an Augmented Reality Board Game on Adolescent Wellbeing

Abstract

Promoting wellbeing plays a protective role in mental health of adolescents and there is a need for mental health support for this age group. Schools, which are ideal settings for community wide preventative interventions due to their reach, have been struggling to meet the increasing mental health needs of their students. Although previously developed wellbeing interventions exist, engagement with such interventions and their effectiveness could be enhanced through gamification and the use of XR technology. However, wellbeing interventions that use XR technology and are based on psychological theories and therapeutic approaches are lacking. A new wellbeing intervention was developed that aims to teach wellbeing skills to adolescents through completing challenges while playing an augmented reality board game. The intervention was tested for its effectiveness and engagement benefits in a school setting with participants between the ages of 13 and 17. The study followed a longitudinal non-randomized controlled design and assessed wellbeing outcomes pre- and post-intervention and at one month follow up. The results showed that there were no differences in self-reported wellbeing of adolescents between those who participated in the intervention and those in the control group, who continued to attend their regular classes as usual. Participants were less motivated to attend the intervention than attending regular classes. Limitations and implications for theory and practice will be discussed.

Introduction

Wellbeing is a multidimensional construct consisting of hedonic and eudaimonic concepts of happiness, satisfaction, self-realization and positive functioning and quality of life (Ryan & Deci, 2000). Although factors positively associated with psychological distress are negatively associated with wellbeing (e.g. having low income), wellbeing and mental distress/mental health problems are not opposite ends of the same continuum (Winefield et al., 2012), with wellbeing representing more than just absence of distress (Bech et al., 2003). Aside from mental health, wellbeing impacts on various areas of life including physical health, employment, interpersonal relationships and educational attainment (Department of Health, 2014). Wellbeing is particularly important in adolescence, which is considered to be a critical developmental period due to the significant amount of changes individuals go through in this period and the lasting impact these changes can have if the challenges they bring exceed the coping abilities of individuals (Department of Health, 2013b). Due to the importance of wellbeing in this period, several wellbeing interventions have previously been developed with a large majority of them taking place in schools and utilizing face to face delivery methods (Kuosmanen et al., 2019). Whilst these school-based interventions have generally shown effectiveness, challenges have been highlighted in the delivery of these interventions (e.g. around fidelity) (van Agteren et al., 2021). XR technologies have utility in mental health research as seen by treatment studies (Halldorsson et al., 2021), can contribute to the development of novel wellbeing interventions and provide added benefits via increasing engagement and potentially amplifying intervention content; however, available XR wellbeing interventions for adolescents have reported mixed results (as seen by the systematic review). The literature on XR wellbeing

interventions is a very recent and rapidly developing. The effectiveness of these interventions can potentially be enhanced by combining them with other evidence based methods.

The context of mental health and wellbeing of adolescents within schools will be reviewed to pinpoint areas of need that new interventions can target. The role of technology within the educational system will be discussed with a particular focus on use and benefits of XR as a teaching delivery method. Evidence based psychological approaches and gamification will then be suggested as possible ways of enhancing XR wellbeing interventions to improve their effectiveness. Finally, the novel wellbeing game intervention that combines these approaches will be introduced.

Mental Health and Wellbeing in Schools

Adolescence is a critical developmental period for mental health and wellbeing, as young people experience neurological changes affecting their cognitions and behaviors (Steinberg, 2005). These changes can often be difficult for adolescents, who may struggle to adapt and lack the coping skills to do so. First onset of mental illnesses commonly (62.5%) occurs in adolescence with the peak age of onset occurring at age 14 (Solmi et al., 2022) and suicide is the leading cause of death in young people in the UK (Office for National Statistics (ONS), 2014). Aside from its links to mental health, children and adolescents' wellbeing also impacts on their cognitive development and educational attainment (Durlak et al., 2011; Public Health England, 2014), as well as their physical health and social life in adulthood (Department of Health, 2013a; National Institute for Health and Care Excellence [NICE], 2008; NICE, 2009).

Providing adolescents with knowledge and skills to improve their wellbeing is important during this period, as it can promote better mental health, prevent psychological problems and improve their quality of life overall. School settings are ideal and well-placed for preventative

wellbeing interventions, as students attend school full time and spend significant amounts of time at school during important developmental ages of childhood and adolescence (Šouláková et al., 2019). This is recognized by the NHS Long Term Plan (2019), which aims to embed mental health support for children and adolescents in schools, building on existing support available and focusing on prevention and early intervention. Improving wellbeing of students and taking steps to prevent mental health problems can also reduce behavioral problems in schools, increase school attendance and build resilience and confidence (Department for Education, 2018a).

School support around mental health and wellbeing of students has been a particular need over the past couple of years due to the effects of the coronavirus pandemic. UK-wide surveys revealed that adolescents' mental health deteriorated over the course of the pandemic, with the reports showing increases in levels of experienced stress, worry, depression and loneliness, and decreased life satisfaction (Pascual-Sanchez et al., 2020; Department for Education, 2022). Students who had pre-existing mental health conditions, had lower socioeconomic status or were females experienced disproportionately negative impact of the coronavirus pandemic on their wellbeing (Public Health England, 2021a). These reports highlighted the need for more mental health support in schools, including more lessons and resources on mental health. The reports also made recommendations for co-produced wellbeing campaigns for children and adolescents to equip young people to be able to better support themselves, prioritizing prevention and early intervention in schools (Young minds report, 2020). However, a survey of UK secondary school teachers found that nearly half of respondents did not feel confident in addressing the mental health needs of their students (Early Intervention Foundation, 2021) and 65% of schools reported that there is a need for teaching materials for delivery of PSHE content (Department for Education, 2018b), indicating that the existing systems in place may not be well equipped to

meet the current mental health needs of students, further emphasizing a need for programmes to support teachers in the delivery of mental health content.

Use of Technology in Education

One way of improving the delivery of mental health content in schools is through the use of technology. Technology progresses rapidly and children are becoming exposed to it at much younger ages. As exposure to technology is likely to influence how children learn, the argument was made that the use of technology should be a key component when designing educational curriculums to meet student needs (Bruno, 2019). Since the coronavirus pandemic and the shift from face to face teaching to home schooling, schools have adapted to using blended approaches, supplemented by an increasing use of technology in teaching of core subjects. Building on this, similar approaches could be applied to education around wellbeing and mental health.

Learning not only depends on the content of the information but also on how that information is acquired (Drummond et al., 2017). Intrinsic motivation to engage in learning tasks in schools can be fostered when teachers encourage autonomy in classes by supporting and empowering students to take initiative, explore and come up with solutions for problems (Niemiec & Ryan, 2009). One way of achieving is through the use of XR by allowing students to navigate digital environments. Using augmented reality and digital story telling has been shown to promote autonomy, creativity and self-directed learning, with the aim of improving engagement with educational material (such as history content) in classrooms (Rammos & Bratitsis, 2019). Augmented reality games can also increase willingness to engage with educational material by promoting socialization through increased interaction with others in game play (Prithwijit et al., 2018). One study found that even though students attending XR supplemented classes learned just as well as students attending traditional classes, they reported

higher engagement and more positive emotions when learning, suggesting that using XR in education can be beneficial by increasing interest in learning (Allcoat et al., 2021).

There is growing evidence that the use of immersive XR tools (such as VR learning environments) can increase student engagement and improve learning of content (Cobb et al., 2009; Lee et al., 2010; Webster, 2016; O'Brien & Toms, 2008; Snelling, 2016; Prithwijit et al., 2018) and there is an increasing focus on incorporating technological tools into teaching in secondary schools for these reasons (Santos Garduno et al., 2021). Digital interventions can be accessible and can provide additional support in schools alongside what is already offered without requiring additional staff time or input. Whilst the general consensus is that integrating technology in education would be beneficial, limitations in the literature exist around generalizability of findings due to costs associated with technology use and tech-savviness of school staff impacting on their willingness and ability to implement such tools (Al Farsi et al., 2021). Furthermore, XR technology's effectiveness in improving learning content in relation to wellbeing was only tested in one intervention (included in the systematic review) (Hadley et al., 2019), which specifically focused on emotion regulation and reported mixed findings, suggesting that further research is needed.

To build on this existing literature on the use of XR in wellbeing promotion in schools and to enhance these interventions, it may be beneficial to refer to literature on 1) psychological approaches to adolescent mental health and wellbeing, and 2) the impact of games/gamification on mental health.

Evidence Based Psychological Interventions in Adolescent Mental Health and Wellbeing

Mental health interventions can be grouped into three broad categories of prevention, treatment and continuing care (Mrazek & Haggerty, 1994). According to Mrazek and Haggerty's

model (1994), prevention interventions include three levels, universal, selective and indicated depending on how broad or specific to at risk target population the interventions are. School based programs that do not target specific at risk student groups, but rather the entire cohort, fall under universal interventions and aim to provide population level education around mental health. As previously outlined in the systematic review, these interventions are often informed by psychological approaches, such as ACT or positive psychology; however, to date, there have been no interventions that have attempted to combine the evidence base around psychologically informed universal interventions and XR technology, to our knowledge. This is a gap in the literature around XR use in mental health, as the effectiveness of these wellbeing interventions could potentially be improved when they are guided by evidence based psychological approaches used in adolescent mental health.

Psychological approaches that are commonly used in the adolescent mental health context are cognitive behavioral therapy (CBT), systemic therapy and narrative therapy. These three approaches have established effectiveness in the treatment of mental health problems in adolescents and can each make unique contributions to interventions around prevention of mental health problems and promotion of wellbeing.

Systemic therapy focuses on interpersonal interactions within groups and aims to create change within the system as a whole rather than focusing on a single individual. Systemic interventions have been found to be effective for children and adolescents, with studies showing small to medium effect sizes for interventions with adolescents experiencing a range of mental health disorders (Cottrell & Boston, 2002; Riedinger et al., 2017). As positive relationships with others plays an important role in wellbeing of adolescents (Mertika et al., 2020), systemic

interventions focusing on interpersonal aspects are important to consider in the development of interventions targeting wellbeing.

CBT, on the other hand, is often focused on the individual and aims to create positive changes in thoughts and behaviors to address mental health difficulties. CBT techniques can be adapted to developmental stages of adolescents (Holmbeck et al., 2006) and interventions using CBT have been shown to be effective in significantly reducing symptoms of mental health disorders in controlled trials (Klein et al., 2007; Waldron et al., 2004; Dalle Grave et al., 2013). CBT has been shown to be compatible with technological delivery formats and gamification (has been integrated into a computer game format for adolescents for the treatment of mental health disorders) and has been found to be acceptable and effective in this context, enhancing motivation for therapy and therapeutic alliance (Brezinka, 2014; Coyle et al., 2011).

Narrative therapy addresses appraisals of mental health problems and aims to create change through focusing on areas, such as language and internalization. Narrative therapy for adolescents can lead to improvements in disorder specific outcomes (as well as CBT on some measures) (Lopes et al., 2014). Narrative therapy interventions can aid the development of emotional and interpersonal skills to support present functioning and to build resilience (Gudiño et al., 2017). One longitudinal controlled narrative intervention using storytelling and creativity to promote wellbeing in a primary school, found that the intervention led to lasting increases in wellbeing and decreases in depression, anxiety and somatization (Ruini et al., 2020). The study found that narrative techniques helped students identify personal resources and assimilate the concept of eudaimonic wellbeing.

Although the systematic review did not identify any studies combining psychologically informed approaches with XR technology for wellbeing, similar approaches were adopted for

treatment of mental health disorders. Using techniques from CBT, systemic therapy and narrative therapy can enhance interventions in a theoretically driven way.

Games, Gamification, Mental Health and Wellbeing

Aside from using psychologically enhanced content in XR interventions targeting wellbeing of adolescents, one further way of improving the effectiveness of new interventions in this field could be through the use of games and gamification. The systematic review identified some studies that used games as their interventions; however these were not purposefully developed to target wellbeing. Several of the studies in the systematic review were also single session interventions, where the authors reflected on the possibility of reduced effects over repeated administrations of the interventions. One way of maintaining engagement is through gamification, for which there is extensive research on its applications in mental health.

Over the years, digital games have continuously increased in popularity and the global gaming industry is now a multibillion dollar business, appealing to more diverse groups of individuals every year (Accenture, 2021). As a result of this, there is growing interest in the use of digital technologies and gamification in the field of mental health and wellbeing across various participant demographics. Games that are developed for specific purposes other than purely for entertainment are called serious games, which include subgenres such as exergames (games that promote health by requiring players to be physically active and exercise while playing the game).

Game based interventions have multiple benefits. Serious games in mental health can reach populations that wouldn't usually access mental health content and can keep users motivated and engaged (Fleming et al., 2017), increasing the accessibility of interventions using these strategies. Digital gamification has been incorporated into interventions around skills

development in children and young people, including teaching facial recognition skills to children with autism (Tanaka et al., 2010) and teaching problem solving skills and psychoeducation to children with diabetes, improving their treatment adherence and clinical outcomes (DeShazo et al., 2010). These studies have shown that the use of serious games is compatible with mental health content and can facilitate delivery of such interventions.

Another benefit of games on wellbeing is in relation to the social component. Loneliness has been found to be one of the main factors contributing to emotional skills deficits in a sample of British adolescents (Wols et al., 2015) and has been identified as a mental health area of concern overall. Research has shown that the social interaction element of video games can provide significant psychosocial benefits for its players (Granic et al., 2014) and that children and adolescents enjoy playing multiplayer serious games due to the social interactions these games provide (Chin et al., 2008). Similar findings were observed in an AR intervention, Pokemon Go, where adolescents reported that they preferred to play with each other rather than alone and experienced an improvement in their relationships with their peers through the game play, compared to those who did not play the game (Hsieh & Chen, 2019). Considering the role social interactions play in maintaining wellbeing of adolescents (Orben et al., 2020), using gamification in wellbeing interventions may enhance wellbeing through the process of gameplay (via increasing interpersonal interactions with peers), in addition to the content of the interventions.

Previous studies have found positive effects of serious games on mental health. Gaming and exergaming interventions have been found to have a positive impact on sense of self and wellbeing (Joronen et al., 2016) and to increase satisfaction with life, positive emotions and sense of emotional balance (Azevedo et al., 2014). One systematic review (Fleming et al., 2017)

on serious games for mental health found that therapeutic principles can be applied to games and that games may lead to therapeutic change even without being based on traditional therapies (e.g. Tetris for PTSD). Previous studies indicate good potential for use of game based interventions in mental health, but research into this is in its early stages (Fleming et al., 2017; Shah, 2018). Whilst it is believed that gaming interventions may be more attractive and engaging that non-gaming interventions, further research is needed for trials longitudinally comparing game based and non-game based interventions for mental health (Fleming et al., 2017). The game based XR interventions identified in the systematic review were not serious games and were not psychologically driven in their approach, suggesting an untapped potential for these interventions. As an additional layer to engagement benefits provided by XR interventions, which might wear off as the novelty of the technology decreases over repeated use, gamification in combination with XR can promote prolonged and recurrent use and potentially longer term engagement with interventions (Lindner et al., 2019). Therefore, psychologically driven XR wellbeing interventions can potentially be made more engaging and effective through combining these with gamification.

The Wellbeing Game- Dragons of Afterlands

AR games, based on evidence-based psychological principles, have the potential to provide psychoeducation and aid in the development of knowledge and skills to improve wellbeing within an immersive format. Immersive games are likely to increase engagement with the content of the game, leading to potentially increased learning and longer lasting effects. Previous studies have pointed out the need for research on the ecological validity of XR interventions for children and adolescents and for studies to focus on co-produced, user centered

products, using study designs that prioritize user needs to optimize uptake, efficacy and cost effectiveness (Parsons et al., 2017; Badawy et al., 2017).

With the aim of addressing this need, we designed an intervention to improve wellbeing in adolescents using a newly-developed board game, grounded in well-established psychological models and theory (CBT, narrative and systemic approaches). The wellbeing game Dragons of Afterlands aims to improve wellbeing through various wellbeing tasks. It is novel, as it combines social elements of a multiplayer board game with AR technology to create an immersive experience to improve engagement with the wellbeing tasks (see Intervention Content and Intervention Materials and Technology sections under Methods for more details). To our knowledge, no study has tested the impact of psychologically informed AR games on youth wellbeing.

Aims

The purpose of the present study is to investigate if Dragons of Afterlands is successful in improving wellbeing in adolescents in a school setting and how it compares to usual teacher-delivered wellbeing education in schools in terms of engagement and effectiveness. The aims and hypotheses of the present trial were the following:

Aim 1: a) To test whether participating in the Dragons of Afterlands intervention improves self-reported wellbeing of adolescents compared to those attending regular classes over the course of the intervention, and

b) whether any effects of the wellbeing intervention can be maintained over the one month follow up period.

Hypothesis 1: Adolescents, who play the AR game, will report higher wellbeing on standardised wellbeing outcome measures at the end of intervention and at the one month follow up assessment compared to those, who attend regular classes.

Aim 2: To test whether Dragons of Afterlands can improve engagement and motivation of students towards classes that cover wellbeing content.

Hypothesis 2: Adolescents, who play the AR game, will report higher engagement with the process (score higher on the motivation outcome measure), compared to those, who attend regular classes.

Methods

Design

The study was a non-randomized controlled trial following a longitudinal pre-post experimental design. A non-randomized design was selected following requests from participating schools (e.g. due to staff availability). The AR wellbeing game intervention took place during class hour and lasted five to six weeks, which was followed by a one month follow up period. The control group attended regular classes as usual.

Power Analysis and Sample Size Calculation

As there were no previous studies looking at the effectiveness of psychologically informed AR interventions on adolescent wellbeing, the power analyses for the calculation of sample size could not be guided by effect sizes in previous literature. The studies included in the systematic review on VR interventions to improve wellbeing in adolescents found small to medium effects, where effects were found and reported. Of the studies included in the systematic review, the closest studies in similarity to the present study in terms of the use of AR and

longitudinal design (Ruiz Ariza et al., 2018 and Hsieh & Chen, 2019) did not find any effects on wellbeing; although the interventions they used were solely game based and not psychologically informed. Literature on non-XR psychological interventions to improve wellbeing show low to moderate effect sizes (van Agteren et al., 2021). In light of these findings, we opted to base our analyses on a medium effect size. Sample size was determined based on a-priori power analyses, which showed that 86 participants are needed in total (alpha=0.05) to have sufficient power (0.8) to detect medium effect sizes (Cohen, 1988). The required sample size was determined using the G-Power software.

Recruitment

333 secondary schools in London were contacted via emails addressed to SENCOs and headteachers (see Appendices F and G). The study was also advertised on a UK-wide teachers Facebook group and on the South East Research Network for Schools (SERNS) website and bulletin. Following advertisement, thirteen schools expressed interest, two of which (one state funded academy in London and one private international school in Surrey) proceeded with their participation in the study. Drop out reasons for other schools included high staff turnover in schools, difficulties in organizing practical components of participation, pressures on schools during the pandemic and lack of willingness to commit to longitudinal interventions.

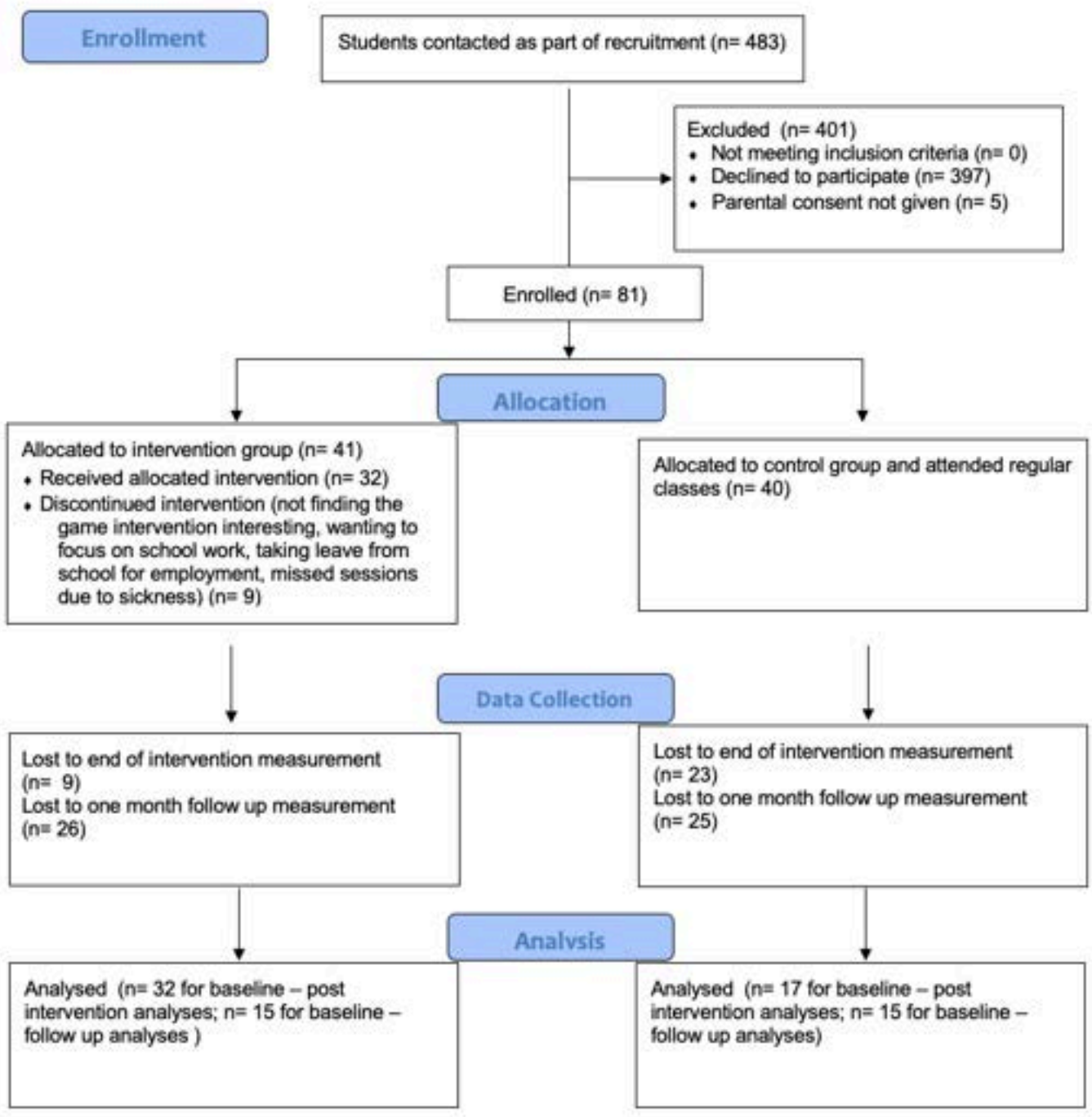
Following confirmation of participation and approval from schools, lead teachers facilitating the project identified students who were approached to take part in the trial (n = 483, control and intervention groups combined). In the state funded academy, participants for the intervention group were selected by the lead teacher from various classes across year groups. The lead teacher was advised on the importance on selecting a representative sample and aimed to do so by basing their selection on internal reports of students. A random selection method was

offered to the school whereby the researcher would attend to randomly select students from a list using a random number generator; however, this offer was declined by the school. The remaining students in the year groups (those not selected to participate in the intervention group) were invited to participate in the control group. In the private international school, all of the students who were attending the lead teacher's class were asked to participate in the intervention group and students attending a separate class (similar in lesson content) were asked to take part as the control group. Recruitment and retention of participants was outlined in the Consort Flow Diagram in Figure 1. Parents of identified students were contacted by school staff with information and consent forms, and had the opportunity to opt their child out of participating in the trial by filling out the form and returning it to their teacher (Appendix H). Students, who were not opted out by their parents, were then provided with Qualtrics links, containing participant information and consent forms, which they were asked to sign if they wished to enroll into the trial (Appendix I). A shopping voucher prize draw was used to incentivize participation. Five parents withdrew their children from the study. Of the students contacted, 396 declined to participate. While students were not obliged to provide reasons for not wishing to participate, some reasons reported were pressures of academic work at school and wanting to attend regular classes, not finding the game intervention interesting and not wishing to fill out questionnaires.

Students who were excluded, withdrawn or those who chose not to participate in the trial or dropped out after enrolment continued to do their usual coursework, either by attending regular classes with the control group students or by independently working on their coursework in the classroom or at the library under the supervision of a school teacher.

Figure 1.

Consort Flow Diagram



Participants

Eighty one students were recruited to participate in the study (n= 41 in the intervention group and n= 40 in the control group). Forty six of these eighty one participants were recruited from the private school and thirty six were recruited from the state school. Following drop out (63% overall), 30 remained in the study at follow up (n= 15 in the intervention group and n= 15 in the control group). Participants were between the ages of 13 and 17. Exclusion criteria for participation were blindness and lack of proficiency in written and spoken English, as the game included visual and auditory components in English language.

Intervention Content

Dragons of Afterlands was selected to be tested in the present trial as it is the first intervention of its kind to combine XR, gaming and psychological content, all of which individually have been shown to positively impact wellbeing (as outlined in the Introduction), opening up the possibility of further increased intervention effects while being fun and engaging for young people. The game intervention was developed using cognitive behavioral, systemic and narrative approaches, and is presented as an AR board game to create a multisensory immersive serious game experience. The wellbeing challenges in the game are based on Personal, Social, Health and Economic education (PSHE) teaching on wellbeing intended to address the four domains of wellbeing in adolescents; physical, social, cognitive and emotional, as defined by Lippman et al. (2011) and in line with the WHO conceptualization of wellbeing (WHO, 2004b). The content of the vignette challenges are outlined in Appendix E. The topics and skills covered in the game intervention included cognitive topics, such as developing an awareness of one's thoughts and how these impact on views of self and influence behaviors, developing alternative thinking styles and becoming more flexible with thoughts. Emotion

related topics focused on identifying emotions and developing language around them as well as strategies of handling difficult emotions (like anxiety or low mood). Social topics included reflecting on challenging situations (like facing criticism or bullying), the emotions these bring up and developing ways of coping with these situations (like assertiveness) as well as improving communication skills overall by reflecting on different communication styles. Finally, physical topics covered factors that contribute to maintaining physical health (like nutrition, exercise) as well as developing an awareness around addictions. Through completing the challenges, participants learn about wellbeing and apply these knowledge and skills in the form of storytelling and roleplay with other participants in the game. Throughout the game, players work with other players, promoting self-reflection (e.g. trying to understand what their avatars may be experiencing in the vignettes by reflecting on their own similar experiences), communication skills (e.g. by trying to explain challenge prompts to each other succinctly and effectively), problem solving skills (e.g. coming up with solutions to dilemmas presented in the vignettes by weighing up various options presented), collaborative working skills (e.g. challenge taker and ally working together to problem solve vignettes) and empathy. The intervention relies on Self Determination Theory principles of autonomy, competence and relatedness (Ryan & Deci, 2000) to motivate players to continue engaging with the content and developing their skills, whereby players are active participants in the learning process (rather than passive listeners in a lecture), can build on and evaluate their skillset through completing challenges and receiving feedback on these and work jointly with others on common goals and difficulties.

Psychological skills training has been shown to lead to improved wellbeing in various contexts with young people (Golby & Wood, 2016; Steyn et al., 2016) and have been shown to increase resilience (van Agteren et al., 2018). The present intervention was designed to target

core skills identified by the educational system (PSHE curriculum) and by the main therapeutic interventions for this age group. The challenge vignettes provide psychoeducation around coping strategies, which include introducing ideas such as behavioral activation. Research has shown that education around coping strategies is important, as utilizing helpful coping strategies, like problem focused coping, leads to better wellbeing whereas others do not (Mayordomo-Rodriguez et al., 2015). The game presents vignettes that allow for problems to be deconstructed into more easily understandable and solvable problems. These vignettes are followed by multiple different perspectives on the scenario, which get discussed between participants with the aim of solving the challenges. The presentation of multiple perspectives is meant to help participants become aware of thinking errors and engage in thought challenging and cognitive reframing around common life struggles. The cognitive flexibility that these skills bring has been shown to have a positive relationship with wellbeing of individuals (Polat et al., 2022) and predicts the use of coping strategies when faced with difficult situations (Asici & Sari, 2021). Working collaboratively on challenges on interpersonal scenarios, participants are intended to become aware of unhelpful patterns of relating to one another and the circularity of effects in these scenarios. Through completing challenges via avatars and immersing themselves in the fantasy world of Dragons of Afterlands via the AR elements and the storytelling, participants are able to externalize problems, gain new perspectives on them and trial different ways of handling situations without having to face scrutiny or stigma about their own mental health. This is based on narrative theories (Madigan, 2011) that propose that individuals make sense of themselves and their difficulties through the stories they (and society) create around them and therefore, engaging in the same storytelling process is a natural and effective way of re-authoring these narratives where needed.

Intervention Materials and Technology

A newly developed AR game called Dragons of Afterlands was used as the wellbeing game intervention in the present study. The game is played through the interaction between the smart phone app and the physical materials of the game, which include a game board, challenge taker and ally cards, monster cards, tokens and dices (Figure 2).

Figure 2.

The Game Materials

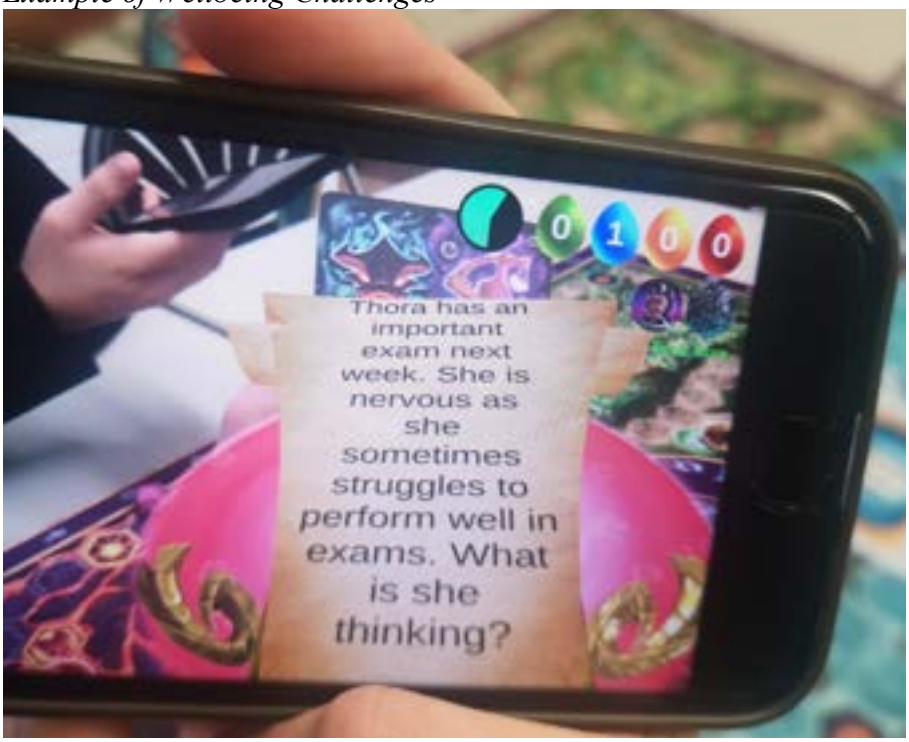
Note. Materials include game board, challenge taker and ally cards, monster card, tokens and dices.

The game allows players to personalize their experience by selecting their own avatars and starting terrain points. Wellbeing challenges (Figures 3 and 4) get triggered when players land on specific tiles, which prompt players to work through vignettes with their allies (other players), using verbal clues and acting skills. The game provides feedback and reinforces correct

ways of problem solving vignettes by giving points. AR components of the game become visible when players hold their phones towards the board (Figures 5, 6 and 7).

Figure 3

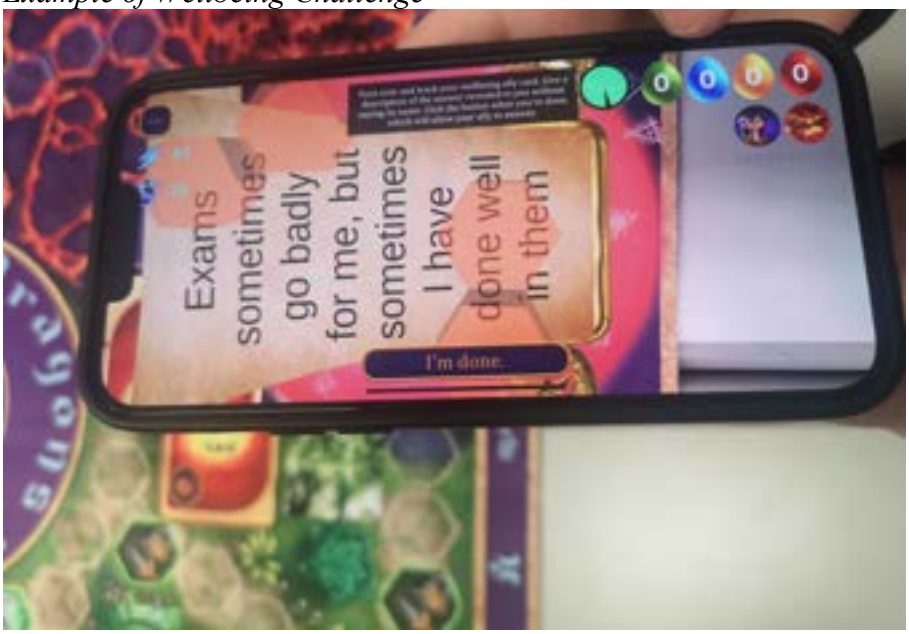
Example of Wellbeing Challenges



Note. One of the challenge vignettes.

Figure 4.

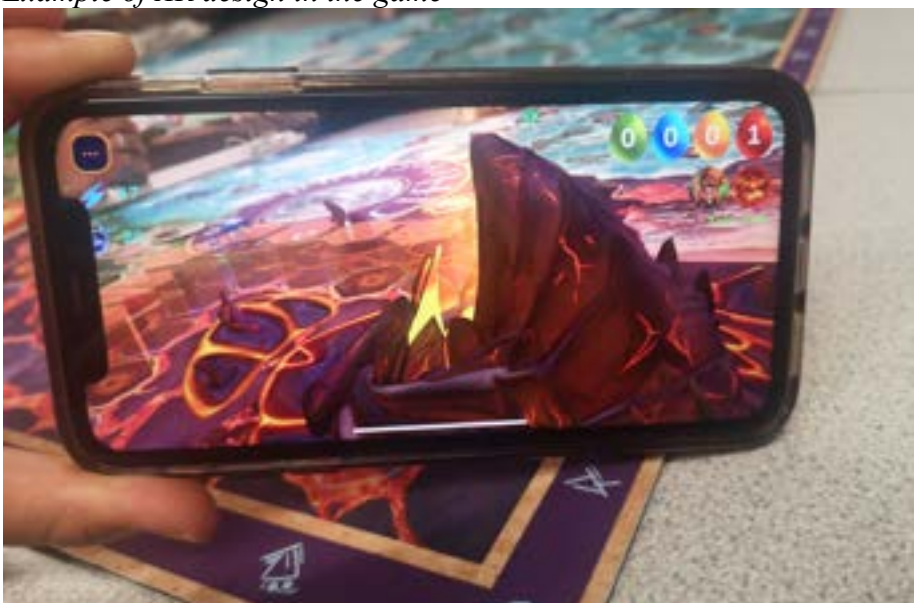
Example of Wellbeing Challenge



Note. Ally prompt of the challenge vignette. This promotes discussions between challenge taker and ally, facilitates joint problem solving and suggests an alternative thinking style.

Figure 5

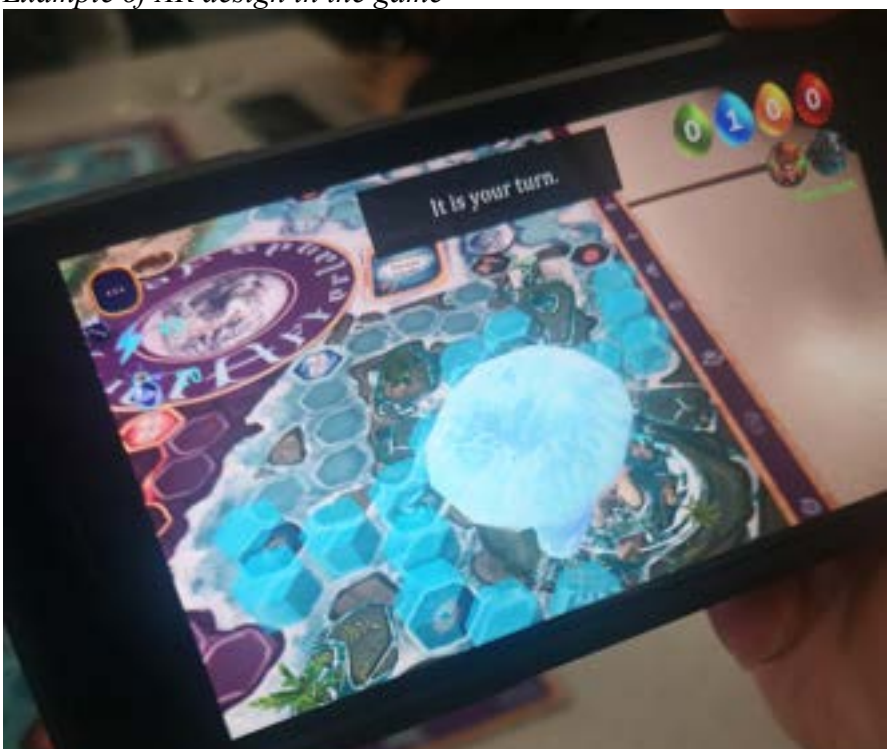
Example of AR design in the game



Note. AR terrain of the Volcano.

Figure 6

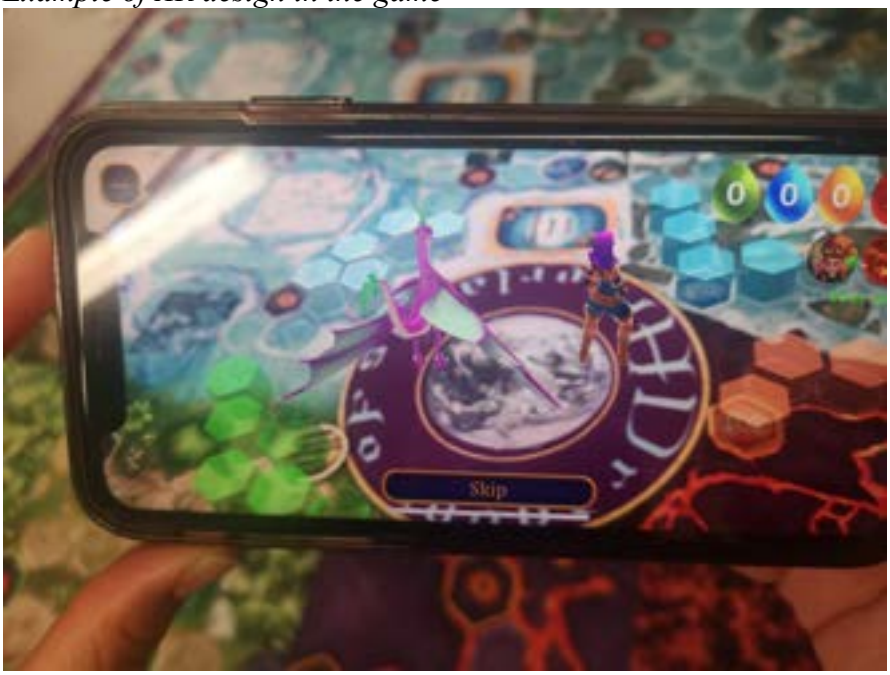
Example of AR design in the game



Note. AR terrain of the Whirlpool.

Figure 7

Example of AR design in the game



Note. AR character, the Dragonwitch.

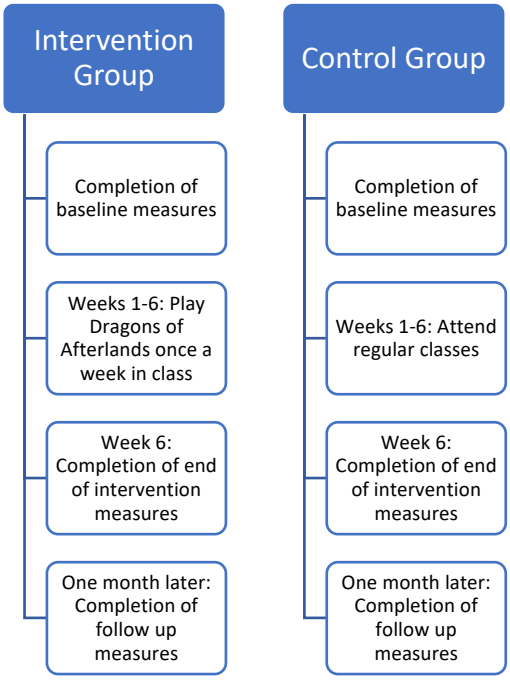
Procedure

Trial Timeline

The trial lasted 10-11 weeks and took place between September 2021 and April 2022. Baseline questionnaires and consent forms were collected one week prior to the start of the intervention. The intervention lasted five to six weeks (depending on school term length to minimize disruptions over term breaks). During this time, the intervention group attended weekly sessions to play the wellbeing game during their class hour in groups of three to four players. The control group continued to attend their regular classes. End of intervention questionnaires were completed at the end of this period and follow up questionnaires were completed one month after (Figure 8).

Figure 8

Trial Timeline



Intervention Group

The study design accommodated for structural differences in both schools to ensure recruitment. In the state school, a randomly selected group of students across three year groups (selected by the school) was allocated to the intervention group, with volunteers from the remaining students from those year groups serving as controls. In the private school, two classes from the same year group were selected by the school to participate in this project, with one class serving as the intervention group and the other as the control group.

Protocol for Gameplay

Participants allocated to the intervention group played the game at school during class hour. Instructions on how to download and install the game app were circulated a week before the start of the intervention and students were instructed to do so prior to the first intervention session.

Participants played the game in small groups consisting of three to four participants. Participants complete challenges in the game by landing their tokens on challenge tiles, selecting one ally from the rest of the players (ally selected only for the duration of the challenge and can be switched for the next challenge). Challenge takers are then asked a multiple choice question and the allies are given prompts (which they act out or explain to the challenge taker) to promote thinking.

The classrooms where the game play took place were set up at the beginning of each session to allow for participants to be seated with their groups. Necessary materials to play the game (e.g. game boards) were provided. In cases where students did not have access to smart phones or could not play on their own devices (e.g. due to compatibility issues with the

technology) phones were provided by the research team. A WiFi dongle was provided by the research team to resolve any issues around internet connection with school WiFis.

In the first intervention session, participants were guided through the app registration process to create their accounts. They were then given written and verbal instructions on how to play the wellbeing game and were instructed to start playing. Participants had the opportunity to ask questions or raise difficulties with the researcher throughout the game play.

The engagement of students with the wellbeing game was ensured and monitored by the researcher. The researcher and a class teacher were present during the game sessions and oversaw the game play. A register was kept by the school teacher to record attendance of participating students.

Data Collection

Data collection took place at three time points; baseline (collected following allocation and one week prior to the start of the intervention), end of intervention and one month follow up. Baseline and end of intervention questionnaires were completed in class for the intervention group, with the latter taking place immediately following the last game session. The control group received the questionnaire Qualtrics links via email in the same weeks as the intervention group and were given a week to complete the measures. Participants were contacted one month after the end of the intervention via email and asked to complete the follow up questionnaires. Email prompts were sent to maximize data completeness.

Outcome Measures

Data on demographics, including age, school year group, gender, ethnicity, history of mental health problems, history of interventions for mental health problems, long term physical illnesses, socioeconomic status (via obtaining information on parental education level and

occupation) were collected at baseline for descriptive purposes. These demographic variables were selected, as they were indicated as protective or risk factors for child and adolescent wellbeing according to an England-wide survey (NHS Digital, 2017). Office for National Statistics demographics questions were used as a guide to ensure that the categorizations were appropriate and meaningful.

Participants were asked to complete the Strengths and Difficulties questionnaire (SDQ) as a baseline measure to characterize the sample (see Appendix B). Strengths and Difficulties Questionnaire (SDQ) is a commonly used self-report emotional and behavioral screening questionnaire for children between the ages of 11 and 17. It is used for identifying internalizing and externalizing problems in children and adolescents. This measure was selected as a baseline measure to examine potential differences between groups at baseline in relation to mental health problems, as these may influence the effectiveness of the wellbeing intervention and may need to be controlled for between groups. The SDQ includes 25 items rated on a three point Likert scale (Appendix). The total difficulties score ranges from zero to 40, with 0-15 classified as *normal*, 16-19 as *borderline* and 20-40 as *abnormal* for the self-report UK version. It is a well-established questionnaire that has been shown to have good interrater reliability (scales ranging $0.51 < r < 0.64$), acceptable cross informant consistency ($r > 0.30$), acceptable test-retest reliability (self-report scales ranging $0.60 < r < 0.73$), acceptable specificity ($\geq 70\%$) and acceptable negative predictive value ($\geq 70\%$) (sensitivity and positive predictive value deemed unacceptable) according to a recent systematic review (Bergström & Baviskar, 2021).

Participants completed wellbeing outcome measures at baseline, end of intervention and at one month follow up. The wellbeing questionnaires used in the present study as main outcome variables were the following:

- WHO-5 is a measure of overall subjective wellbeing, covering aspects of physical and mental health, developed by the World Health Organisation (1998). It can be used with children over the age of nine and consists of five items that participants rate based on their frequency on a six point Likert scale (Appendix). The raw score is calculated by totaling the figures of the five answers. The raw score ranges from zero to 25, zero representing worst possible and 25 representing best possible quality of life (scores can be multiplied by four for percentages). The questionnaire has been shown to have high reliability cross culturally (ρ = .83-.93) (Philipp et al., 2020) and good construct validity in adolescent samples (correlation with Beck Depression Inventory (BDI-6) r = -0.49 (p=0.0001), coefficient of homogeneity = 0.52, unidimensionality p > 0.05) (Blom et al., 2012). It was deemed to be an appropriate tool to measure wellbeing over time and between groups in research trials by a comprehensive systematic review (Topp et al., 2015) (see Appendix A).
- Warwick Edinburgh Mental Wellbeing Scale (WEMWBS) is a positive mental wellbeing scale consisting of 14 items rated on a five point Likert scale (Appendix). WEMWBS includes affective-emotional aspects, cognitive and psychological functioning and can be used with individuals over the age of 11. A higher score on WEMWBS indicates better mental well-being, reflecting more positive thoughts, behaviors and feelings. WEMWBS has been shown to have high internal consistency (Cronbach's alpha = 0.87, 95% CI [0.85; 0.88]), acceptable test-retest reliability (Intraclass correlation coefficient (ICC) 0.66 (95% CI [0.59; 0.72] n = 212)) and construct validity (r = 0.59 for psychological wellbeing domain of the Kidscreen-27 (95% CI [0.55; 0.62]); r = 0.65 for the Mental Health Continuum Short Form (MHC-SF) 95% CI [0.62; 0.69]; r = -0.44 for the SDQ 95% CI [-0.49; -0.40]) in a UK based sample of adolescent school

students (Clark et al., 2011). WEMWBS shows a normal distribution in the general population avoiding floor and ceiling effects (Clarke et al., 2011; Tennant et al., 2007) (see Appendix D).

This combination of measures was selected due to their fit with intervention content and their comprehensive assessment of wellbeing, indicating that any improvements are likely to be meaningful. WHO-5 and WEMWBS have been deemed appropriate for positive psychology studies due to their strengths focused language and have been recommended to be used in studies with this population (Taggart & Stewart-Brown, 2019). Both of these wellbeing scales combined allow for an assessment of optimal wellbeing, covering hedonic and eudaimonic aspects of wellbeing, as well as subjective appraisals on the four domains of wellbeing in adolescents (physical, social, cognitive, emotional) that the intervention aims to address.

Participants in the intervention and the control group also completed the Intrinsic Motivation Inventory (IMI) (Ryan, 1982) at end of intervention as a measure of their engagement. Participants rate their subjective experience of participating in an activity using a seven point Likert scale and IMI scales can be customized depending on the area of study. IMI has been shown to have good internal consistency and adequate reliability ($\alpha = 0.85$) and validity (convergent, discriminant and face validity) (McAuley et al., 1987; Ostrow & Heffernan, 2018). For the present study, the value/usefulness and the interest/enjoyment scales were used. Items from both scales were presented to participants in alternating order and the wording was adapted to the research topic (e.g. “I think that doing this activity is useful for….(fill in the blank)” was presented as “I think that doing this activity is useful for my wellbeing.”) (see Appendix C). The interest/enjoyment scale is considered to be the primary measure of intrinsic motivation and the value/usefulness scale has been associated with internalization (Deci et al., 1994). Together, these scales aim to measure the degree to which participants find the tasks rewarding to engage

with and can adopt and integrate the concepts they cover. Furthermore, the IMI was selected to understand the willingness of participants to play the game and whether they feel it meets their needs to make inferences about the acceptability and usefulness of the intervention.

Ethical Approval

The study received ethical approval from Royal Holloway University Research Ethics Committee on 19 April 2021 (REC ProjectID: 2566). Applications for minor amendments were made to collect email addresses for the distribution of prize draw vouchers and to extend participant sample group by age to include older students. These were approved on 20 September 2021 and 11 October 2021 (see Appendices J-M).

Service User Involvement

Two participants were selected from the intervention group to provide feedback on their experience of the game. These youth advisors were selected from the group of participants, who volunteered to provide feedback at the end of the trial, and the selection was made using an online random name generator. Participants were verbally informed about their participation (topics that will be covered, information around time commitments and use of data) and provided verbal informed consent. Parental consent was obtained via consent forms sent to parents of selected students (Appendix N). The feedback covered topics, such as the content of the game, the acceptability and usability of the technology, as well as their opinions on the research design, such as playing the game in class hour, duration of the trial and any factors that they believed might have impacted on the analysis of the results. The selected participants were consulted regarding efficient ways of disseminating the findings within their age groups and schools. These feedback interviews were conducted with co-production principles in mind, to empower young people to have a say in interventions aimed at them and influence various stages of research

development, which has been highlighted as an area of importance (Durose et al., 2012; Norton, 2021; Shamrova & Cummings, 2017). Through these interviews we were able to gain insights into their experiences and reflections of taking part in this project to aid in the improvement of the product and interpretation of our findings, as well as identify key areas for our population of interest, so that the research conducted can be of value and meaning to our stakeholders.

Informal feedback was gathered throughout the study, including at recruitment meetings with teachers and during game sessions from students. These were passed onto the developers of the game or future improvements.

Data Analyses

Data was analyzed using IBM SPSS Statistics version 25. Intervention group participants were considered to have completed the intervention if they attended at least four sessions of game play. Those who completed less than four sessions were excluded from the dataset. Descriptive statistics were calculated for demographic variables and SDQ outcomes at baseline using χ^2 tests and independent samples t-tests to determine and address any significant differences for demographic variables between intervention and control groups at baseline and to assess the need to control for these differences in future analyses. Differences between intervention and control groups on motivation scores were assessed using an independent samples t-test. Separate two-way mixed ANOVAs were used to analyze differences between intervention and control groups in relation to wellbeing outcome measures (WHO-5 and WEMWBS) at baseline, end of intervention and follow up assessments. Change scores were calculated for the wellbeing outcome measures to aid in the interpretation of the findings.

Results

Missing Data

At the end of the trial, 30 participants remained in the study, meaning that the statistical analyses proposed were underpowered based on a priori power analyses. As previously mentioned (see Methods), multiple avenues of recruitment were explored to recruit the proposed sample size, including contacting 333 secondary schools in London multiple times and advertising on social media and on research networks. The recruitment period was extended past the proposed timeline to allow for more schools to participate. Unfortunately, it was not possible to recruit sufficient schools within the timeframe of this doctorate. 13 schools expressed interest and all of these schools were given opportunities to discuss their participation. 11 of these 13 schools were not able to participate eventually due to organizational and covid related factors (e.g. staff turnover (including staff who are interested in participating leaving their posts), reluctance from staff and management due to pressures on delivering curriculums, lack of interest in longitudinal interventions, lack of classroom availability due to spaces being used for vaccinations). As there were no remaining potential schools to recruit from at the end of the extended recruitment period, decision was made to end recruitment even though participant numbers were low.

Although recruitment was limited to only two schools, we were able to enrol 81 students from these schools into the trial, which meant that the study was initially only underpowered by five participants. Unfortunately, a significant number of participants dropped out of the trial, which was caused by a variety of factors, including staff level challenges (e.g. facilitating teacher leaving her post and loss of contact with some participants due to this, difficulties in getting staff on board to facilitate control group data collection in class hour to increase participation and

ensure data completeness), participant level challenges (e.g. increasing academic pressures following covid and students wanting to prioritize their curriculum work, students not finding the game interesting) and technical challenges (e.g. initial difficulties with WiFi access in one school and glitches in the game app). These challenges were informative with regards to gathering real world application data, however led to difficulties in progressing the study with a sufficient sample size.

To manage missing data following high numbers of drop out, Little's Test of Missing Completely at Random (MCAR) (Little, 1988) was conducted to assess if data was missing completely at random and missingness was not dependent on observable or unobservable variables. The results were non-significant ($\chi^2(7) = 4.41$, $p = 0.73$) indicating that the data missing from the end of the intervention and follow up assessments happened randomly. Therefore, the results of the statistical analyses were unlikely to be biased by missing data. Listwise deletion of missing cases (where participants with missing data are removed from the statistical analyses) was applied, as this is the recommended method for unbiased estimates and conservative results when MCAR assumptions are satisfied (Kang, 2013). Imputations were not considered for missing data, as aside from MCAR assumptions being met, there were large amounts of missing data (39% loss between end of intervention and follow up), which could have led to erroneous results after imputations (Jakobsen et al., 2017). Instead, analyses were conducted on complete cases and data missingness was noted as a limitation of the findings and interpretations. Participants who only provided data at baseline were removed from the analyses. Forty nine participants had complete data at end of intervention and thirty participants had complete data at follow up.

Demographics

Participant demographics were outlined in Table 1 and Figures 9, 10, 11 and 12. Descriptive analyses indicated that participants in the intervention and control groups did not differ in their composition on demographic variables in terms of participant age ($t(47) = 1.63$, $p = 0.11$), school site ($\chi^2(1) = 0.12$, $p = 0.73$), school year ($.\chi^2(3) = 5.54$, $p = 0.14$), ethnicity $\chi^2(4) = 3.71$, $p = 0.45$), gender ($\chi^2(3) = 1.19$, $p = 0.76$), history of mental health diagnoses ($\chi^2(5) = 3.39$, $p = 0.64$) or history of mental health treatment ($\chi^2(5) = 4.97$, $p = 0.42$), current long term physical illness ($\chi^2(5) = 4.66$, $p = 0.46$), parental qualification ($\chi^2(3) = 1.72$, $p = 0.63$) or profession of highest earning parent ($\chi^2(5) = 6.00$, $p = 0.31$).

Participants were asked to complete SDQ at baseline as a descriptive measure of their mental health at the start of the trial. Participants in the intervention and control groups did not differ in their overall mental health as measured by their total SDQ scores ($t(47) = -0.45$, $p = 0.66$) and their SDQ categories (normal, borderline or abnormal) ($\chi^2(2) = 0.77$, $p = 0.68$). As such, there were no demographic or baseline factors identified that could have led to differences between groups in relation to their wellbeing.

Table 1

Demographics of the Participant Sample (n=49)

Demographics	Control Group (n = 17)	Intervention Group (n = 32)	Statistics[a]
	Mean (SD)		
Age	16.2 (1.2)	15.6 (1.5)	p = 0.11
	Number of Participants		
School site (Private school)	13	23	p = 0.73
School Year (Year 12)	13	23	p = 0.14
Gender (Female)	12	20	p = 0.76
Ethnicity (White)	10	13	p = 0.45
History of Mental Health Diagnoses			p = 0.64
Anxiety	2	1	
Depression	1	2	
Autism	0	1	
History of Mental Health Treatments			p = 0.42
Antidepressants	1	1	
Talking therapies	2	3	
Current long term physical illnesses			p = 0.46
Epilepsy	0	1	
Hip dysplasia	1	0	
Scoliosis	0	2	
Highest Parental Qualification (At least one has degree level qualification)	16	27	p = 0.63
Profession of Highest Earning Parent (Senior managerial or administrative role)	8	12	p = 0.31

[a] Statistics shows p-values of χ^2 tests or independent samples t-tests, depending on data properties.

Figure 9

Gender Distribution in the Sample

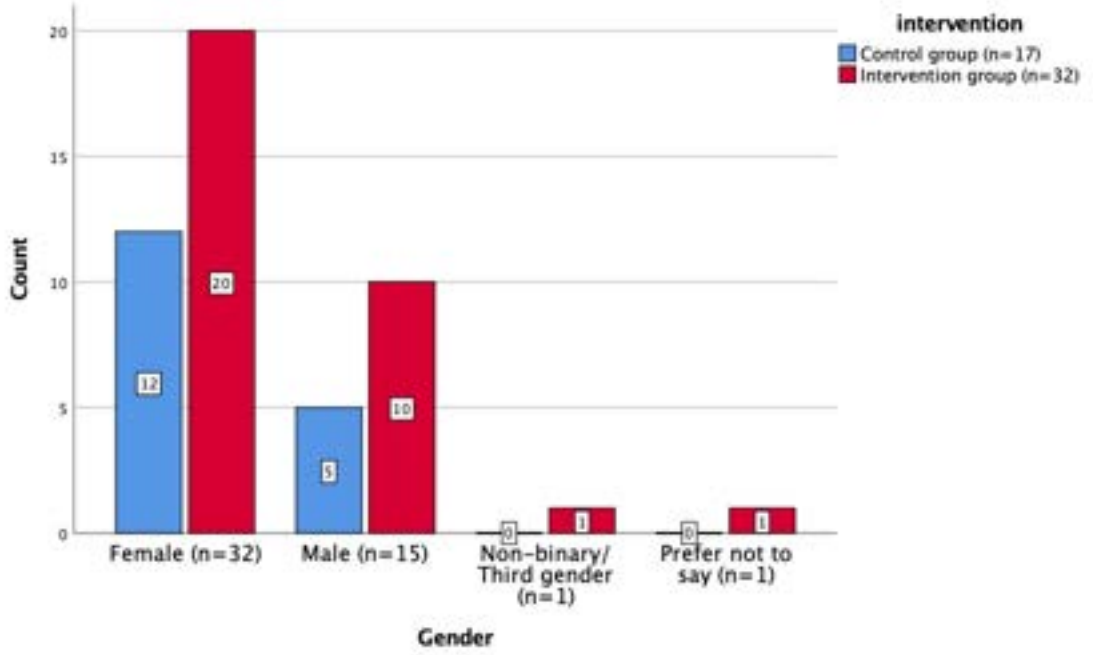


Figure 10

School Year Group Distribution in the Sample

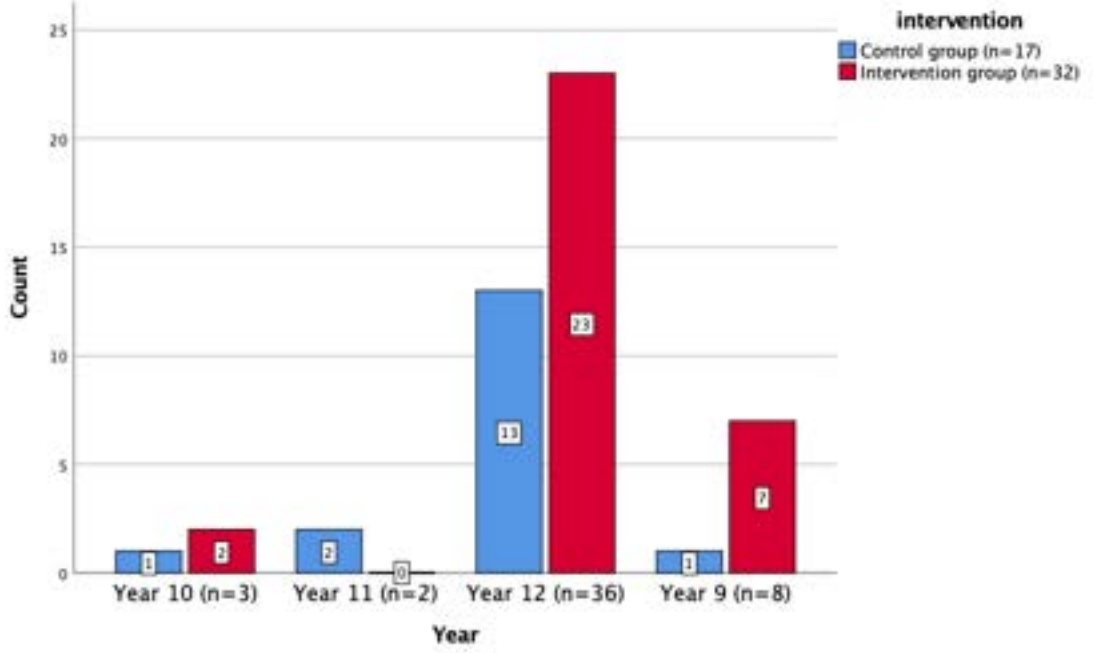


Figure 11

Parental Profession Distribution in the Sample

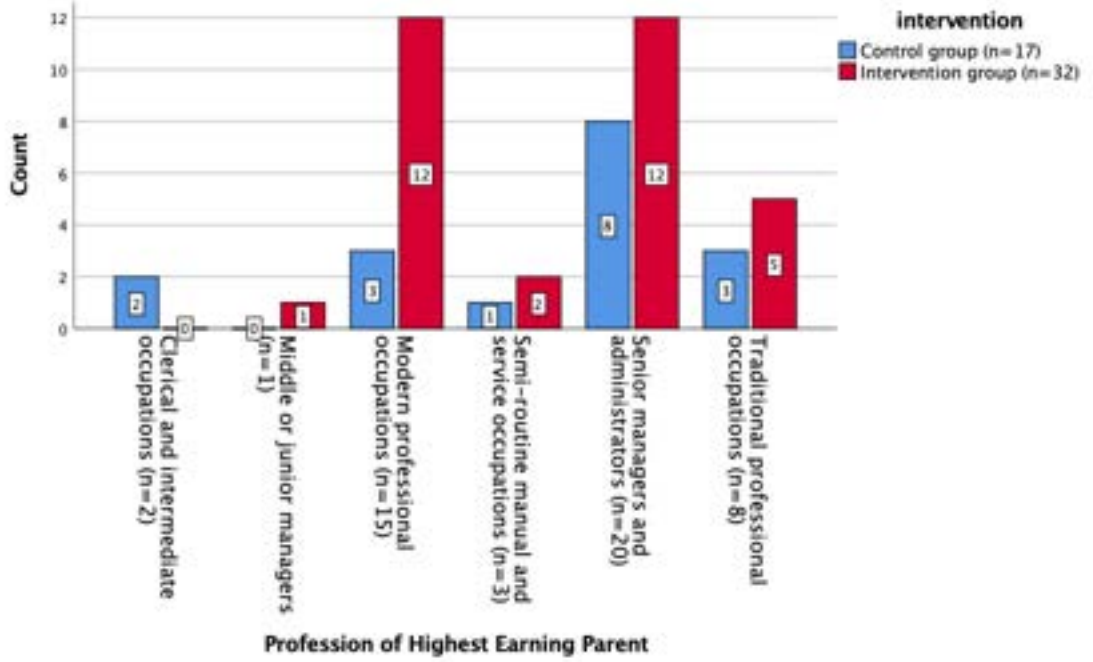
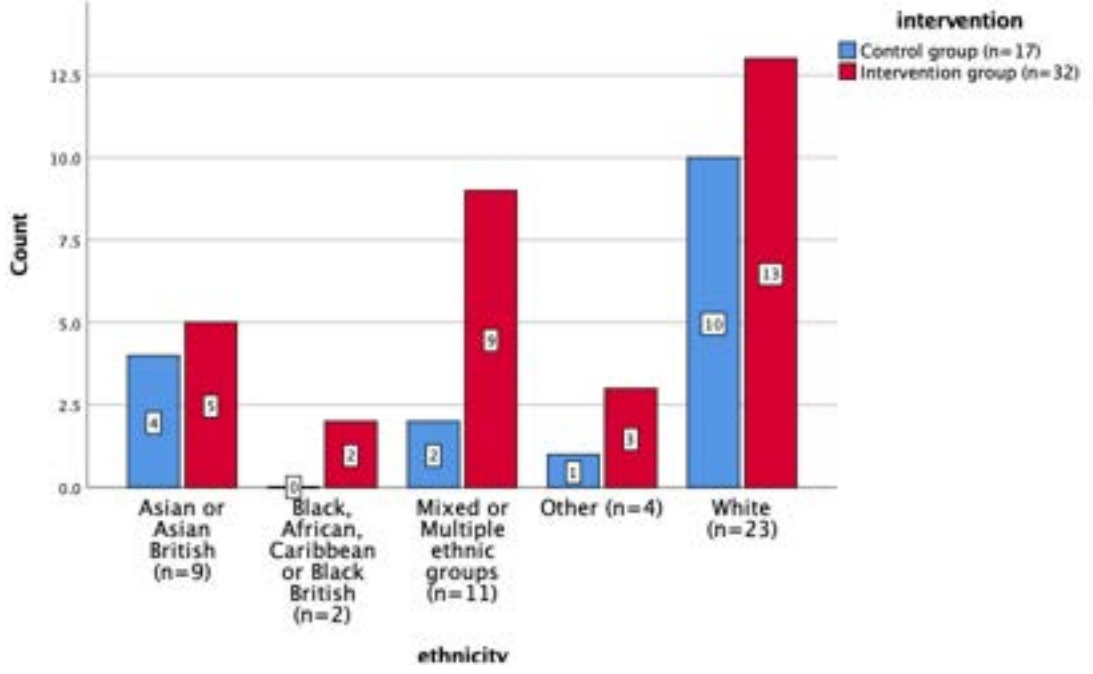


Figure 12

Ethnicity Distribution in the Sample



Assumptions and Analyses

The distribution of continuous variables were assessed for normality by plotting the graphs and assessing for skewness and kurtosis. z scores were calculated and all scores were between -1.96 and 1.96 ($p > 0.01$) indicating that the variables were normally distributed (Kim, 2013).

Data were plotted using boxplots to identify outliers. Three outliers were identified in the control group at the follow up measurement for the WEMWBS questionnaire. These values deviated more than three standard deviations from the mean of WEMWBS at follow up for the control group (Dave & Varma, 2014). Analyses were conducted with and without the outliers, which did not lead to differences in outcomes. The decision was made to include these outliers in the reported analyses. No other outliers were identified at baseline or in the intervention group.

Usability and Engagement- Differences in Motivation

Differences between control and intervention groups on how motivated they were to attend their allocated classes (intervention - game sessions or control -regular classes) were assessed to test whether students who played the game in the intervention group were more engaged with the process and more motivated to attend classes compared to the control group. Motivation was assessed and analyzed via the IMI scores collected at the end of the intervention period, which asked participants to rate their experience of attending either intervention or regular classes in relation to interest/enjoyment and value/usefulness of the tasks. Groups were compared using an independent samples t-test. Levene's test indicated that equal variances could be assumed ($F = 1.04$, $p = 0.31$) and the results showed that there were significant differences between groups ($t(47) = 2.50$, $p = 0.02$), with the control group (mean: 68.35, standard deviation: 17.40) reporting higher levels of motivation compared to the intervention group (mean: 53.06,

standard deviation: 21.74). The highest possible score on this measure was 98, with higher scores indicating higher levels of motivation.

Effectiveness of the Wellbeing Intervention- Changes to Wellbeing Outcomes

Changes to wellbeing scores were analyzed to test 1) whether participants who received the wellbeing game intervention reported higher wellbeing compared to the control group at the end of the intervention and 2) whether any improvements in wellbeing were maintained at the follow up assessment. Baseline wellbeing scores of control and intervention groups were compared using independent samples t-tests. There were no significant differences between groups in WHO-5 scores ($t(47) = 0.54$, $p = 0.59$) or WEMWBS scores ($t(47) = 1.26$, $p = 0.21$) at baseline and therefore these were not controlled for in the analyses.

Changes Between Baseline and End of Intervention

To test if the wellbeing game had a positive effect on wellbeing pre- and post-intervention, changes in wellbeing scores on WHO-5 and WEMWBS measures were analyzed over time and across the control and intervention groups using mixed ANOVAs. These analyses were conducted on all 49 participants who had complete data at end of intervention. Wellbeing scores at baseline and end of intervention for these participants were outlined in Table 2.

Table 2

Wellbeing Scores at Baseline and End of Intervention

		Baseline (n = 49)	End of Intervention (n = 49)
Control	WHO-5	14.53 (5.44)	14.29 (4.97)
	WEMWBS	49.71 (9.33)	48.12 (6.85)
Intervention	WHO-5	13.72 (4.81)	13.66 (5.23)
	WEMWBS	46.09 (9.63)	46.81 (9.74)

Note. Values show means and standard deviations.

A mixed ANOVA for WHO-5 scores between baseline and end of intervention across groups revealed that there were no significant main effects for time ($F(1, 47) = 0.05$, $p = 0.82$) or for group ($F(1, 47) = 0.28$, $p = 0.60$) and no significant interaction effects ($F(1, 47) = 0.02$, $p = 0.90$). Similar results were observed in the mixed ANOVA for WEMWBS scores between baseline and end of intervention, where there were no significant main effects for time $F(1, 47) = 0.12$, $p = 0.73$) or for group ($F(1, 47) = 1.00$, $p = 0.32$) or for group x time interaction ($F(1, 47) = 0.83$, $p = 0.37$). These results indicated that attending the wellbeing intervention group did not lead to any significant differences in the wellbeing of participants at the end of the intervention period compared to those attending regular school classes.

Changes Between Baseline and Follow Up

Further analyses were conducted to assess if there were any differences between groups at the follow up assessment. The following analyses were conducted on cases with complete data at follow up (n = 30) after listwise removal of missing data. Wellbeing scores of these participants were outlined in Table 3.

Table 3

Wellbeing Scores at Baseline and Follow up

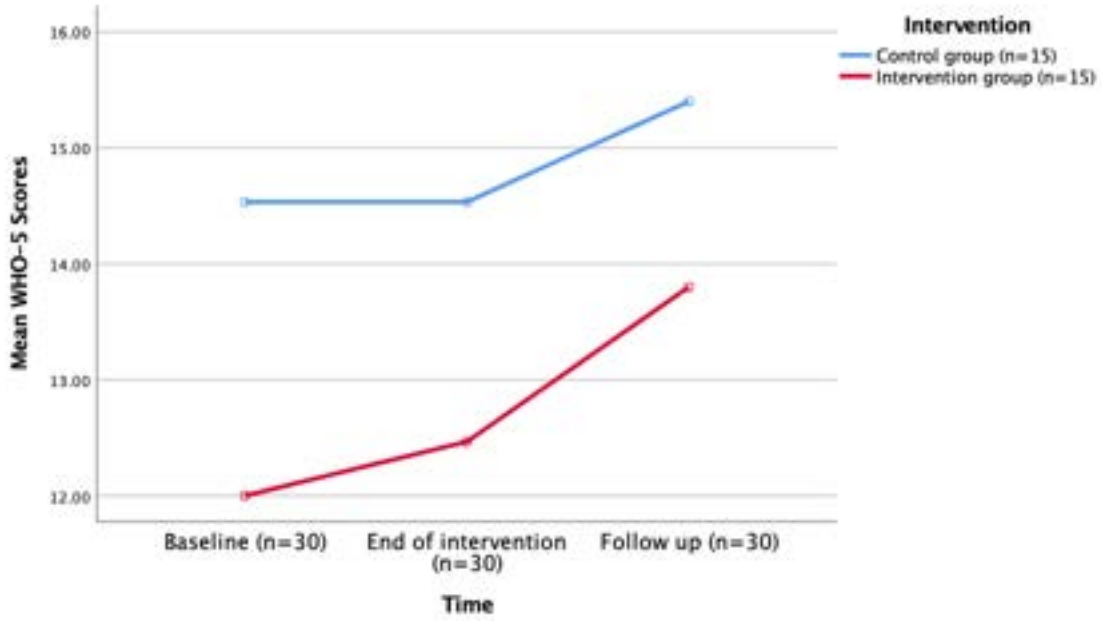
		Baseline (n = 30)	End of Intervention (n = 30)	Follow up (n = 30)
Control	WHO-5	14.53 (1.32)	14.53 (2.32)	15.40 (1.07)
	WEMWBS	49.14 (2.58)	47.79 (1.87)	50.14 (2.17)
Intervention	WHO-5	12.00 (1.32)	12.47 (1.16)	13.80 (1.07)
	WEMWBS	42.47 (2.49)	45.20 (1.81)	45.53 (2.09)

Note. Values show means and standard deviations

A mixed ANOVA was used to examine changes over time on the WHO-5 wellbeing questionnaire across control and intervention groups. Mauchly's Test of Sphericity was significant ($p = 0.04$) and therefore sphericity was not assumed and Huynh-Feldt analyses were taken into consideration. There were no significant main effects for time ($F(1.79, 50.27) = 1.54$, $p = 0.23$) or for group ($F(1, 28) = 2.35$, $p = 0.14$) and no significant interaction effects ($F(1.79, 50.27) = 0.16$, $p = 0.83$). Therefore, participants in the wellbeing game intervention group did not differ from the control group at the end of intervention or at the follow up assessment with regards to their self-reported wellbeing (Figure 13).

Figure 13

WHO-5 Scores Across Time in Intervention and Control Groups

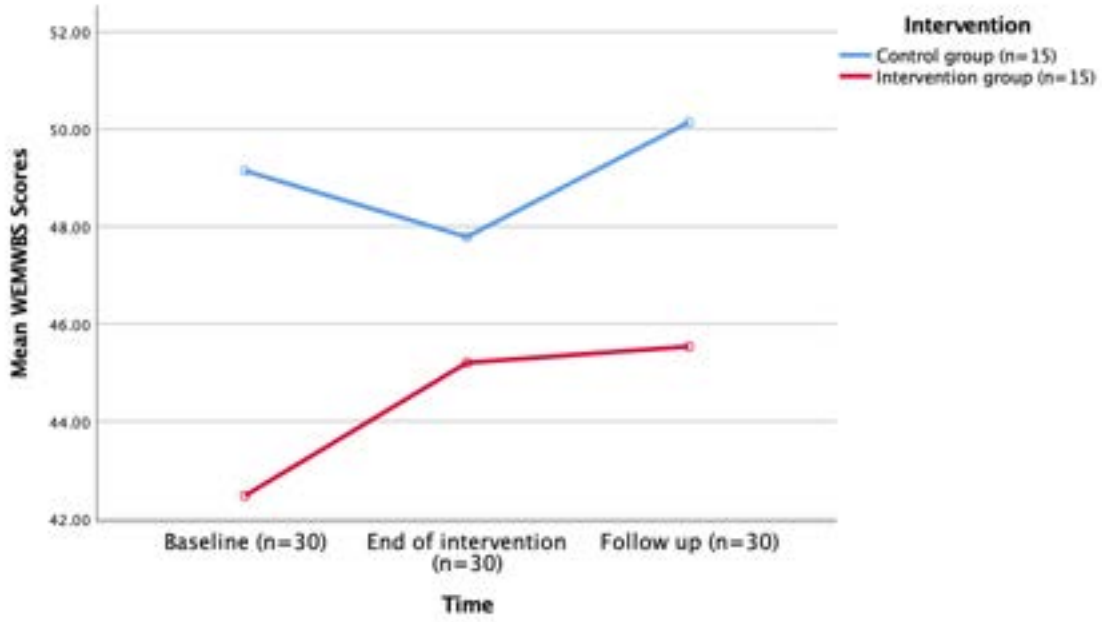


A second mixed ANOVA was used to examine changes over time on the WEMWBS wellbeing questionnaire across control and intervention groups. Mauchly's Test of Sphericity was not significant ($p = 0.20$) and therefore sphericity was assumed. There were no significant

main effects for time (F(2, 54) = 0.83, p = 0.44) or for group (F(1, 27) = 3.79, p = 0.06) and no significant interaction effects (F(2, 54) = 0.81, p = 0.45), indicating that participants in the wellbeing game intervention group and the control group rated their wellbeing similarly at the end of intervention or at the follow up assessment on this wellbeing measure as well (Figure 14).

Figure 14

WEMWBS Scores Across Time in Intervention and Control Groups



Change scores were calculated by subtracting baseline scores from end of intervention scores and follow up scores to allow for a clinical interpretation of the findings. These change scores are depicted in Table 4. Mean WEMWBS scores of the participants in the intervention group increased by 2.73 points by the end of intervention and by 3.07 points between baseline and follow up (compared to a decrease of 1.2 points at end of intervention and an increase of 1 point by follow up in the control group). A difference of three points on this measure is needed to meet the threshold for minimally important level of change (Maheswaran et al., 2012), with

higher scores indicating higher levels of wellbeing. Similarly, WHO-5 scores showed some small positive trends in the intervention group (0.47 point improvement by end of intervention and 1.8 point improvement by follow up) compared to the control group. A 2.5 point difference in the raw score of this measure (equivalent to 10% change) is considered to be a clinically relevant improvement on this measure (Bech et al., 2007), with higher scores indicating higher wellbeing. However, independent samples t-tests conducted on these change scores confirmed previous analyses that these small differences between groups were not statistically significant.

Table 4

Changes in Wellbeing Scores Over Time

Wellbeing Measure	Group	Change between Baseline and End of Intervention	Change between Baseline and Follow up
WHO-5	Control Group	0.0 (4.90) (95% CI [-9.60, 9.60])	0.87 (5.64) (95% CI [-10.18, 11.92])
	Intervention Group	0.47 (4.34) (95% CI [-8.04, 8.98])	1.80 (4.66) (95% CI [-7.33, 10.93])
	Independent samples t-test	p = 0.78	p = 0.63
WEMWBS	Control Group	-1.20 (7.39) (95% CI [-16.28, 12.68])	1.00 (12.34) (95% CI [-23.19, 25.19])
	Intervention Group	2.73 (8.93) (95% CI [-14.77, 20.23])	3.07 (7.05) (95% CI [-10.74, 16.89])
	Independent samples t-test	p = 0.20	p = 0.58

Note. This table includes data from only the participants, who had complete data at follow up. Values show change scores (calculated by subtracting baseline scores from end of intervention or follow up scores), standard deviations and 95% confidence intervals of these change scores. Independent samples t-tests shows differences between intervention and control groups with regards to their change scores.

Discussion

Summary of Findings

The present trial assessed the effectiveness of a newly developed intervention in improving wellbeing and examined whether it improved motivation and engagement compared to regular classes.

Wellbeing scores for those young people who used the game were not significantly different to their peers, who participated in regular classes. This finding was consistent at the end of the intervention and at one month follow up. There were some non-statistically significant trends in the wellbeing scores which suggested that participants using the game intervention may have experienced some small positive changes in their wellbeing over the course of the trial. The changes in WHO-5 scores were larger for the intervention group but still under 2.5 (as a raw score, or equivalent to 10%), which is considered to be the clinically relevant change on this outcome measure (Bech et al., 2007). On the other hand, a three point difference, observed in the intervention group between baseline and follow up, met the lower cut-off for minimally important level of change on the WEMWBS (Maheswaran et al., 2012). These findings should be interpreted with caution due to the small sample size and high rates of attrition and potential biases of students remaining in the trial at the follow up assessment. Overall, it can be concluded

that there was not sufficient evidence to suggest that the intervention reliably improved wellbeing in this cohort of adolescents.

Unexpectedly, participants in the control group reported significantly higher levels of motivation to attend their regular classes than participants attending the intervention classes. The IMI does not utilize clinical cut-offs, which prevented a categorical interpretation of the findings.

Interpretation of Findings on Motivation and Engagement

The findings of this study showing a lack of engagement and motivational benefits of this AR game contradict with general consensus around the effectiveness of AR (Drljević et al., 2022) and VR interventions (Huang et al., 2021) in increasing engagement. Although engagement is perceived as a central argument in favor of digital and XR interventions, evidence around whether technology enhances engagement is limited and mixed due to variations in research design (Fairburn & Patel, 2017). Evidence is often confounded by novelty effects and lack of longitudinal interventions (Martin, 2014). The present trial took a longitudinal approach and measured engagement, as recommended, for digital intervention studies for children and adolescents (Georgeson et al., 2020) to explore this relationship and failed to find positive effects. Whilst this can be due to various trial specific factors, it is also possible that technological interventions are not as unique as they were thought to be for generations that grew up with technology. This is supported by the Hopp & Gangadharbatla (2016) study, which found that higher levels of technological self-efficacy was negatively associated with motivation to engage with AR content. The findings of the present study highlight the need to maintain a critical approach towards potential engagement benefits of XR and indicate the need for more longitudinal research in this area.

Nevertheless, there were also trial specific factors in the present study that could have led to the lack of motivation of participants. Discussions with youth advisors revealed that students felt under pressure to succeed academically due to important upcoming assessments and may have perceived attending regular classes as more useful and valuable towards achieving higher grades. These attitudes may have also been influenced by the impact of the coronavirus pandemic. A recent UK-wide survey revealed that many students felt that they fell behind in their school work during the lockdowns and school closures, felt unprepared for exams and worried about not getting good grades (Young minds report, 2020). Intrinsic as well as extrinsic motivation play roles in maximizing the effectiveness of serious games (Drummond et al., 2017) and it is possible that competing demands reduced the extrinsic motivational aspect.

Youth advisors reported getting a sense that participants enjoyed the game overall and enjoyed playing with their classmates; but noted some discrepancies between the lack of complexity of some of the wellbeing tasks and the participants' developmental stage as a potential contributor to lower usefulness ratings. Using the value/usefulness scale in the IMI was beneficial in highlighting these concerns, which will be important targets of change for future versions of the game. It is possible that the assessment of motivation using only the value/usefulness and interest/enjoyment subscales of the IMI may not have been able to capture motivation as comprehensively within this context and it is possible that including additional scales, such as relatedness, could have provided a more rounded picture of engagement.

Lastly, motivation scores could have been influenced through the self-selection bias in trial participation in the control group, as these students were approached more broadly unlike the intervention group participants who were targeted to participate by their teachers (either cluster randomized based on their classes or individually selected). It is possible that the control

group consisted of more motivated or conscientious students to begin with. Baseline scores were not collected for the IMI as it was designed to assess engagement based on having participated in a specific activity. Nevertheless, it is possible that the lack of motivation and engagement in the game intervention hindered performance and persistence in the wellbeing challenges and prevented the content from being internalized, contributing to the lack of effects on wellbeing outcome measures in the intervention group (Ryan & Deci, 2000).

Interpretation of Findings on Effectiveness

The research literature on wellbeing interventions with adolescents are generally positive in their findings, although there is evidence of mixed effectiveness highlighted in some reviews. One systematic review looking at school based wellbeing interventions (including digital interventions) in a global context, found that 56% of the included trials reported improvement in wellbeing of adolescents and 28% reported deterioration, with mindfulness and positive psychology interventions particularly showing effectiveness (Cilar et al., 2020). However, only 7% of the included studies were deemed to be high quality, further emphasizing the lack of conclusiveness in the literature. Our systematic review looking at the effectiveness of XR wellbeing interventions in adolescents also found mixed results on effectiveness (60% of studies showing some improvement in wellbeing and 40% showing no improvement or decline), with the two AR-based interventions (both using Pokemon Go) not reporting changes in wellbeing. The present findings of the empirical study fit within this context. Although statistically significant differences in wellbeing measures were not found in the present trial, there was evidence of some positive trends in wellbeing scores of the intervention group, which shows promise for future versions of the game intervention.

Interestingly, positive trends in wellbeing scores of the intervention group occurred despite participants reporting feeling less engaged or motivated towards the task compared to the control group. One possible explanation is that the process of playing the game contributed to small improvements in wellbeing by consolidating wellbeing skills. Participants in the study likely had some knowledge around wellbeing (e.g. through the lectures they attended at school around mental health) and the game intervention may have provided them with an opportunity to apply this knowledge in a practical way as they were solving the wellbeing challenges in the game. The alternate reality created in XR interventions may partially lead to a depersonalization of game performances, allowing participants to be more truthful in their answers and explore different strategies without social pressures (Hadley et al., 2019). The storyline, use of avatars and immersive elements would have contributed to this effect, but these could be further developed in future versions of the game by making them more salient and visible. A further benefit of the gameplay process was the social gains of playing a collective game that encourages communication among peers and the sharing of common thoughts and interests. This may have helped by normalizing some difficulties or providing an opportunity to socialize and de-stress by engaging in a non-school work related activity.

Several factors could explain the lack of effectiveness of the game intervention overall. The lack of power is a potential explanation for the non-significant findings, increasing the risk of Type 2 error and reducing the chances of finding true effects where there might be one. A priori power analyses indicated that 86 participants were required for the study to be sufficiently powered, meaning that, due to recruitment difficulties and drop out, the analyses conducted for end of intervention and follow up data were underpowered. Underpowered studies are more likely to lead to biased conclusions, such as incorrectly retaining the null hypothesis and

suggesting that interventions tested were not effective. Although the findings of the statistical analyses in the present study were non-significant, it is possible that these findings may not be reliable and valid representations of treatment effects and larger, sufficiently powered studies may find significant differences pre and post intervention. As unknown biases or sampling errors may also have larger effects on outcomes in smaller samples, it is possible that the findings may have been skewed. It is well known that running underpowered trials is not advised (due to using up participant time without being able to deliver desired outcomes, contributing to replicability issues and creating discrepancies in the field) (Button et al., 2013), however, these studies are common in the field as seen in the systematic review due to reasons outlined previously. As the field of XR wellbeing interventions for adolescents is newly developing, it can be argued that analysis of data from small studies, such as the present trial, can still contribute to the accumulation of knowledge in the field, be meaningfully collated as meta-analyses in the future and may contribute to more precise overall estimates of outcomes by representing a diversity of findings (Turner et al., 2013; Lilford & Stevens, 2002), as long as the results are reported with an explicit reference to statistical power (Abbott, 2014), which we have aimed to convey throughout the report.

There were non-statistically significant differences in baseline scores for wellbeing outcome measures between the control and experimental groups. Although participation was voluntary for all participants, the control group had a more optional approach to participation, as the control group consisted of students, who opted to repeatedly engage with data collection over a period of time at their own will (introducing a self-selection bias), whereas the intervention group was selected by the participating schools, either by individually selecting students or by allocating a whole class for the intervention. Therefore, the control group may have consisted of

individuals with different characteristics (perhaps higher functioning in ways not measured by the scales collected in the trial, more conscientious or responsible students, who felt able and willing to complete multiple questionnaires), who may have had different motivations or attitudes towards participating, whereas participants in the intervention group may have felt more obliged to stay in the trial due to staff involvement around their participation. The control group had slightly higher wellbeing scores on both wellbeing outcome measures compared to the intervention group at baseline and these remained higher throughout the trial. Further possible explanations around this could be that the control group consisted of more agreeable student more likely to score positively, whereas some members of the intervention group may have felt unsure about their participation (e.g. due to stress around academic achievement and competing demands as outlined previously), which may have influenced how they scored on these outcome measures. The analyses that were used aimed to account for these differences by calculating change scores, whereby the focus of the investigation was the improvement or deterioration in wellbeing and the findings were less influenced by baseline scores.

Another possible explanation for the lack of intervention effectiveness is that the novelty of the game wore off by the end of the intervention through repeated sessions of gameplay. One way researchers testing a VR intervention in physical rehabilitation overcame this was introducing more challenging tasks as users became used to the game, which helped them maintain engagement and motivation (Elor et al., 2022). In the present version of the game, the challenges were presented in random order; however, they can be presented with increasing difficulty to maintain engagement. A challenge in developing serious games for populations, such as adolescents, who are exposed to commercial games that have high specs, is developing and maintaining comparable interventions (Fleming et al., 2017), as individuals who are used to

high-tech computer games and visuals and may require more advanced stimuli to capture their attention (Cowan, 2000). It is possible that the game intervention was not as engaging as the games adolescents are used to playing in their spare time. Finally, the game intervention was designed with skills development and learning in mind. It is possible that ceiling effects existed in the sample with regards to wellbeing knowledge, which was not assessed, and that the game therefore could not contribute to further improvements.

Strengths

This study was the first trial of Dragons of Afterlands, which is the first AR serious game intervention that is psychologically informed and specifically designed to target wellbeing in adolescents. As the field of XR technology is rapidly growing, rapid testing of new XR interventions is encouraged to prevent interventions from being outdated or unappealing by the time they are made available to the populations of interest (Fleming et al., 2016). Therefore, the choice was made to test the intervention in schools, in a naturalistic setting to explore ecological validity of the game as well as understand barriers and improvements that need to made to the game to improve its effectiveness. This trial design did not allow for an evaluation of the contribution of the various aspects of the game interventions to our findings; however allowed for us to highlight different areas needing improvement in a more time efficient way (e.g. difficulties in recruitment, sampling from a younger population, improvement on the technical glitches).

Despite difficulties in recruitment and the lack of power associated with the final sample size, being able to recruit for and run a longitudinal trial in a school setting for an intervention that had not been trialed previously was an ambitious goal that was achieved in the present study. There is a need to develop an evidence base for interventions embedded in school contexts, in

line with the NHS Long Term plan (2019), which aims to make mental health support more accessible in schools. Studies, such as the present trial, testing a novel intervention, contribute to the development of this evidence base.

Effective digital health interventions are iteratively longitudinally developed, are a product of collaborative working with stakeholders and professionals across different disciplines and adapt to technological advances (Davies & Bergin, 2021). A strength of the game intervention was the collaborative working aspect. The game itself was a product of interdisciplinary collaborations between professionals from psychology, education and computer science backgrounds to ensure that advances in each field were reflected in the prototype.

Collaborative working with young people was central to the development of the intervention. Meeting young people's needs is an important component of developing new digital interventions and research has highlighted that young people often do not feel that these interventions are relevant to their mental health needs (Halldorsson et al., 2021). Children and adolescents can meaningfully co-produce all aspects of wellbeing research and often find these experiences enjoyable and valuable (Ben-Arieh, 2005; Moltrecht et al., 2021). The present study had two youth advisors recruited from the participating schools, who provided feedback and ideas for all stages of the trial design and interpretation of findings. The consultations with the youth advisors helped shed light on processes within these schools that the researcher would not have been aware of and significantly aided the interpretation of findings and the identification of future directions for research.

Limitations

Participants were not randomized to ensure engagement and retention of participating schools following difficulties around recruitment. Randomization would have strengthened this

study design and improved the reliability of its findings by eliminating accidental biases (including selection biases) and provided a basis to apply the probability theory on (Lim & In, 2019). Selection biases in the present study could have been introduced by the influence of teachers in the decision making on allocations (e.g. they may have selected students who they think might benefit more from the intervention because they struggle more with their mental health, or chosen students who they think are more cooperative and therefore would engage better with the intervention). Self-selection bias could have also been present in the control group, who were asked to sign up to complete questionnaires around their mental health (e.g. those struggling might have not opted to participate). Rationales behind allocation to intervention and control groups in non-randomized trials such as the present one can be varied and unknown and randomization would have reduced any systematic selection biases (Reeves, 2008). A further challenge of non-randomized trials is confounding, where the intervention and control groups may differ on prognostic factors that can lead to differences in intervention effects and increase heterogeneity in the data and treatment effects (Higgins et al., 2022). This can be caused by a number of factors and it is not possible to measure and account for all of these factors in a non-randomized design. The demographic data collected at baseline did not lead to an identification of any observable differences between intervention and control groups; however, due to lack of randomization, it is possible that factors that were not identified at baseline could have been unequally distributed between groups and potentially affected results of the study. One such factor that was not controlled for was existing wellbeing skills and knowledge at baseline. Another factor that was not controlled for and could have potentially influenced motivation to engage with the AR game, was previous gaming experience of participants, which could have had a positive (participants feeling more eager to engage as they already find gaming enjoyable,

are excited to try out new games and are more inclined to learn through gaming) or negative effect (participants being used to higher level computer gaming systems and being underwhelmed in the trial). Future research should aim to take these factors into account.

Whilst trials that take place in real world settings are unavoidably less controlled than laboratory settings, it is important to understand and address barriers to recruitment for future trials of similar interventions. It is possible that attitudes and skepticism in education settings towards game based teaching or interventions prevents uptake of such trials. A further possibility is additional demands put on schools during and following the coronavirus pandemic, which likely contributed to lack of support from school staff around scheduling and planning aspects of the study and retention of participants.

Retention of participants was a limitation of the present study as drop out between baseline and follow up assessment was 63%. Dropout rate was similar in both control and intervention groups and the Little's MCAR (1988) analyses showed that drop out was completely at random. Nevertheless, drop out in the sample meant that the sample size was smaller, lowering the power of the statistical analyses and the conclusions drawn were based on the reports from a subset of the original sample. Students were not obliged to provide rationale for dropping out; however, some reasons for drop out reported in the intervention group were around wishing to attend regular classes to succeed academically. Opinions of participants regarding the game itself likely influenced decisions to drop out as well, which need to be explored further in qualitative studies. Recruitment and retention was particularly challenging in the control group despite efforts to motivate participants with incentives (vouchers). Reasons for these difficulties are unknown, as contact with this group was limited and only over email. Recruitment and retention in the control group could have possibly been improved through more

staff support (e.g. collection of outcome measures in class hour for the control group); however this was challenging to negotiate with schools due to it being perceived as an additional burden.

The exclusion criteria for participation in the trial was kept to minimum to increase applicability and generalizability of findings. Nevertheless, the following limitations with regards to applicability and generalizability should be noted. The participants of the present study were selected from a community sample of students. Whilst this allowed for conclusions to be drawn on the use of the intervention in school settings, the findings are not sufficient to shed light on possible therapeutic applications of the intervention with clinical populations.

The majority of participants in the study (73.5%) were attending Year 12 in a private international school, representing a narrow and specific demographic within the UK population. Children, who come from low income families, are more likely to have mental health problems and worse mental wellbeing (NHS Digital, 2017), indicating a higher need for research and interventions for these populations; however, diversity in socioeconomic status is lacking in research on preventative digital mental health interventions (Bergin et al., 2020). In this regard, the present study is lacking in demographic diversity as well. Children from underprivileged backgrounds may be more likely to benefit from such psychoeducational interventions; however, barriers to research with these populations were evident in the present trial's recruitment and retention process. In addition to this, the majority of the students in the study had had several lectures dedicated to mental health and wellbeing by Year 12, which likely provided them with a better baseline understanding of psychoeducation around wellbeing. This has implications on the generalizability of findings as well as on the size of the effects found, as students without similar levels of wellbeing input may have experienced more benefits. Schools that put more emphasis on wellbeing are more likely to participate in trials on wellbeing, and therefore, may be more

represented in the research literature, potentially skewing findings. Similarly, a younger cohort of students who had less mental health input may have found the intervention content more novel and engaging and potentially benefited more from it. Younger children can navigate and make use of XR interventions on psychoeducation and emotional literacy (Flujas-Contreras et al., 2020), suggesting that the technological components of the present AR game are unlikely to present a barrier to its application in younger populations in future trials.

The game app requires a higher level operating system in smartphones, functional cameras and a reliable connection to the internet. This may not be accessible to all adolescents, particularly those coming from underprivileged backgrounds, who may benefit from wellbeing interventions. Data and smartphones were made available throughout this study to ensure that participation of students was not constrained by this; however, consideration needs to be given to equal access, should this intervention be rolled out on a larger scale.

Although the majority of the participants in this study played the game using smartphones provided by the research team (to overcome barriers to participation in relation to battery levels, compatibility of processing systems with AR), some participants opted to use their own devices. These students' devices were compatible with AR; however, minor differences in performance of the software on different smartphone models could have existed, potentially contributing to some of the glitches experienced during gameplay sessions, and impacting on the level of immersiveness of the game (which could in turn potentially influence intervention effectiveness). These challenges could have been overcome by ensuring that all participants played the game on identical devices that had been previously performance tested by the developers. The use of different devices in this study was helpful in providing information on challenges that would be faced in real world delivery of this intervention, as well as information

for the developers to improve the software to increase its compatibility where needed. Based on researcher observations during delivery of the gameplay, there were no obvious or particularly disruptive differences in individual participants' experiences of the game that could be attributed to the smartphone model used.

As this was the first trial of the intervention, several learning points and limitations were around practical elements of the trial. In retrospect, running an initial feasibility or pilot study could have helped with anticipating and potentially problem solving some of the challenges that occurred in the implementation of the trial. These preliminary trials could have helped identify recruitment issues (and potentially problem solve this by building relationships with schools), understand concerns around randomization and selection of groups, estimate resource needs (like smartphones), develop protocols for the delivery of the intervention in schools, inform selection of participant age groups and understand concerns around acceptability (e.g. competing demands with academic work) and dropouts (Abbott, 2014). Having these data would have informed product development, improved methodological rigor and trial design (Ismail et al., 2018). At the time of the development of the research questions and trial protocol, the longer term impacts of covid on the educational system were unknown and previous trials on digital wellbeing interventions (although not using AR or gaming) in schools had been successfully conducted by other researchers within the research team. Therefore, an effectiveness trial felt feasible to run at the time and the trial was designed with this aim in mind. At baseline, the sample size recruited was nearly sufficiently powered; however, by the end of intervention assessment the impact of organizational level challenges compounded by covid, became apparent and contributed to high dropout rates. As the initial study design, the outcome variables collected and the aim of assessing effectiveness were not aligned with aims of a feasibility or pilot trial (Whitehead et al.,

2014), the decision was made to proceed with the original trial design that the school and participants agreed to take part in. Feasibility and study design related data were collected along the way in addition to inform future trials. Although the initial decision to pursue the present trial design was led by beliefs around the possibility of successfully completing the trial, feasibility or pilot trials are not commonly conducted due to system level issues in the field, such as the short lifespan of startups and the need to establish working products swiftly (particularly in the field of XR with rapidly evolving technologies), the reluctance of journal editors to publish studies labelled as "pilot" (Arain et al., 2010) and the encouragement from journals in the retrospective labelling of articles with inadequate sample sizes, inconclusive findings or uncertainty around generalizability as "pilots" (Shanyinde et al., 2011). This, coupled with the fact that pilots do not guarantee success of future effectiveness trials, particularly in dynamic and constantly evolving situations, such as covid, as well as due to individual school or participant level differences, prevented the utilization of the more conservative route of initial feasibility and pilot testing in the present trial. Overall, this highlights the need for systemic changes in research, as feasibility and pilot trials have the potential to increase efficacy of trials and interventions (Eldridge et al., 2013) and could have led to higher retention of participants and more conclusive and valid findings in the present trial.

Future research

The present trial has highlighted the need for further collaborations with stakeholders with a focus on usability and effectiveness. Convergent mixed methods design, where multiple cycles of qualitative and quantitative data collection and analysis takes place throughout the development of a product, is one of the recommendations to guide innovation in the development of digital health interventions (Alwashmi et al., 2019). A next step for the development of

Dragons of Afterlands could include formal qualitative interviews with adolescents, which can provide valuable information on improving engagement as well as understanding barriers and needs and attitudes towards this topic of research (Eysenbach et al., 2019).

Joint discussions with youth advisors revealed that some areas to explore with stakeholders could include: 1) when the intervention takes place (ad hoc in wellbeing related classes rather than a long term intervention, or made available by the school mentors when needed), 2) which age groups would benefit most from it (considering that children are exposed to technology at much younger ages and receive mental health and wellbeing input in schools as part of PSHE earlier on in their education, and considering added pressures of the education system (e.g. exams) at different stages in school), 3) whether the game could have different utility at different developmental stages (relaxation and distraction for older adolescents and skills development for younger students and adapting research designs to test for these (e.g. pre-post affect assessments for ad hoc use in older adolescents)), 4) improvements that can be made to the delivery and content of the game to improve engagement (addressing technological glitches, making storyline more salient, adapting the complexity and phrasing of the wellbeing tasks to developmental stages of participants).

Additional steps to improve the effectiveness of the intervention can involve in-app data collection on participant use of the app. These were not included in the present analyses due to technological challenges experienced during the implementation of the trial (e.g. app glitches, lack of battery on phones). Future trials, having addressed these issues, can utilize in-app data on success in wellbeing challenges and engagement with the game to explore different trajectories and potential mediators and moderators of change in wellbeing of participants. In-app data can also be used to personalize the game to each player's developmental level and target their

specific needs by customizing the complexity of the wellbeing game challenges based on their entry level scores in the first game play session. This would address issues around lack of engagement due to a mismatch between developmental level of participants and the difficulty of wellbeing challenges as well as increase learning and skills development by presenting students with challenges that are within their capacity to learn. Research has shown that effectiveness and acceptability of wellbeing interventions is heavily dependent on their age appropriateness, perceived helpfulness and engagingness (Francis et al., 2021).

The majority of participants in the trial overall (and in the intervention group, 62.5%) identified as female. Whilst there is evidence that both boys and girls benefit from educational games (Papastergiou, 2009), it was historically found that there were gender differences in gaming preferences, with adolescent boys being more motivated by achievement and drawn to games with action and competition, whereas girls preferring games with logic, puzzle solving or skills training and being motivated by the social interactions these games provide (Romrell, 2013). It was believed that these preferences were the result of a male dominated gaming environment (with boys having significantly more gaming experiences at earlier ages), where sexism and hostility towards girls in certain gaming genres may have created a discrepancy in experiences and preferences (Fox & Tang, 2014). It was also believed that preferences were changeable and based on access, experience and knowledge of gaming (Vermeulen et al., 2011). Recent research has shown that the previous gender differences around gaming motivations may no longer be present as a result of shifts in gaming culture and gender stereotypes over the past decade, making gaming more accessible and socially acceptable to individuals from different genders (Wohn et al., 2020). One study evaluating an educational iPad based AR game found that girls outperformed boys, suggesting that the lack of gaming experience girls may have had

due to reasons, such as cultural norms, may not necessarily pose barriers to their performance (Atwood-Blaine & Huffman, 2017). Although gaming is becoming more prevalent overall in the society, gender and cultural norms are socially construed and may impact students from various sociodemographic backgrounds differently. Accessibility of gaming and the extent and variety of previous experience can also be dependent on other factors, such as affordability. Gender differences in experiences of participating in the game intervention were not explored in the present trial; however, future versions of the trial with qualitative assessments may benefit from investigating this topic to ensure that the game is designed to cater to all genders. There were no gender differences in those who dropped out of the present trial. In the present game intervention, considering historically held beliefs around gender influences, it is possible that boys may have enjoyed the goal-oriented/competitive challenges and the girls may have appreciated the social component with allies. Better understanding differential preferences, if there are any, could lead to design developments whereby more choice could be offered to individual players throughout the game to increase enjoyment and engagement.

The present trial was pragmatic and limited by resource constraints. Further developed versions of the intervention could eventually be tested in a larger, multisite, cluster randomized trial to gather more conclusive evidence of its effectiveness across settings and populations. The present trial and intervention design provided learning points for future larger longitudinal trials. Delivering an intervention that requires two to four players to interact within school lesson time over an extended period poses some challenges, such as ensuring that students remain focused on the game tasks when they may feel tempted to interact with each other. In the present trial, engagement with the intervention was maintained through supervision and prompts from the researcher and facilitating teacher. Supervision by facilitating teachers must be made part of the

game delivery protocol in future deliveries of the game intervention. A further challenge is that same students may not always be present to play the game together over extended periods of time (e.g. due to sickness absence) or may not continue to wish to play with each other (e.g. due to interpersonal conflicts in friendships). The present game was designed so that each session is a new session rather than a continuation of the previous week, meaning that different groups of students can play together at different times and continue to benefit from the intervention. Furthermore, the game can be played with two, three or four participants each time to account for smaller groups due to uneven class sizes or missing students. The game intervention in the present trial was delivered in groups of four, as this allowed for more challenges to be completed by each player group (as each participant had to complete a certain number of challenges each game play session) during each game play session, increasing opportunities for social learning. This flexibility in the intervention design limited disruptions to game play in the present trial and is an invaluable feature for other similar trials in the field.

Clinical Implications

The present study tested a universal wellbeing intervention aimed at promoting wellbeing and preventing mental health decline in an adolescent sample within schools. According to the THRIVE Framework that is being implemented across England, supporting adolescents around usual challenges of life through wellbeing interventions during this critical developmental stage can help them thrive and empower them to cope better when faced with systemic or individual adversities (poverty, abuse, inequality, physical health problems as well as adjusting to changes in life circumstances or mild/temporary difficulties) (Wolpert et al., 2019). Universal wellbeing interventions, such as Dragons of Afterlands, can contribute to the evidence base around the development of these strengths-based skills development interventions at the community level.

The game-based depersonalized approach of Dragons of Afterlands can help improve accessibility of these interventions by appealing to harder to reach adolescents, who may be less likely to receive help through traditional pathways until problems develop. The contents of the intervention can also help reduce stigma around talking about emotions and mental health, although this was not assessed in the present trial.

The findings of the present trial provided insight into novel and innovative ways of embedding mental health support in schools, in line with NHS Long Term Plan (2019). Research has shown that implementation of such interventions, can help streamline access to mental health support by reducing waiting times in NHS services as seen by data provided by trusts where similar interventions were implemented (Tavistock and Portman NHS Foundation Trust, 2022) and therefore have impact on personal- and wider system levels.

Policy Implications

School based research is valuable as it can provide access to large cohorts of samples. Research evidence and policy around mental health and wellbeing interventions in schools strongly recommend "whole school approaches" to promoting wellbeing in schools, where all parts of the school work together coherently and committedly, creating a culture of change (O'Reilly et al., 2018; Public Health England, 2021b; NICE, 2008; NICE, 2009, Department for Education & Department of Health, 2017). The present study highlights challenges and shortfalls of school based interventions when whole school approaches are not taken and parts of the school may have conflicting attitudes towards the proposed intervention. These challenges included difficulties around recruitment (e.g. difficulties recruiting and retaining control group participants due to lack of staff facilitating contact with this group, staff not wishing for their students to miss classes to attend intervention), students feeling that participating in the

intervention is a conflict of demands put on them by the school (and this potentially affecting their motivation) and practical difficulties in the implementation of the intervention (e.g. staff availability for co-facilitation of intervention, room availability). It became evident in the recruitment for this study that attitudes and beliefs held by teachers and school administrators can significantly facilitate or hinder recruitment and retention of participants.

The role of leadership and management in schools in supporting efforts to promote wellbeing is central to success in ensuring changes are embedded and accepted (Public Health England, 2021b). Developing connections with schools can allow for researchers to have regular consistent access to schools and develop interventions that are context specific and in line with needs, although requiring longer involvement and more commitment from the staff (Moltrecht et al., 2021). It is, therefore, important to understand the views of staff to address any barriers and to ensure that the research conducted can be guided by the needs of the systems that it aims to work with. For the present intervention, qualitative studies can be used to explore these beliefs and attitudes of school staff and particularly focus on 1) their approach to game or XR based interventions and their perceived benefits, 2) any challenges they anticipate when participating in such trials (including organizational challenges), and 3) any contexts that they would find beneficial to have access to such interventions (content of interventions, their use and fit into existing teaching plans, ideas around implementation in schools).

Whilst digital health apps are held to high standards prior to being approved for public usage (e.g. NICE, 2021; Public Health England, 2017), there is a lack of guidance around the development stages of these apps. This is particularly an issue in XR, considering the push towards early testing of prototypes with client populations to avoid outdated technologies. Steps need to be taken to ensure that adverse effects on vulnerable populations are avoided. In the

present trial, these included managing expectations of participants at the beginning of the trial by informing them that technological glitches as well as issues with the gameplay can be expected as it is pilot testing. A psychologist (the researcher) was present during testing sessions to monitor emotional wellbeing of participants and manage disappointment or frustration should technical failures occur during the gameplay. This highlighted the importance of support from mental health clinicians at testing and development stages of mental health apps to minimize risk to participants. Further guidelines around safety and usability of prototypes in developmental stages of digital mental health apps are needed.

Part IV:

Integration, Impact and Dissemination Plan

Integration

The overall aim of this thesis was to contribute to the literature on wellbeing interventions for adolescent populations using XR technology. The topic of this project drew my attention, as it moved away from problem focused thinking that is common in mental health, and towards flourishing of young people and equipping them with skills to allow them to live their best lives. Children and adolescents are often referred to mental health services once problems have developed and often have to wait significant periods of time on waiting lists, leading to deterioration of their mental health (Smith et al., 2018b). Prevention in mental health is important and has a valuable impact (Durlak & Wells, 1997) and innovative interventions, such as this one, can potentially empower adolescents to thrive (according to the THRIVE Framework, explained previously in Introduction) (Wolpert et al., 2019). I found it particularly meaningful and enjoyable to work on a project that embraces creativity and playfulness when trying to create change, as these qualities often get lost or neglected in the seriousness of clinical psychology, despite their therapeutic benefits (Berger & Lahad, 2010).

The Systematic Review

The systematic review and the empirical study topics were closely linked and integrated well together. The systematic review explored the current state of the evidence with regards to available XR interventions for adolescents and their effectiveness in improving wellbeing. It identified a significant gap in the literature around the lack of XR wellbeing interventions overall and particularly around psychologically driven interventions, which then the empirical study

contributed to. The aim of choosing this particular systematic review topic was to use its findings to guide the design of the empirical study and to provide a basis for comparison.

Most studies in the systematic review looked at immediate effects of interventions on wellbeing (mostly in relation to emotions) and were not psychologically driven in content, but were rather interested in examining the effects of XR technology on wellbeing. Although identifying this gap in the literature was helpful in highlighting the uniqueness and originality of our empirical intervention, the findings of the systematic review did not provide much guidance for the design of the empirical study. A small number of studies met the inclusion criteria of the systematic review and 70% of these were deemed to have at least some risk of bias in their designs.

The systematic review could have been helpful in the identification and selection of outcome measures for the empirical study; however, most of the outcome measures used in the included studies assessed for immediate effects of interventions on wellbeing, which were not suited for our longitudinal design. Furthermore, the systematic review highlighted that outcome measures assessing wellbeing in a more global sense, rather than focusing on specific aspects (e.g. stress, positive affect etc.), were lacking in XR literature. As the systematic review was not helpful in the selection of outcome measures, further searches on literature around assessment of wellbeing in children and adolescents were conducted and WHO-5 and WEMWBS were identified, as the use of these together could assess wellbeing in a well-rounded manner.

Similarly, the lack of previously conducted studies meant that medium effect sizes were used in sample size calculations for the empirical study. Although in retrospect, small effect sizes would have been the more likely outcome of our intervention, the fact that the final recruited

sample was too small to be sufficiently powered to detect medium effects, suggests that aiming to recruit for small effects would not have been feasible.

The inclusion criteria for the systematic review were kept broad to capture an overview of the field. XR interventions in mental health are common; however, the majority of these focus on disorder-specific treatments for children and adolescents rather than on wellbeing. The XR interventions that were developed to be used as- or in conjunction with treatments were often developed under the influence of psychotherapeutic principles, which was lacking in the field of XR wellbeing interventions, as shown by the systematic review. These treatment interventions could have been interesting to examine as part of the systematic review, considering that we were aiming to evaluate a psychologically-driven intervention as part of the empirical study. However, scoping searches in the early stages of the systematic review revealed that previous systematic reviews had already reviewed interventions looking at mental health disorder related outcomes (Mesa-Gresa et al., 2018; Eijlers et al., 2019; Romero-Ayuso et al., 2021). Therefore, the scope of our systematic review was limited to wellbeing outcomes only to avoid duplication.

Wellbeing as a Construct

Searching the selected systematic review databases for wellbeing interventions revealed that wellbeing, as a construct, was defined using different definitions and assessed using a range of outcome measures in previous studies. This impacts the quality of the evidence base in the literature and posed challenges for the selection of measures in the empirical study and the inclusion criteria for the systematic review. As the focus of this project was primarily around promotion of wellbeing and prevention of mental health difficulties in the future rather than treatment, we opted for a positive psychological approach and avoided disorder or symptom oriented outcome measures for the empirical study and excluded studies using these in the

systematic review. In the empirical study, we used the definition of optimal wellbeing (described previously in Introduction) as a guide for the selection of outcome measures and aimed to select measures that would capture hedonic and eudaimonic aspects, as we felt that a comprehensive approach like this would be more meaningful. In the systematic review searches, several studies reported that they were measuring "wellbeing" but only including physical health outcomes or mental health disorder specific measures. Discussions with the second reviewer and supervisors were particularly helpful in this context in determining what can constitute as wellbeing interventions and outcome measures.

Although wellbeing is well defined within specific schools of psychology, lack of cohesion between these definitions prevents it from being a well-defined construct (Dodge et al., 2012; Baldwin et al., 2021). Whilst reviews of wellbeing outcome measures exist (Bentley et al., 2019), good practice guidelines are not available unlike for depression and anxiety (Wolpert, 2020, Funders agree first common metrics for mental health science). This is a need in the field to ensure that findings can be compared across studies and standards can be developed around effectiveness.

Recruitment

Recruitment and retention difficulties are common in school based research and lack of understanding around interventions offered, requirements for scientific rigor (randomization, sample sizes), competing demands on schools for delivery of curriculums as well as lack of contacts within schools to facilitate interventions are common reasons for these recruitment challenges, as experienced in the present trial (Smith & Petosa, 2016). Multiple recruitment avenues were explored to increase sample size; however, the number of participants recruited did not provide sufficient power for the analyses, as detailed previously in the empirical study.

Several factors hindered recruitment including the coronavirus pandemic and changing guidelines for schools around provision. Several of the schools that were approached for recruitment purposes experienced high staff turnover during the pandemic, which made building and maintaining connections with schools challenging. Staff changes hindered recruitment and retention, as some schools, that initially showed interest, were no longer reachable following staff members leaving their posts. Similarly, one of the schools that participated in the trial became less responsive after the facilitating teacher's contract ended, leading to difficulties with data collection and retention of participants. Engaging schools earlier on in the trial and co-producing the intervention and trial design together to meet their needs could have been a more effective way of recruiting (Smith & Petosa, 2016).

It is likely that the game itself may have been perceived with skepticism due to beliefs and attitudes around the possible effectiveness of game based interventions, the lack of previous trials proving the specific intervention's effectiveness, and the novelty and unfamiliarity of school staff around AR technology. Teacher attitudes when faced with novel technological game-based interventions is a determining factor of uptake and success (Yam et al., 2015). One of the ideas for service user involvement in the present study was asking a couple of participants to make a short video of their experiences of playing the game, as well as short instructions on the gameplay to demonstrate AR components. This video could have been a helpful way of explaining what an AR boardgame looks like, as well as provide information on the learning content and offer reassurance to school staff that it was well-received by previous participants. The video would have been used as part of recruitment efforts when reaching out to schools to briefly inform them of the trial in a more engaging way than study flyers. Research on recruitment and retention on participants in schools has shown that recruiting representatives in

schools who take ownership of projects and promote them within schools can help facilitate recruitment, retention and delivery of interventions (Smith & Petosa, 2016). Using these videos could have helped identify interested teachers, who could have then been appointed as members of the research team to co-produce and facilitate the trial together. Unfortunately, due to scheduling difficulties with schools around the filming of the video, it was not possible to proceed with this plan.

One challenge of the empirical study was balancing recruitment efforts and requests from schools regarding the details of their participation in the study, while maintaining research integrity and upholding best practice principles as much as possible. An example of this was when schools were not keen on randomization for their students. Following discussions with schools and attempts to problem solve any challenges, it was not possible to proceed with randomization of participants, which may have introduced biases. Schools wished to select the students that would participate in the intervention group and made attempts to make sure that group is representative where possible. This was a compromise that had to be made to retain the two interested schools.

Participation in the trial was skewed towards the private school, as many students from the state-funded school did not opt into the trial and several dropped out during the trial (86% drop out compared to 46% drop out in the private school). Majority of students in the state-funded school were from ethnically diverse backgrounds. Difficulties in engaging these students in mental health research could be related to higher prevalence of stigma around mental health (Gary, 2005) or skepticism and mistrust towards research within the context of inequalities in mental health care (Alvidrez & Arean, 2002). Coming from a diverse background myself and believing in the importance of extending mental health research beyond traditional Western

principles, this was a disappointing outcome for me. Co-production with adolescents from diverse backgrounds in earlier stages of future studies, particularly around methods of advertisement and sharing of relevant study information (e.g confidentiality, anonymity) can help address these barriers (The two youth advisors in present study reported their ethnicity as "white other".). One way that we could have engaged the student body better could have been through advertising the study in schools more by attending assemblies, events or putting up posters to familiarize students with the intervention (Smith & Petosa, 2016), which could have helped ease potential worries around participation.

Lack of power was common in the literature on XR interventions as shown by the systematic review and the empirical study was similarly affected by this issue as well. A further possible explanation, aside from recruitment difficulties detailed above, could be around the need for physical materials and equipment for these studies (e.g. headsets in VR studies or game board in ours), which limit the reach of participants compared to fully online interventions that can be circulated on social media platforms. Although the present empirical study's sample size is smaller compared to other studies included in the systematic review, this is likely influenced by its longitudinal design compared to the mostly short intervention studies included in the review. One systematic review on longitudinal digital mental health interventions for children and young people found that a third of the included RCTs (10 out of 30) had sample sizes smaller than 30 (the empirical study's sample size) (Hollis et al., 2017), showing that small sample sizes are a challenge for the field overall.

Service User Involvement

Service user involvement was achieved through the recruitment of two youth advisors. These advisors provided insight into all stages of the empirical study and their views were used

to interpret the findings and evaluate the strengths and weaknesses of the trial overall. Feedback gathered from the youth advisors was passed on to the researchers working on the next iteration of the Dragons of Afterlands project. Unfortunately, due to difficulties engaging the school staff of one of the participating schools, both youth advisors were recruited from the same year group in the private international school, which likely limited the breadth of input we received. Service user involvement was conceptualized as involvement of the target population of the intervention (i.e. adolescents); however, experiences of recruiting and delivering the intervention also revealed that this should have perhaps been conceptualized more broadly to include school staff and potentially parents. Teachers aiding in the facilitation of the intervention sessions had valuable insights by observing students play (e.g. concerns around age-appropriateness of phrasing of some of the wellbeing tasks) and these were noted as informal feedback and passed on to developers and the rest of the research team. Teachers also reported openness to participation as advisors or participants of qualitative studies in the future. These reflections were outlined to be kept in mind for future school based trials.

Impact

Wellbeing is important and has been associated with better physical health and healthier aging (Steptoe et al., 2015), better quality of relationships with others (Goswami, 2012) and academic success (Douglass & Islam, 2009).

Both the systematic review and the empirical trial covered original and novel content. The empirical study reported on the first trial of Dragons of Afterlands, the first known psychologically informed XR intervention that was specifically developed to improve wellbeing in adolescents. Similarly, the systematic review study was the first review on the topic of XR

wellbeing interventions for adolescents. The impacts of these studies on academic and non-academic fields are discussed below.

Academic Impact

The wellbeing intervention was influenced by three evidence based psychological models (CBT, narrative and systemic). These therapeutic approaches were integrated into an AR game both in relation to its psychoeducational and self-reflexive content but also in relation to the design of the game (e.g. depersonalization using avatars to aid externalization of problems) and in relation to the process of gameplay (e.g. playing in groups with allies and learning from each other).

The empirical study discussed the process of developing a new intervention, which provided valuable insights into research methods and approaches in the field, such as trialling early effectiveness testing of prototypes, co-production with adolescents and barriers to administration of such interventions in schools. It also contributed to the discussion in the XR literature regarding the motivational benefits of such interventions in longitudinal trials; although this could not be taken as definitive conclusions due to potential influences of choice of outcome measures and selection of participants.

The field of XR is rapidly growing, with various ongoing projects with varying similarities. Future research needs to build on what has been done to develop effective and acceptable interventions. Reflections reported in the present empirical study on what went well and didn't go well, as well as learning points and actions for future iterations of the trial, can be used as a guide for other researchers in the field to build on our experiences.

The development of the wellbeing intervention was the result of interdisciplinary collaborations between professionals from computer science, education and psychology

backgrounds, which has been pointed out as a significant need in the literature (Davies & Bergin, 2021). The process was an example of how storytelling and technology can be used creatively and cohesively to enhance the delivery of psychological techniques. The findings of the study highlight learning points for not only the field of mental health, but also for education and computer science as well and therefore has cross-disciplinary impact academically.

This experience of an interdisciplinary collaboration also demonstrated the different responsibilities psychologists can take on in these dynamics. For the present trial, these roles included leadership, consultation, as well as providing mental health containment and support during testing of the intervention. The beneficial use of transferable skills in this context, beyond the traditional role of mental health provision with clients in clinical settings, broadens the range of domains psychology can be applied within a leadership framework (British Psychological Society, 2010)

While interdisciplinary collaborations are encouraged in the field, one barrier that was identified in the scoping searches for the systematic review was that different disciplines utilize different platforms to disseminate their research, limiting collaborative development of ideas. For instance, when searching for XR wellbeing interventions for adolescents, using PubMed did not yield many relevant findings; whereas IEEE, a database not commonly used by psychologists, delivered much more useful articles to include in the review. This has highlighted the need to improve communication of ideas and findings between disciplines working on XR interventions to increase the impact of research in this area, potentially leading to faster growth, more innovation and higher likelihood of beneficial collaborations. Researchers working in the field should maintain an awareness that research content may be published in a wider range of

journals and research conferences should be circulated more inclusively of professionals from different backgrounds.

Non-academic Impact

This trial contributes to the development of future versions of the game, which if successful, can be implemented in schools. There is increasing focus on mental health in schools as reflected in government policies (Department for Education & Department of Health, 2017), emphasizing the importance of prevention and raising awareness of mental health and wellbeing, and recognizing the crucial role that schools play in this regard. The intervention, if successful, has the potential to improve mental wellbeing by increasing the effectiveness of wellbeing education in schools, and play a protective role in maintaining mental health, potentially reducing the need for mental health support in schools. The delivery of the intervention requires minimal staff input, as the game is played independently by groups of students. Therefore, the game would not compound existing responsibilities of school staff, who have been facing increasing burdens especially following the coronavirus pandemic (Gottenborg et al., 2021), and potentially reduce some of their workload (around teaching mental health content), which is a priority for the Department for Education (Department for Education, 2018b). The game intervention could potentially be scheduled into regular curriculum around wellbeing (e.g. as part of PSHE classes) or be made available as an extra-curricular activity for students to engage with in their free time (e.g. used by substitute teachers, or during breaks).

The recruitment difficulties with the empirical study highlighted barriers that need to be addressed in schools to improve access to research. Schools are uniquely placed to make wellbeing interventions accessible on a larger scale and target students who are at risk of experiencing mental health difficulties and poor wellbeing (Swick & Powers, 2018). However,

given the recruitment challenges and the consequences of low powered trials, alternative settings, where participants may be more readily accessible, can be considered in early stages of product development to gather evidence of effectiveness. These could potentially include youth centres or charities. Even though populations in these settings and their needs may differ from school-based cohorts, the impact of providing a wellbeing intervention, that is game based, depersonalized and hopefully easier to engage with can be significant in these settings as well. These trials can then provide evidence of effectiveness that can be used to broaden recruitment.

The use of the game is associated with some initial costs around the purchasing of the board, smartphones and access to data. If future versions of the game demonstrate effectiveness, it could be a cost-effective prevention intervention in the long term due to lack of ongoing running costs for the game and reduced burden on school staff in relation to mental health promotion. Prevention of mental health problems can also provide economic benefits, as the economic and social costs of mental health problems are estimated to be £105 billion each year in England (Knapp et al., 2011). It is important to ensure that any future interventions made available in the community, do not compound existing health inequalities in underprivileged populations who may not be able to afford the initial costs or have access to data or smartphones (Radovic & Badawy, 2020). Research with schools across a range of settings is necessary prior to rolling out this intervention in the future to ensure that any financial concerns are understood and addressed.

The process of conducting the study also highlighted some gaps in policies and implementation of policies, such as the difficulties in the implementation of whole school wellbeing interventions, the lack of guidance on early stage development and testing of mental health apps. These gaps were detailed in the reports of the study and can be used to guide the

development of future policies. As the field is rapidly growing, the need for such policies and guidelines will become more evident.

Dissemination Plan

The findings of the empirical study were presented virtually to first, second and third year trainee clinical psychologists at Royal Holloway University of London on 6th May 2022. The empirical study, with references to the systematic review findings, will also be presented as part of a continuing professional development (CPD) session on adolescent wellbeing and the use of AR interventions at the Chelsea and Westminster Hospital and South Kensington & Chelsea Mental Health Center, where the researcher is currently on placement. The CPD session will take place on 29th June 2022 and will be attended by clinicians working across disciplines.

Although the findings of the studies were inconclusive or non-significant, both studies addressed a significant gap in the literature that could benefit future research in the field. The publication bias has been identified as a significant concern in mental health and calls were made for researchers to publish non-significant findings to address failures to replicate and provide insights into validity of theories (Mehler et al., 2019; Munafò & Neill, 2016). Therefore, both the systematic review and the empirical study will be submitted for publication following the viva. Journals, such as JMIR Serious Games, which has a multidisciplinary audience in health, education, medicine, sociology and computer science, or International Journal of Wellbeing will be considered, as similar research was previously published on these platforms. We will also consider disseminating the findings at The International Society for Research on Internet Interventions (ISRII) conference, which focuses on the use of information and communication technologies in mental health.

The findings will be disseminated to the two participating schools, students and their families. Youth advisors were consulted on best methods of disseminating the study findings within their schools and age groups. The suggestions included asking staff to share findings in advisory assemblies in school and sharing findings on social media. They felt that it would be important to disseminate the findings to the school administrators as well as the students, so that the school administration would be open to participation in similar future trials. Following from these discussions, a lay summary of the study findings will be shared with the schools and addressed to the facilitating teachers as well as school administrators. An age appropriate child-friendly version of the lay summary will also be sent and school staff will be asked to share this as they see appropriate, with a suggestion to disseminate as part of the assemblies and via emails to students and parents. A lay audience blog summary will be posted on social media.

References

Abbott J. H. (2014). The distinction between randomized clinical trials (RCTs) and preliminary feasibility and pilot studies: what they are and are not. *The Journal of orthopaedic and sports physical therapy*, *44*(8), 555–558. https://doi.org/10.2519/jospt.2014.0110

Accenture. (2021). *Global Gaming Industry Value Now Exceeds $300 Billion, New Accenture Report Finds*. https://newsroom.accenture.com/news/global-gaming-industry-value-now-exceeds-300-billion-new-accenture-report-finds.htm

Adhyaru, J. S., & Kemp, C. (2022). Virtual reality as a tool to promote wellbeing in the workplace. *Digital Health*, 8, 205520762210844–205520762210844. https://doi.org/10.1177/20552076221084473

Adi, Y., Killoran, A., Janmohamed, K., & Stewart-Brown, S. (2007). Systematic review of the effectiveness of interventions to promote mental wellbeing in children in primary education. Report 1: Universal approaches: non-violence related outcomes. *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews.*

Al Farsi, G., Yusof, A. B. M., Fauzi, W. J. B., Rusli, M. E. B., Malik, S. I., Tawafak, R. M. & Jabbar, J. (2021). The Practicality of Virtual Reality Applications in Education: Limitations and Recommendations. *Journal of Hunan University Natural Sciences*, 48(7).

Allcoat, D., Hatchard, T., Azmat, F., Stansfield, K., Watson, D., & von Mühlenen Adrian. (2021). Education in the digital age: learning experience in virtual and mixed realities. *Journal of Educational Computing Research*, 59(5), 795–816. https://doi.org/10.1177/0735633120985120

Almog, I., Wallach, H. & Safir, M. (2009). Ethnicity and Sense of Presence in a Virtual

Environment: Arab Women - A Case in Point. 2009 Virtual Rehabilitation International Conference, VR 2009. 78 - 82. 10.1109/ICVR.2009.5174209.

Alvidrez, J., & Arean, P. (2002). Psychosocial treatment research with ethnic minority populations: ethical considerations in conducting clinical trials. *Ethics & Behavior*, 12(1), 103–116.

Alwashmi, M. F., Hawboldt, J., Davis, E., & Fetters, M. D. (2019). The iterative convergent design for mobile health usability testing: mixed methods approach. *JMIR mHealth and uHealth*, 7(4), 11656. https://doi.org/10.2196/11656

Arain, M., Campbell, M., Cooper, C. & Lancaster, G. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC medical research methodology*. 10. 67. 10.1186/1471-2288-10-67.

Asıcı, E. & Sarı, H. İ. (2021). The mediating role of coping strategies in the relationship between cognitive flexibility and well-being. *International Journal of Scholars in Education,* 4(1), 38-56. doi:10.52134/ueader.889204

Atwood-Blaine, D. & Huffman, D. (2017). Mobile Gaming and Student Interactions in a Science Center: the Future of Gaming in Science Education. *International Journal of Science and Mathematics Education. 1*5. 10.1007/s10763-017-9801-y.

Azevedo, L. B., Burges-Watson, D., Haighton, C., & Adams, J. (2014). The effect of dance mat exergaming systems on physical activity and health e related outcomes in secondary schools: Results from a natural experiment. *BMC Public Health*, 14(1), 951. http://doi.org/10.1186/1471-2458-14-951.

Badawy, S. M., Thompson, A. A., & Kuhns, L. M. (2017). Medication adherence and technology-based interventions for adolescents with chronic health conditions: a few key

considerations. *JMIR mHealth and uHealth*, 5(12), 202. https://doi.org/10.2196/mhealth.8310

Baldwin, D., Sinclair, J., & Simons, G. (2021). What is mental wellbeing? *BJPsych Open*, 7(S1), S236-S236. doi:10.1192/bjo.2021.631

Baños, R. M., Etchemendy, E., Mira, A., Riva, G., Gaggioli, A., & Botella, C. (2017). Online positive interventions to promote well-being and resilience in the adolescent population: a narrative review. *Frontiers in Psychiatry*, 8, 10–10. https://doi.org/10.3389/fpsyt.2017.00010

Bech P, Lunde M, Bech-Andersen G, Lindberg L, & Martiny K. (2007). Psychiatric outcome studies (POS): does treatment help the patients? A Popperian approach to research in clinical psychiatry. *Nordic Journal of Psychiatry*, 61:sup46, 4-34, DOI:10.1080/08039480601151238

Bech, P., Olsen, L. R., Kjoller, M., & Rasmussen, N. K. (2003). Measuring well-being rather than the absence of distress symptoms: a comparison of the sf-36 mental health subscale and the who-five well-being scale. *International Journal of Methods in Psychiatric Research*, 12(2), 85–91. https://doi.org/10.1002/mpr.145

Ben-Arieh, A. (2005). Where are the children? children's role in measuring and monitoring their well-being. *Social Indicators Research*, 74(3), 573–596.

Ben-Zur, H. (2003). Happy adolescents: the link between subjective well-being, internal resources, and parental factors. *Journal of Youth and Adolescence : A Multidisciplinary Research Publication*, 32(2), 67–79. https://doi.org/10.1023/A:1021864432505

Bentley, N., Hartley, S., & Bucci, S. (2019). Systematic review of self-report measures of general mental health and wellbeing in adolescent mental health. *Clinical Child and Family Psychology Review*, 22(2), 225–252. https://doi.org/10.1007/s10567-018-00273-x

Benzing, V., & Schmidt, M. (2018). Exergaming for children and adolescents: strengths, weaknesses, opportunities and threats. *Journal of Clinical Medicine*, 7(11). https://doi.org/10.3390/jcm7110422

Berenguer, C., Baixauli, I., Gómez, S., Andrés, M., & De Stasio, S. (2020). Exploring the Impact of Augmented Reality in Children and Adolescents with Autism Spectrum Disorder: A Systematic Review. *International Journal of Environmental Research and Public Health*, 17(17), 6143. https://doi.org/10.3390/ijerph17176143

Berger, R., & Lahad, M. (2010). A safe place: ways in which nature, play and creativity can help children cope with stress and crisis - establishing the kindergarten as a safe haven where children can develop resiliency. *Early Child Development and Care*, 180(7), 889–900. https://doi.org/10.1080/03004430802525013

Bergin, A. D., Vallejos, E. P., Davies, E. B., Daley, D., Ford, T., Harold, G., Hetrick, S., Kidner, M., Long, Y., Merry, S., Morriss, R., Sayal, K., Sonuga-Barke, E., Robinson, J., Torous, J., & Hollis, C. (2020). Preventive digital mental health interventions for children and young people: a review of the design and reporting of research. *NPJ Digital Medicine*, 3(1). https://doi.org/10.1038/s41746-020-00339-7

Bergström, M. & Baviskar, S. (2021) A Systematic Review of Some Reliability and Validity Issues regarding the Strengths and Difficulties Questionnaire Focusing on Its Use in Out-of-Home Care, *Journal of Evidence-Based Social Work*, 18:1, 1-31, DOI: 10.1080/26408066.2020.1788477

Best, P., Manktelow, R., & Taylor, B. (2014). Online communication, social media and adolescent wellbeing: a systematic narrative review. *Children and Youth Services Review,* 41, 27–36. https://doi.org/10.1016/j.childyouth.2014.03.001

Bird, J. M. (2020). The use of virtual reality head-mounted displays within applied sport psychology. *Journal of Sport Psychology in Action*, 11(2), 115–128. https://doi.org/10.1080/21520704.2018.1563573

Blakemore, S. J. (2019). Adolescence and mental health. *The Lancet*, 393(10185), 2030-2031.

Blom, E. H., Bech, P., Hogberg, G., Larsson, J. O., Serlachius, E. (2012). Screening for depressed mood in an adolescent psychiatric context by brief self-assessment scales - testing psychometric validity of WHO-5 and BDI-6 indices by latent trait analyses. *Health Qual Life Outcomes*, 10:149.

Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health*, 13(119), 1–20.

Bosworth, K. (2006) *Application of computer technology to drug abuse prevention. Handbook of drug abuse prevention.* Handbooks of sociology and social re-search. Boston, MA: Springer.

Bouchard, S. (2011). Could virtual reality be effective in treating children with phobias?. *Expert Review of Neurotherapeutics*, *11*(2), 207-213.

Brezinka, V. (2014). Computer games supporting cognitive behaviour therapy in children. *Clinical Child Psychology and Psychiatry*, 19(1), 100-110

British Psychological Society (Division of Clinical Psychology). (2010). *Clinical Psychology Leadership Development Framework*. Leicester:.

Browning, M. H. E. M., Mimnaugh, K. J., van Riper, C. J., Laurent, H. K., & LaValle, S. M. (2019). Can simulated nature support mental health? comparing short, single-doses of 360-degree nature videos in virtual reality with the outdoors. *Frontiers in Psychology*, *10*, 2667–2667. https://doi.org/10.3389/fpsyg.2019.02667

Bruijniks, S. J. E., Lemmens, L. H. J. M., Hollon, S. D., Peeters, F. P. M. L., Cuijpers, P., Arntz, A., Dingemanse, P., Willems, L., van Oppen, P., Twisk, J. W. R., van den Boogaard, M., Spijker, J., Bosmans, J., & Huibers, M. J. H. (2020). The effects of once- versus twice-weekly sessions on psychotherapy outcomes in depressed patients. *British Journal of Psychiatry,* 216(4), 222–230.

Bruno, L.E. (2019). *Get Gamified: Promoting Augmented Reality and Digital Game Technology in Education*. In: Geroimenko, V. (eds) Augmented Reality Games I. Springer, Cham. https://doi.org/10.1007/978-3-030-15616-9_15

Butti N, Biffi, E., Genova, C., Romaniello, R., Redaelli, D. F., Reni, G., Borgatti, R., & Urgesi, C. (2020). Virtual reality social prediction improvement and rehabilitation intensive training (vr-spirit) for paediatric patients with congenital cerebellar diseases: study protocol of a randomised controlled trial. *Trials*, *21*(1). https://doi.org/10.1186/s13063-019-4001-4

Button, K., Ioannidis, J., Mokrysz, C., Nosek, B., Flint, J., Robinson, E. & Munafò, M. (2013). Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience. *Nature reviews. Neuroscience.* 14. 10.1038/nrn3475.

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure

therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders*, 61, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003

Carroll, J., Hopper, L., Farrelly, A. M., Lombard-Vance, R., Bamidis, P. D., & Konstantinidis, E. I. (2021). A scoping review of augmented/virtual reality health and wellbeing interventions for older adults: redefining immersive virtual reality. *Frontiers in Virtual Reality*, 2. https://doi.org/10.3389/frvir.2021.655338

Cebolla, A., Herrero, R., Ventura, S., Miragall, M., Bellosta-Batalla, M., Llorens, R., & Baños, R. M. (2019). Putting oneself in the body of others: a pilot study on the efficacy of an embodied virtual reality system to generate self-compassion. *Frontiers in Psychology*, *10*, 1521–1521. https://doi.org/10.3389/fpsyg.2019.01521

Cheng, V. W. S., Davenport, T., Johnson, D., Vella, K., & Hickie, I. B. (2019). Gamification in apps and technologies for improving mental health and well-being: systematic review. *Jmir Mental Health*, *6*(6). https://doi.org/10.2196/13717

Chin, M. J. M., Jacobs,W. M., Vaessen, E. P. G., Titze, S., & van Mechelen,W. (2008). The motivation of children to play an active video game. *Journal of Science and Medicine in Sport*, 11(2), 163e166. http://doi.org/10.1016/j.jsams.2007.06.001.

Chung, CH., Chen, CH. (2017). Augmented Reality Based Social Stories Training System for Promoting the Social Skills of Children with Autism. In: Soares, M., Falcão, C., Ahram, T. (eds) Advances in Ergonomics Modeling, Usability & Special Populations. *Advances in Intelligent Systems and Computing*, vol 486. Springer, Cham. https://doi.org/10.1007/978-3-319-41685-4_44

Cilar, L., Štiglic, G., Kmetec, S., Barr, O., & Pajnkihar, M. (2020). Effectiveness of school-based mental well-being interventions among adolescents: a systematic review. *Journal of Advanced Nursing*, *(20200504)*. https://doi.org/10.1111/jan.14408

Clarke, A., Friede, T., Putz, R., Ashdown, J., Martin, S., Blake, A., Adi, Y., Parkinson, J., Flynn, P., Platt, S., & Stewart-Brown, S. (2011). Warwick-Edinburgh Mental Well-being Scale (WEMWBS): validated for teenage school students in England and Scotland. A mixed-methods assessment. *BMC Public Health*, 11(1), 1–9. https://doi.org/10.1186/1471-2458-11-487

Cobb, S., Heaney, R., Corcoran, O., & Henderson-Begg, S. (2009). The learning gains and student perceptions of a Second Life virtual lab. *Bioscience Education*,13(1), 1–9.

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

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 386-396.

Cooper, A. & Petrides, K. V. (2010). A psychometric analysis of the Trait Emotional Intelligence Questionnaire-Short Form (TEIQue-SF) using Item Response Theory. *Journal of Personality Assessment*, 92, 449-457.

Cottrell, D., & Boston, P. (2002). Practitioner review: the effectiveness of systemic family therapy for children and adolescents. *Journal of Child Psychology and Psychiatry*, 43(5), 573–586. https://doi.org/10.1111/1469-7610.00047

Cowan, R. L., Frederick, B. B., Rainey, M., Levin, J. M., Maas, L. C., Bang, J., Hennen, J., Lukas, S. E., Renshaw, P. F. (2000). Sex differences in response to red and blue light in

human primary visual cortex: a bold fMRI study. *Psychiatry Research*. 22;100(3):129-38. doi: 10.1016/s0925-4927(00)00074-3.

Coyle, D., McGlade, N., Doherty, G., & O'Reilly, G. (2011). *Exploratory evaluations of a computer game supporting cognitive behavioural therapy for adolescents*. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2937-2946).

Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology*, 43(Pt 3), 245–65.

Crosswell, L., & Yun, G. W. (2022). Examining virtual meditation as a stress management strategy on college campuses through longitudinal, quasi-experimental research. *Behaviour & Information Technology*, *41*(4), 864–878. https://doi.org/10.1080/0144929X.2020.1838609

Croudace, T., Wolpert, M., Deighton, J., Patalay, P., Brown, J., & Fonagy, P. (2014). Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: a review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health*, *8*(1), 1–14. https://doi.org/10.1186/1753-2000-8-14

Csikszentmihalyi, M., & Csikzentmihaly, M. (1990). *Flow: The psychology of optimal experience* (Vol. 1990). New York: Harper & Row.

D'Cunha, N. M., Nguyen, D., Naumovski, N., McKune, A. J., Kellett, J., Georgousopoulou, E. N., Frost, J., & Isbel, S. (2019). A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. *Gerontology*, 65(4), 430–440. https://doi.org/10.1159/000500040

Dalle Grave, R., Calugi, S., Doll, H. A., & Fairburn, C. G. (2013). Enhanced cognitive behaviour therapy for adolescents with anorexia nervosa: an alternative to family therapy? *Behaviour Research and Therapy*, 51(1), 12. https://doi.org/10.1016/j.brat.2012.09.008

Dave, D., & Varma, T. (2014). A review of various statistical methods for outlier detection. *Int. J. Comput. Sci. Eng. Technol*, *5*(2), 137-140.

Davies, E. B., & Bergin, A. D. (2021). Commentary: let's get digital: a commentary on Halldorsson et al.'s call for more rigorous development and evaluation of immersive digital interventions for children and young people's mental health. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *62*(5), 606–609. https://doi.org/10.1111/jcpp.13423

Daykin, N., Mansfield, L., Grigsby, D. L., D'Innocenzo, G., Kay, T., Meads, C., Julier, G., Tomlinson, A., Lane, J., Payne, A., Burnett, A., Dolan, P., Testoni, S., & Victor, C. (2018). What works for wellbeing? a systematic review of wellbeing outcomes for music and singing in adults. *Perspectives in Public Health*, *138*(1), 39–46. https://doi.org/10.1177/1757913917740391

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, 62, 119-142.

Demers, M., Martinie O., Winstein, C., & Robert, M. T. (2020). Active video games and low-cost virtual reality: an ideal therapeutic modality for children with physical disabilities during a global pandemic. *Frontiers in Neurology*, *11*. https://doi.org/10.3389/fneur.2020.601898

Department for Education. (2022). *The Impact of the COVID-19 pandemic on adolescent Mental Health.*

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063619/The_impact_of_the_COVID-19_pandemic_on_adolescent_mental_health.pdf

Department for Education. (2018a). *Mental health and behaviour in schools: Departmental advice for school staff*. London

Department for Education. (2018b). *Teacher Voice Omnibus Survey*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/687010/Teacher_Voice_report_Summer_2017.pdf

Department for Education & Department of Health. (2017). *Transforming children and young people's mental health provision: a Green Paper*. London.

Department of Health. (2013a). *Annual report of the chief medical officer: public mental health priorities: investing in the evidence*. London.

Department of Health. (2013b). *Developing Well – 11-19 Years.. A Compendium of Factsheets: Wellbeing Across the Lifecourse.* London.

Department of Health. (2013c). *Starting Well*. *A Compendium of Factsheets: Wellbeing Across the Lifecourse*. London.

Department of Health. (2014). *Wellbeing and Why it matters to health.* London. https://www.gov.uk/government/publications/wellbeing-and-health-policy

DeShazo J, Harris L, Pratt W. (2010) Effective intervention or child's play? A review of video games for diabetes education. *Diabetes Technol Ther*. 12(10):815-22. doi: 10.1089/dia.2010.0030. PMID: 20807119.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*(3), 542–575.

Diener, E., Pressman, S. D., Hunter, J., & Delgadillo-Chase, D. (2017). If, why, and when subjective well-being influences health, and future needed research. *Applied Psychology: Health and Well-Being*, *9*(2), 133–167. https://doi.org/10.1111/aphw.12090

Dodge, R., Daly, A., Huyton, J., & Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing*, 2(3), 222-235. doi:10.5502/ijw.v2i3.4

Douglass, L., & Islam, M. R. (2009). Emotional wellbeing of first year university students: Critical for determining future academic success. Non Refereed Paper.

Dray, J., Bowman, J., Campbell, E., Freund, M., Wolfenden, L., Hodder, R. K., McElwaine, K., Tremain, D., Bartlem, K., Bailey, J., Small, T., Palazzi, K., Oldmeadow, C., & Wiggers, J. (2017). Systematic review of universal resilience-focused interventions targeting child and adolescent mental health in the school setting. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(10), 813–824. https://doi.org/10.1016/j.jaac.2017.07.780

Drljević, N., Botički, I., & Wong, L.-H. (2022). Investigating the different facets of student engagement during augmented reality use in primary school. *British Journal of Educational Technology*, (20220214). https://doi.org/10.1111/bjet.13197

Drummond, D., Hadchouel, A.,, & Tesnière, A. (2017). Serious games for health: three steps forwards. *Advances in Simulation*, 2(1), 3–3. https://doi.org/10.1186/s41077-017-0036-3

Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: a meta-analytic review. *American Journal of Community Psychology*, 25(2), 115–152. https://doi.org/10.1023/A:1024654026646

Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Weissberg, R. P., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432.

Durose, C., Beebeejaun, Y., Rees, J., Richardson, J., & Richardson, L. (2012). *Towards Co-production in Research with Communities.* (Connected Communities). Arts and Humanities Research Council.

Early Intervention Foundation. (2021). *Half of UK teachers don't feel confident helping pupils with their mental health*. https://www.eif.org.uk/press-release/half-of-uk-teachers-don-t-feel-confident-helping-pupils-with-their-mental-health

Eijlers, R., Utens, E., Staals, L. M., de Nijs, P., Berghmans, J. M., Wijnen, R., Hillegers, M., Dierckx, B., & Legerstee, J. S. (2019). Systematic Review and Meta-analysis of Virtual Reality in Pediatrics: Effects on Pain and Anxiety. *Anesthesia and Analgesia*, 129(5), 1344–1353. https://doi.org/10.1213/ANE.0000000000004165

Eldridge, S., Bond, C., Campbell, M., Lancaster, G., Thabane, L. & Hopwell, S. (2013). Definition and reporting of pilot and feasibility studies. *Trials.* 14. O18. 10.1186/1745-6215-14-S1-O18.

Elor, A., Powell, M., Mahmoodi, E., Teodorescu, M., & Kurniawan, S. (2022). Gaming beyond the novelty effect of immersive virtual reality for physical rehabilitation. *IEEE Transactions on Games*, 14(1). https://doi.org/10.1109/TG.2021.3069445

Eysenbach, G., Kracht, C., Gallagher, S., Guetterman, T., Faric Nuša, Yorke, E., Varnes, L., Newby, K., Potts, H. W. W., Smith, L., Hon, A., Steptoe, A., & Fisher, A. (2019). Younger adolescents' perceptions of physical activity, exergaming, and virtual reality:

qualitative intervention development study. *JMIR Serious Games*, 7(2). https://doi.org/10.2196/11960

Fairburn, C. G., & Patel, V. (2017). The impact of digital technology on treatments and their dissemination. *Behaviour Research and Therapy*, 88, 19–25. https://doi.org/10.1176/appi.focus.16405.

Farič, N., Smith, L., Hon, A., Potts, H. W. W., Newby, K., Steptoe, A., & Fisher, A. (2021). A virtual reality exergame to engage adolescents in physical activity: mixed methods study describing the formative intervention development process. *Journal of Medical Internet Research,* 23(2), 18161. https://doi.org/10.2196/18161

Farrah, K., Young, K., Tunis, M. C., & Zhao, L. (2019). Risk of bias tools in systematic reviews of health interventions: an analysis of PROSPERO-registered protocols. *Systematic Reviews*, *8*(1), 1–9. https://doi.org/10.1186/s13643-019-1172-8

Fleming, T. M., de Beurs, D., Khazaal, Y., Gaggioli, A., Riva, G., Botella, C., Baños, R. M., Aschieri, F., Bavin, L. M., Kleiboer, A., Merry, S., Lau, H. M., & Riper, H. (2016). Maximizing the impact of e-therapy and serious gaming: time for a paradigm shift. *Frontiers in Psychiatry*, 7, 65–65. https://doi.org/10.3389/fpsyt.2016.00065

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

Flujas-Contreras, J. M., Ruiz-Castañeda, D., & Gómez, I. (2020). Promoting emotional well-being in hospitalized children and adolescents with virtual reality: usability and

acceptability of a randomized controlled trial. *Computers, Informatics, Nursing* : Cin, 38(2), 99–107. https://doi.org/10.1097/CIN.0000000000000586

Fox, Jesse & Tang, Wai Yen. (2014). Sexism in Online Video Games: The Role of Conformity to Masculine Norms and Social Dominance Orientation. Computers in Human Behavior. 33. 314-320. 10.1016/j.chb.2013.07.014.

Francis, J., Vella-Brodrick, D., & Chyuan-Chin, T. (2021). Effectiveness of online, school-based Positive Psychology Interventions to improve mental health and wellbeing: A systematic review. *International Journal of Wellbeing*, 11(4).

Galina, Y. M., Olga, A. S., & Yury, P. Z. (2018). The study of ethnic attitudes during interactions with avatars in virtual environments. *Psychology in Russia: State of Art*, *11*(1), 18–29. https://doi.org/10.11621/pir.2018.0102

Gary, F. A. (2005). Stigma: barrier to mental health care among ethnic minorities. *Issues in Mental Health Nursing*, 26(10), 979–99

Gee, B., Wilson, J., Clarke, T., Farthing, S., Carroll, B., Jackson, C., King, K., Murdoch, J., Fonagy, P., & Notley, C. (2021). Review: delivering mental health support within schools and colleges - a thematic synthesis of barriers and facilitators to implementation of indicated psychological interventions for adolescents. *Child and Adolescent Mental Health*, 26(1), 34–46. https://doi.org/10.1111/camh.12381

Georgeson, A. R., Highlander, A., Loiselle, R., Zachary, C., & Jones, D. J. (2020). Engagement in technology-enhanced interventions for children and adolescents: current status and recommendations for moving forward. *Clinical Psychology Review*, *78*, 101858–101858. https://doi.org/10.1016/j.cpr.2020.101858

Golby, J. & Wood, P. (2016). The Effects of Psychological Skills Training on Mental Toughness

and Psychological Well-Being of Student-Athletes. *Psychology*, **7**, 901-913. doi: 10.4236/psych.2016.76092.

Goswami, H. (2012). Social relationships and children's subjective well-being. *Social Indicators Research*, 107(3), 575–588.

Gottenborg, E., Yu, A., Naderi, R., Keniston, A., McBeth, L., Morrison, K., Schwartz, D., & Burden, M. (2021). Covid-19's impact on faculty and staff at a school of medicine in the us: what is the blueprint for the future? *BMC Health Services Research*, 21(1), 395–395. https://doi.org/10.1186/s12913-021-06411-6

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist*, 69(1), 66e78. http://doi.org/10.1037/a0034857

Greenleaf, W., Fowler, C., Torous, J., Birckhead, B., Khalil, C., Liu, X., Conovitz, S., Rizzo, A., Danovitch, I., Bullock, K., & Spiegel, B. (2019). Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. *Jmir Mental Health*, *6*(1). https://doi.org/10.2196/11973

Gudiño, O. G., Leonard, S., Stiles, A. A., Havens, J. F., & Cloitre, M. (2017*). STAIR narrative therapy for adolescents*. In Evidence-based treatments for trauma related disorders in children and adolescents (pp. 251-271). Springer, Cham

Hadley, W., Houck, C., Brown, L. K., Spitalnick, J. S., Ferrer, M., & Barker, D. (2019). Moving beyond role-play: evaluating the use of virtual reality to teach emotion regulation for the prevention of adolescent risk behavior within a randomized pilot trial. *Journal of Pediatric Psychology*, *44*(4), 425–435. https://doi.org/10.1093/jpepsy/jsy092

Halldorsson, B., Hill, C., Waite, P., Partridge, K., Freeman, D., & Creswell, C. (2021). 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. *Journal of Child Psychology and Psychiatry,* 62(5), 584–605. https://doi.org/10.1111/jcpp.13400

Hatta, M. H., Sidi, H., Siew Koon, C., Che Roos, N. A., Sharip, S., Abdul Samad, F. D., Wan Xi, O., Das, S., & Mohamed Saini, S. (2022). Virtual reality (VR) technology for treatment of mental health problems during covid-19: a systematic review. *International Journal of Environmental Research and Public Health*, 19(9). https://doi.org/10.3390/ijerph19095389

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.

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*(4), 474-503.

Holmbeck, G. N., O'Mahar, K., Abad, M., Colder, C., & Updegrove, A. (2006). *Cognitive-Behavioral Therapy with Adolescents: Guides from Developmental Psychology.* In P. C. Kendall (Ed.), Child and adolescent therapy: Cognitive-behavioral procedures (pp. 419–464). The Guilford Press

Hopp, T., & Gangadharbatla, H. (2016). Novelty effects in augmented reality advertising environments: the influence of exposure time and self-efficacy. *Journal of Current Issues & Research in Advertising*, 37(2), 113–130. https://doi.org/10.1080/10641734.2016.1171179

Houlden, V., Weich, S., Porto de Albuquerque João, Jarvis, S., Rees, K., & Schooling, C. M. (2018). The relationship between greenspace and the mental wellbeing of adults: a systematic review. *Plos One*, *13*(9). https://doi.org/10.1371/journal.pone.0203000

Hsieh, C.-Y., & Chen, T. (2019). Effect of Pokémon Go on the cognitive performance and emotional intelligence of primary school students. *Journal of Educational Computing Research*, *57*(7), 1849–1874.

Huang, W., Roscoe, R. D., Johnson-Glenberg, M. C., & Craig, S. D. (2021). Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *Journal of Computer Assisted Learning*, 37(3), 745–758. https://doi.org/10.1111/jcal.12520

Hyland, M. E., Lewith, G. T., & Wheeler, P. (2008). Do existing psychologic scales measure the nonspecific benefit associated with cam treatment? . *J Altern Complement Med*. 14(2), 185–189. https://doi.org/10.1089/acm.2007.7050

Ismail, N., Kinchin, G. & Edwards, J. (2018). Pilot Study, Does It Really Matter? Learning Lessons from Conducting a Pilot Study for a Qualitative PhD Thesis. *International Journal of Social Science Research*, 6(1) pp. 1–17. https://doi.org/10.5296/ijssr.v6i1.11720

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*, *17*(1), 1–12. https://doi.org/10.1016/0197-2456(95)00134-4

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical

guide with flowcharts. *BMC Medical Research Methodology*, *17*(1), 162–162. https://doi.org/10.1186/s12874-017-0442-1

Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: a systematic review of the literature. *Internet Interventions*, *6*, 89–106. https://doi.org/10.1016/j.invent.2016.10.002

Joronen, K., Aikasalo, A., & Suvitie, A. (2016). Nonphysical effects of exergames and child and adolescent well-being: A comprehensive systematic review. *Scandinavian Journal of Caring Sciences,* 31(3), 449–461. doi:10.1111/ scs.12393

Kang H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, 64(5), 402–406. https://doi.org/10.4097/kjae.2013.64.5.402

Kavanagh, S., Luxton-Reilly, A., Wuensche, B., & Plimmer, B. (2017). A systematic review of virtual reality in education. *Themes in Science and Technology Education*, *10*(2), 85-119.

Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: the empirical encounter of two traditions. *Journal of Personality and Social Psychology*, *82*(6), 1007–22.

Kim, H-Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry and Endodontics 38*, 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Klein, J. B., Jacobs, R. H., & Reinecke, M. (2007). Cognitive-behavioral therapy for adolescent depression: a meta-analytic investigation of changes in effect-size estimates. *Journal of the American Academy of Child & Adolescent Psychiatry,* 46(11), 1403–1413. https://doi.org/10.1097/chi.0b013e3180592aaa

Knapp, M, McDaid, D, Parsonage, M (2011) *Mental health promotion and mental illness prevention: the economic case.* . Department of Health, London, UK.

Kothgassner, O. D., & Felnhofer, A. (2021). Lack of research on efficacy of virtual reality exposure therapy (VRET) for anxiety disorders in children and adolescents. *Neuropsychiatrie*, *35*(2), 68-75.

Kothgassner, O. D., Goreis, A., Kafka, J. X., Van Eickels, R. L., Plener, P. L., & Felnhofer, A. (2019). Virtual reality exposure therapy for posttraumatic stress disorder (PTSD): A meta-analysis. *European Journal of Psychotraumatology*, 10, 1. https://doi.org/10.1080/20008198.2019.1654782

Kuosmanen, T., Clarke, A.M. & Barry, M.M. (2019). Promoting adolescents' mental health and wellbeing: evidence synthesis. *Journal of Public Mental Health*, Vol. 18 No. 1, pp. 73-83. https://doi.org/10.1108/JPMH-07-2018-0036

Kuykendall, L., Tay, L., & Ng, V. (2015). Leisure engagement and subjective well-being: a meta-analysis. *Psychological Bulletin*, 141(2), 364–403. https://doi.org/10.1037/a0038508

Kwan, B., & Rickwood, D. J. (2015). A systematic review of mental health outcome measures for young people aged 12 to 25 years. *BMC psychiatry*, *15*, 279. https://doi.org/10.1186/s12888-015-0664-x

Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424–1442.

Lee, L. N., Kim, M. J., & Hwang, W. J. (2019). Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. *Applied Sciences*, 9(17), 3556–3556. https://doi.org/10.3390/app9173556

Lehtimaki, S., Martic, J., Wahl, B., Foster, K. T., Schwalbe, N. (2021). Evidence on Digital Mental Health Interventions for Adolescents and Young People: Systematic Overview *JMIR Ment* Health ;8(4):e25847 doi: 10.2196/25847

Lenhard, W. & Lenhard, A. (2016). Computation of effect sizes. https://www.psychometrica.de/effect_size.html. *Psychometrica*. DOI: 10.13140/RG.2.2.17823.92329

Leventhal, K. S., Gillham, J., DeMaria, L., Andrew, G., Peabody, J., & Leventhal, S. (2015). Building psychosocial assets and wellbeing among adolescent girls: a randomized controlled trial. *Journal of Adolescence*, *45*(1), 284–295. https://doi.org/10.1016/j.adolescence.2015.09.011

Lilford, R., & Stevens, A. J. (2002). Underpowered studies. *British Journal of Surgery, 89,* 129–131.

Lim, C.-Y., & In, J. (2019). Randomization in clinical studies. *Korean Journal of Anesthesiology*, *72*(3), 221–232. https://doi.org/10.4097/kja.19049

Lindner, P., Miloff, A., Hamilton, W., & Carlbring, P. (2019). The potential of consumer-targeted virtual reality relaxation applications: descriptive usage, uptake and application performance statistics for a first-generation application. *Frontiers in Psychology*, 10, 132–132. https://doi.org/10.3389/fpsyg.2019.00132

Little, R.J.A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202.

Liu, Q., Wang, Y., Yao, M. Z., Tang, Q., & Yang, Y. (2020). The effects of viewing an uplifting 360-degree video on emotional well-being among elderly adults and college students under immersive virtual reality and smartphone conditions. *Cyberpsychology, Behavior and Social Networking*, *23*(3), 157–164. https://doi.org/10.1089/cyber.2019.0273

Lopes, R. T., Gonçalves, M. M., Machado, P. P. P., Sinai, D., Bento, T., & Salgado, J. (2014). Narrative therapy vs. cognitive-behavioral therapy for moderate depression: empirical evidence from a controlled clinical trial. *Psychotherapy Research : Journal of the Society for Psychotherapy Researc*h, 24(6), 662–74. https://doi.org/10.1080/10503307.2013.874052

Lorenzo, G., Gómez-Puerta Marcos, Arráez-Vera Graciela, & Lorenzo-Lledó Alejandro. (2019). Preliminary study of augmented reality as an instrument for improvement of social skills in children with autism spectrum disorder. *Education and Information Technologies : The Official Journal of the Ifip Technical Committee on Education*, 24(1), 181–204. https://doi.org/10.1007/s10639-018-9768-5

Lotan, M., Yalon-Chamovitz, S., & Weiss, P. L. (T. (2010). Virtual reality as means to improve physical fitness of individuals at a severe level of intellectual and developmental disability. *Research in Developmental Disabilities*, *31*(4), 869–874. https://doi.org/10.1016/j.ridd.2010.01.010

Ma, N., Roberts, R., Winefield, H., & Furber, G. (2015). Utility of qualitative metasynthesis: Advancing knowledge on the wellbeing and needs of siblings of children with mental health problems. *Qualitative Psychology, 2*(1), 3–28. https://doi.org/10.1037/qup0000018

Madigan, S. (2011). *Narrative therapy.* American Psychological Association.

Maheswaran, H., Weich, S., Powell, J., & Stewart-Brown, S. (2012). Evaluating the responsiveness of the Warwick Edinburgh mental well-being scale (WEMWBS): group and individual level analysis. *Health and Quality of Life Outcomes*, *10*(1), 1–8. https://doi.org/10.1186/1477-7525-10-156

Marsch, L. A., & Borodovsky, J. T. (2016). Technology based interventions for preventing and treating substance use among youth. *Child and Adolescent Psychiatric Clinics of North America,* 25, 755–768.

Martin, A. (2014). *Motivating Learners with Virtual Reality*. In J. Viteli & M. Leikomaa (Eds.), Proceedings of EdMedia 2014--World Conference on Educational Media and Technology (pp. 2631-2636). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/147851/.

Mayordomo-Rodríguez, T., Meléndez-Moral, J. C., Viguer-Segui, P., & Sales-Galán, A. (2015). Coping strategies as predictors of well-being in youth adult. *Social Indicators Research, 122*(2), 479–489. https://doi.org/10.1007/s11205-014-0689-4

McAuley, E. D., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, 60(1), 48-58. https://doi.org/10.1080/02701367.1989.10607413

Mehler, D. M. A., Edelsbrunner, P. A., & Matić, K. (2019). Appreciating the Significance of Non-significant Findings in Psychology. *Journal of European Psychology Students*, 10(4), 1–7. Doi:10.5334/jeps.e2019a

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, 70, 29–40.

Mertika, A., Mitskidou, P., & Stalikas, A. (2020). "Positive Relationships" and their impact on wellbeing: A review of current literature. *Psychology: the Journal of the Hellenic Psychological Society*, 25(1), 115-127

Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J.-A., & Gil-Gómez, J.-A. (2018). Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: an evidence-based systematic review. *Sensors*, 18(8). https://doi.org/10.3390/s18082486

Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and immersion in video games: The aftermath of a conceptual challenge. *Frontiers in Psychology*, 9, 1682.

Moltrecht, B., Patalay, P., Bear, H. A., Deighton, J., & Edbrooke-Childs, J. (2021). Interdisciplinary Development of a Transdiagnostic Mobile App to Enhance Children's Emotion Regulation: Sharing Insights and Lessons Learned. https://doi.org/10.31234/osf.io/ks7ag

Montana, J. I., Matamala-Gomez, M., Maisto, M., Mavrodiev, P. A., Cavalera, C. M., Diana, B., Mantovani, F., & Realdon, O. (2020). The benefits of emotion regulation interventions in virtual reality for the improvement of wellbeing in adults and older adults: a systematic review. *Journal of Clinical Medicine*, *9*(2). https://doi.org/10.3390/jcm9020500

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105-125.

Morrish, L., Rickard, N., Chin, T. C., & Vella-Brodrick, D. A. (2018). Emotion regulation in adolescent well-being and positive education. *Journal of Happiness Studies : An*

Interdisciplinary Forum on Subjective Well-Being, 19(5), 1543–1564. https://doi.org/10.1007/s10902-017-9881-y

Munafò, M., & Neill, J. (2016). Null is beautiful: on the importance of publishing null results. *Journal of Psychopharmacology*. 30(7), 585–585. https://doi.org/10.1177/0269881116638813

Mrazek, P. J. & Haggerty, R. J. (Eds.). (1994*). Reducing risks for mental disorders: Frontiers for Preventive Intervention Research*.

National Institute for Health and Care Excellence (NICE). (2021). *Evidence Standards Framework for Digital Health Technologies.* https://www.nice.org.uk/corporate/ecd7

National Institute for Health and Care Excellence (NICE). (2008). *Social and emotional wellbeing in primary education.* London.

National Institute for Health and Care Excellence (NICE). (2009). *Social and emotional wellbeing in secondary education*. London.

Naylor, M., Ridout, B., & Campbell, A. (2020). A scoping review identifying the need for quality research on the use of virtual reality in workplace settings for stress management. *Cyberpsychology, Behavior and Social Networking*, 23(8), 506–518. https://doi.org/10.1089/cyber.2019.0287

Naylor, M., Ridout, B., Campbell, A., & Morrison, B. (2019). Augmented experiences: investigating the feasibility of virtual reality as part of a workplace wellbeing intervention. *Interacting with Computers*, 31(5), 507–523. https://doi.org/10.1093/iwc/iwz033

NHS Digital. (2017). *Mental Health of Children and Young People in England*. https://files.digital.nhs.uk/95/AC12EC/MHCYP%202017%20Multiple%20Conditions.pdf

NHS Long Term Plan. (2019). https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

Ni, M. Y., Hui, R. W. H., Li, T. K., Tam, A. H. M., Choy, L. L. Y., Ma, K. K. W., Cheung, F., & Leung, G. M. (2019). Augmented reality games as a new class of physical activity interventions? the impact of Pokémon Go use and gaming intensity on physical activity. *Games for Health Journal*, 8(1), 1–6. https://doi.org/10.1089/g4h.2017.0181

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, 7(2), 133-144.

Norton, M. J. (2021). Co-Production within Child and Adolescent Mental Health: A Systematic Review. *International Journal of Environmental Research and Public Health*, 18(22), 11897. https://doi.org/10.3390/ijerph182211897

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the Association for Information Science and Technology*, 59(6), 938–955.

O'Reilly, M., Svirydzenka, N., Adams, S., & Dogra, N. (2018). Review of mental health promotion interventions in schools. *Social Psychiatry and Psychiatric Epidemiology*, 53(7), 647–662. https://doi.org/10.1007/s00127-018-1530-1

Office for National Statistics (ONS). (2014). *Suicide in the United Kingdom*. Statistical Bulletin 2016:1-33

Opriş, D., Pintea, S., Garc a-Palacios, A., Botella, C., Szamosk zi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and Anxiety*, 29(2), 85-93.

Orben, A., Tomova, L., & Blakemore, S. J. (2020). The effects of social deprivation on adolescent development and mental health. *The Lancet Child & Adolescent Health*, *4*(8), 634-640.

Ostrow, K. S., Heffernan, N. T., & 19th International Conference on Artificial Intelligence in Education, AIED (2018). Testing the validity and reliability of intrinsic motivation inventory subscales within ASSISTments. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10947 Lnai*, 381–394. https://doi.org/10.1007/978-3-319-93843-1_28

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Plos Medicine*, 18(3), 1003583. https://doi.org/10.1371/journal.pmed.1003583

Pallavicini, F., & Pepe, A. (2019). *Comparing player experience in video games played in virtual reality or on desktop displays: Immersion, flow, and positive emotions.* Extended abstracts of the annual symposium on computer-human interaction in play companion extended abstracts (pp. 195–210)

Papastergiou, M. (2009). Digital game-based learning in high school computer science

education: Impact on educational effectiveness and student motivation. Computers & Education, 52(1), 1-12. doi: 10.1016/j.compedu.2008.06.004

Parekh V, Shah D, Shah M (2020) Fatigue detection using artificial intelligence framework. *Augment Hum Res* 5(1):5. https://doi.org/10.1007/s41133-019-0023-4

Park, M. J., Kim, D. J., Lee, U., Na, E. J., & Jeon, H. J. (2019). A literature overview of virtual reality (VR) in treatment of psychiatric disorders: recent advances and limitations. *Frontiers in Psychiatry*, 10, 505–505. https://doi.org/10.3389/fpsyt.2019.00505

Parsons, T. D., Riva, G., Parsons, S., Mantovani, F., Newbutt, N., Lin, L., Venturini, E., & Hall, T. (2017). Virtual reality in pediatric psychology. *Pediatrics*, 140(Suppl 2), 91. https://doi.org/10.1542/peds.2016-1758I

Pascual-Sanchez, A., Nicholls, D., Patalay, P., Crosby, L., McColoud, T., Hudson, L., & Viner, R. (2020). You-COPE. *Mental health consequences experienced by young people aged*, *16-24.*

Philipp, E. S., Andreia, P. C., Georges, S., & Alexander, F. S. (2020). The WHO-5 well-being index – validation based on item response theory and the analysis of measurement invariance across 35 countries. *Journal of Affective Disorders Reports*, 1, 100020–100020. https://doi.org/10.1016/j.jadr.2020.100020.

Plante, Thomas & Frazier, S. & Tittle, A. & Babula, M. & Ferlic, E. & Riggs, E.. (2003). Does virtual reality enhance the psychological benefits of exercise?. *Journal of Human Movement Studies.* 45. 485-507.

Polat, Ş., Afşar Doğrusöz, L., & Yeşil, A. (2022). The relationship between cognitive flexibility and happiness among nurses. *Perspectives in Psychiatric Care*, *58*(4), 2862–2871. https://doi.org/10.1111/ppc.13134

Pollard, E. L., & Lee, P. D. (2003). Child well-being: a systematic review of the literature. *Social Indicators Research : An International and Interdisciplinary Journal for Quality-Of-Life Measurement*, *61*(1), 59–78. https://doi.org/10.1023/A:1021284215801

Pratscher, S. D. (2020). *Transformative Technology: Examining the Capacity of Virtual Reality Induced Awe to Change Self-Transcendent Values and Improve Hedonic and Eudaimonic Well-Being* (Doctoral dissertation, University of Missouri-Columbia).

Prithwijit, D., Meng'ou, Z., Laura, M. L., Zaid, B., & Ruth, L. M. (2018). Augmented reality video games: new possibilities and implications for children and adolescents. *Multimodal Technologies and Interaction*, 2, 8–8. https://doi.org/10.3390/mti1020008

Public Health England. (2021a). *COVID-19 mental health and wellbeing surveillance report, Chapter 7: Children and young people*. London.

Public Health England. (2017). *Criteria for Health App Assessment*. London. https://www.gov.uk/government/publications/health-app-assessment-criteria/criteria-for-health-app-assessment

Public Health England. (2021b). *Promoting Children and Young People's Mental Health and Wellbeing: A Whole School or College Approach*. London

Public Health England. (2014). *The link between pupil health and wellbeing and attainment*. London.

Radovic, A., & Badawy, S. M. (2020). Technology use for adolescent health and wellness. *Pediatrics*, 145(Supplement_2), 194. https://doi.org/10.1542/peds.2019-2056G

Rammos, D., Bratitsis, T., & 8th International Conference on Games and Learning Alliance, GALA 2019 8th 2019 11 27 - 2019 11 29. (2019). *Alternative teaching of history subject in primary school: the case of the 3d hit playful activity.* Lecture Notes in Computer

Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11899 Lncs, 457–467. https://doi.org/10.1007/978-3-030-34350-7_44

Reeves, B. C. (2008). Principles of research: limitations of non-randomized studies. Surgery (Oxford), 26(3), 120–124. https://doi.org/10.1016/j.mpsur.2008.02.004

Rice, S. P. M., & Shorey-Fennell, B. R. (2020). Comparing the psychometric properties of common measures of positive and negative emotional experiences: implications for the assessment of subjective wellbeing. *Journal of Well-Being Assessment*, 4(1), 37–56. https://doi.org/10.1007/s41543-020-00025-1

Riches, S., Azevedo, L., Bird, L., Pisani, S., & Valmaggia, L. (2021). Virtual reality relaxation for the general population: a systematic review. *Social Psychiatry and Psychiatric Epidemiology,* 56(10), 1707–1727. https://doi.org/10.1007/s00127-021-02110-z

Richesin, M. T., Baldwin, D. R., & Wicks, L. A. M. (2021). Art making and virtual reality: a comparison study of physiological and psychological outcomes. *The Arts in Psychotherapy*, *75*. https://doi.org/10.1016/j.aip.2021.101823

Ridout, B., Kelson, J., Campbell, A., & Steinbeck, K. (2021). Effectiveness of virtual reality interventions for adolescent patients in hospital settings: systematic review. *Journal of Medical Internet Research*, *23*(6), 24967. https://doi.org/10.2196/24967

Riedinger, V., Pinquart, M., & Teubert, D. (2017). Effects of systemic therapy on mental health of children and adolescents: a meta-analysis. *Journal of Clinical Child and Adolescent Psychology : The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53,* 46(6), 880–894. https://doi.org/10.1080/15374416.2015.1063427

Romero-Ayuso, D., Toledano-González, A., Rodríguez-Martínez, M. del C., Arroyo-Castillo, P., Triviño-Juárez, J. M., González, P., Ariza-Vega, P., Del Pino González, A., & Segura-Fragoso, A. (2021). Effectiveness of Virtual Reality-Based Interventions for Children and Adolescents with ADHD: A Systematic Review and Meta-Analysis. *Children*, 8(2), 70. https://doi.org/10.3390/children8020070

Romrell, D. (2013). Gender and Gaming: A Literature Review. 36th Annual Proceedings: Selected research and development papers presented at the Annual Convention of the AECT. 170-182.

Rose, T., Joe, S., Williams, A., Harris, R., Betz, G., & Stewart-Brown, S. (2017). Measuring mental wellbeing among adolescents: a systematic review of instruments. *Journal of Child and Family Studies*, *26*(9), 2349–2362. https://doi.org/10.1007/s10826-017-0754-0

Ruini, C., Albieri, E., Ottolini, F., & Vescovelli, F. (2020). Once upon a time: A school positive narrative intervention for promoting well-being and creativity in elementary school children. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000362

Ruiz-Ariza, A., Casuso, R. A., Suarez-Manzano, S., & Martínez-López Emilio J. (2018). Effect of augmented reality game Pokémon go on cognitive performance and emotional intelligence in adolescent young. *Computers & Education*, 116, 49–63. https://doi.org/10.1016/j.compedu.2017.09.002

Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education*, *128*, 256–268. https://doi.org/10.1016/j.compedu.2018.09.015

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450-461.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166. https://doi.org/10.1146/annurev.psych.52.1.141

Sancassiani, F., Pintus, E., Holte, A., Paulus, P., Moro, M. F., Cossu, G., Angermeyer, M. C., Carta, M. G., & Lindert, J. (2015). Enhancing the emotional and social skills of the youth to promote their wellbeing and positive development: a systematic review of universal school-based randomized controlled trials. *Clinical Practice and Epidemiology in Mental Health : Cp & Emh,* 11(Suppl 1 M2), 21–40. https://doi.org/10.2174/1745017901511010021

Santos Garduno, H.A., Esparza Martinez, M.I., Portuguez Castro, M. (2021). Impact of Virtual Reality on Student Motivation in a High School Science Course. *Appl. Sci.* 11, 9516. https://doi.org/10.3390/app11209516

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., Patton, G. C. (2018). The age of adolescence. *Lancet Child Adolesc Health.* 2(3):223-228. doi: 10.1016/S2352-4642(18)30022-1.

Shamrova, D. P., & Cummings, C. E. (2017). Participatory action research (PAR) with children and youth: an integrative review of methodology and PAR outcomes for participants, organizations, and communities. *Children and Youth Services Review*, 81, 400–412.

Shanyinde, M., Pickering, R., & Weatherall, M. (2011). Questions asked and answered in pilot

and feasibility randomized controlled trials. *BMC medical research methodology*. 11. 117. 10.1186/1471-2288-11-117.

Shoshani, A., & Steinmetz, S. (2013). Positive psychology at school: a school-based intervention to promote adolescents' mental health and well-being. *Journal of Happiness Studies,* 15(6), 1289–1311. https://doi.org/10.1007/s10902-013-9476-1

Singh, D. K. A., Rahman, N. N. A. A., Rajiman, S., Yin, C. S., Karim, Z. A., Ruslan, A. S., & Singh, R. K. H. (2017). Impact of virtual reality games on psychological well-being and upper limb performance in adults with physical disabilities: a pilot study. *Medical Journal of Malaysia*, *72*(2), 119–121.

Smith, J., Beauchamp, M. R., Faulkner, G., Morgan, P. J., Kennedy, S. G., & Lubans, D. R. (2018a). Intervention effects and mediators of well-being in a school-based physical activity program for adolescents: the 'resistance training for teens' cluster RCT. *Mental Health and Physical Activity*, *15*, 88–94. https://doi.org/10.1016/j.mhpa.2018.08.002

Smith, J., Kyle, R. G., Daniel, B., & Hubbard, G. (2018b). Patterns of referral and waiting times for specialist child and adolescent mental health services. *Child and Adolescent Mental Health*, 23(1), 41–49. https://doi.org/10.1111/camh.12207

Smith, L. H., & Petosa, R. L. (2016). Effective practices to improve recruitment, retention, and partnerships in school-based studies. *Journal of Pediatric Health Care*, *30*(5), 495–498. https://doi.org/10.1016/j.pedhc.2016.05.004

Snelling, J. (2016). *Virtual reality in K-12 education: How helpful is it? Converge.* http://www.centerdigitaled.com/k-12/Virtual-Reality-in-K-12-Education-Is-It-Really-Helpful.html.

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., & Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, *27*(1), 281–295. https://doi.org/10.1038/s41380-021-01161-7

Šouláková, B., Kasal, A., Butzer, B., & Winkler, P. (2019). Meta-review on the effectiveness of classroom-based psychological interventions aimed at improving student mental health and well-being, and preventing mental illness. *The Journal of Primary Prevention*, 40(3), 255–278. https://doi.org/10.1007/s10935-019-00552-5

Spiegel, B., Fuller, G., Lopez, M., Dupuy, T., Noah, B., Howard, A., Albert, M., Tashjian, V., Lam, R., Ahn, J., Dailey, F., Rosen, B. T., Vrahas, M., Little, M., Garlich, J., Dzubur, E., IsHak, W., Danovitch, I., & Gupta, V. (2019). Virtual reality for management of pain in hospitalized patients: a randomized comparative effectiveness trial. *Plos One*, 14(8). https://doi.org/10.1371/journal.pone.0219115

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in cognitive sciences*, *9*(2), 69-74.

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The Lancet,* 385(9968), 640–648

Sterne, J. A. C., Savović Jelena, Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán Miguel A, Hopewell, S., Hróbjartsson Asbjørn, Junqueira, D. R., Jüni Peter, Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). Rob 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, 366. https://doi.org/10.1136/bmj.l4898

Steyn, B. J. M., Steyn, M. H., Maree, D. J. F., & Panebianco-Warrens, C. (2016). Psychological skills and mindfulness training effects on the psychological wellbeing of undergraduate music students: an exploratory study. *Journal of Psychology in Africa*, *26*(2), 167–171. https://doi.org/10.1080/14330237.2016.1163906

Swick, D., & Powers, J. D. (2018). Increasing access to care by delivering mental health services in schools: the school-based support program. *School Community Journal*, 28(1), 129–144

Taggart, F., & Stewart-Brown, S. (2019). A Review of questionnaires designed to measure mental wellbeing. *Retrieved on*, *25*.

Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., Brown, C., Stahl, S., Kaiser, M. D., & 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. https://doi.org/10.1111/j.1469-7610.2010.02258.x

Tavistock and Portman NHS Foundation Trust. (2022). *Helping Young People THRIVE: Implementing a person-centred model of care for young people's mental health*. London. https://www.health.org.uk/improvement-projects/helping-young-people-thrive-a-person-centred-model-of-care-for-young-people's#:~:text=THRIVE%20is%20an%20integrated%2C%20person,Help%20and%20Getting%20Risk%20Support

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale

(WEMWBS): development and UK validation. *Health and Quality of Life Outcomes*, *5*, 63–63.

Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing, 1(3),* 176-184. https://doi.org/10.1111/j.1524-475X.2004.04006.x

Tong, X., Gromala, D., Amin, A., & Choo, A. (2015). *The design of an immersive mobile virtual reality serious game in cardboard head-mounted display for pain management*. In International Symposium on Pervasive Computing Paradigms for Mental Health (pp. 284-293). Springer, Cham.

Topp C.W., Østergaard S.D., Søndergaard S., & Bech P. (2015). The WHO-5 Well-Being Index: A Systematic Review of the Literature. *Psychotherapy and Psychosomatics*, 84, 167-176.

Turner, R., Bird, S. & Higgins, J. (2013). The Impact of Study Size on Meta-analyses: Examination of Underpowered Studies in Cochrane Reviews. *PloS one*. 8. e59202. 10.1371/journal.pone.0059202.

Tyrer, S., & Heyman, B. (2016). Sampling in epidemiological research: Issues, hazards and pitfalls. *BJPsych Bulletin*, 40(2), 57-60. doi:10.1192/pb.bp.114.050203

van Agteren, J., Iasiello, M., Lo, L., Bartholomaeus, J., Kopsaftis, Z., Carey, M., & Kyrios, M. (2021). A systematic review and meta-analysis of psychological interventions to improve mental wellbeing. *Nature Human Behaviour*, 5(5), 631–652. https://doi.org/10.1038/s41562-021-01093-w

van Agteren, J., Iasiello, M. & Lo, L. (2018). Improving the wellbeing and resilience of health

services staff via psychological skills training. *BMC Res Notes* **11**, 924. https://doi.org/10.1186/s13104-018-4034-x

Ventura, S., Baños, R. M., Botella, C., & Mohamudally, N. (2018). Virtual and augmented reality: New frontiers for clinical psychology. *State of the art virtual reality and augmented reality knowhow*, *10*.

Vermeulen, L., Van Looy, J., De Grove, F., & Courtois, C. (2011). You are what you play?: a quantitative study into game design preferences across gender and their interaction with gaming habits. In *Proceedings of DiGRA 2011 conference : think design play*. Utrecht, The Netherlands: Digital Games Research Association (DiGRA).

Vostanis, P., Humphrey, N., Fitzgerald, N., Deighton, J., & Wolpert, M. (2013). How do schools promote emotional well-being among their pupils? findings from a national scoping survey of mental health provision in English schools. *Child and Adolescent Mental Health*, 18(3), 151–157. https://doi.org/10.1111/j.1475-3588.2012.00677.x

Waldron, H. B., & Kaminer, Y. (2004). On the learning curve: the emerging evidence supporting cognitive-behavioral therapies for adolescent substance abuse. *Addiction*, 99, 93–105. https://doi.org/10.1111/j.1360-0443.2004.00857.x

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, 24(6), 1319–1333.

Whitehead, A., Sully, B., Campbell, M. (2014). Pilot and feasibility studies: Is there a difference from each other and from a randomised controlled trial?. *Contemporary clinical trials*. 38. 10.1016/j.cct.2014.04.001.

Wille, D., Eng, K., Holper, L., Chevrier, E., Hauser, Y., Kiper, D., Pyk, P., Schlegel, S., &

Meyer-Heim, A. (2009). Virtual reality-based paediatric interactive therapy system (pits) for improvement of arm and hand function in children with motor impairment-a pilot study. *Developmental Neurorehabilitation*, *12*(1), 44–52. https://doi.org/10.1080/17518420902773117

Winefield, H. R., Gill, T. K., Taylor, A. W., & Pilkington, R. M. (2012). Psychological well-being and psychological distress: is it necessary to measure both? Psychology of Well-Being. *Theory, Research and Practice*, 2(1), 1–14. https://doi.org/10.1186/2211-1522-2-3

Wohn, D.Y., Ratan, R., Cherchiglia, L. (2020). Gender and Genre Differences in Multiplayer Gaming Motivations. In: Fang, X. (eds) HCI in Games. HCII 2020. Lecture Notes in Computer Science(), vol 12211. Springer, Cham. https://doi.org/10.1007/978-3-030-50164-8_16

Wolpert, M., Harris, R., Hodges, S., Fuggle, P., James, R., Wiener, A.,… Munk, S. (2019). *THRIVE Framework for system change*. London: CAMHS Press

Wolpert. (2020). *Funders agree first common metrics for mental health science,* https://www.linkedin.com/pulse/funders-agree-first-common-metrics-mental-health-science-wolpert

Wols, A., Scholte, R. H. J., & Qualter, P. (2015). Prospective associations between loneliness and emotional intelligence. *Journal of Adolescence*, 39, 40e48. http://doi.org/10.1016/j.adolescence.2014.12.007.

World Health Organization. (2013). *Mental health action plan 2013-2020*. Geneva: WHO Document Production Services.

World Health Organization. (2004a). *Prevention of mental disorders: Effective interventions and policy options: Summary report*. Geneva.

World Health Organization. (2004b). *Promoting Mental Health: Concepts, Emerging evidence, practice.* ISBN 9241591595

World Health Organization WHO. (1998). *Wellbeing Measures in Primary Health Care/The Depcare Project.* WHO Regional Office for Europe: Copenhagen.

Yam, S. C., Swati, M., & Jing, C. O. (2015). Professional development for scaling pedagogical innovation in the context of game-based learning: teacher identity as cornerstone in "shifting" practice. *Asia-Pacific Journal of Teacher Education*, 43(5), 423–437. https://doi.org/10.1080/1359866X.2014.962484

Young Minds. (2020). *Coronavirus: Impact on young people with mental health needs*. London. https://www.youngminds.org.uk/about-us/reports-and-impact/coronavirus-impact-on-young-people-with-mental-health-needs/

Zhang, Y., Mavoa, S., Zhao, J., Raphael, D., & Smith, M. (2020). The Association between Green Space and Adolescents' Mental Well-Being: A Systematic Review. *International Journal of Environmental Research and Public Health*, *17*(18), 6640. https://doi.org/10.3390/ijerph17186640

Zirbel, Chloe & Zhang, Xiaorong & Hughes, Charmayne. (2018). The VRehab System: A Low-Cost Mobile Virtual Reality System for Post-Stroke Upper Limb Rehabilitation for Medically Underserved Populations. 1-8. 10.1109/GHTC.2018.8601885.

Appendices

Appendix A. Outcome Measure for Empirical Study 1

WHO-5 Wellbeing Index

Please respond to each item by marking one box per row, regarding how you felt in the last two weeks.	All of the time	Most of the time	More than half the time	Less than half the time	Some of the time	At no time
I have felt cheerful and in good spirits.	5	4	3	2	1	0
I have felt calm and relaxed.	5	4	3	2	1	0
I have felt active and vigorous.	5	4	3	2	1	0
I woke up feeling fresh and rested.	5	4	3	2	1	0
My daily life has been filled with things that interest me.	5	4	3	2	1	0

Appendix B. Outcome Measure for Empirical Study 2

Strengths and Difficulties Questionnaire (SDQ)

For each item, please mark the box for Not True, Somewhat True or Certainly True. It would help us if you answered all items as best you can even if you are not absolutely certain or the item seems daft! Please give your answers on the basis of how things have been for you over the last six months.

	Not True	Somewhat True	Certainly True
I try to be nice to other people. I care about their feelings.			
I am restless, I cannot stay still for long.			
I get a lot of headaches, stomach-aches or sickness.			
I usually share with others (food, games, pens etc.)			
I get very angry and often lose my temper.			
I am usually on my own. I generally play alone or keep to myself.			
I usually do as I am told.			
I worry a lot.			
I am helpful if someone is hurt, upset or feeling ill.			
I am constantly fidgeting or squirming.			
I have one good friend or more.			
I fight a lot. I can make other people do what I want.			
I am often unhappy, down-hearted or tearful.			
Other people my age generally like me.			
I am easily distracted, I find it difficult to concentrate.			
I am nervous in new situations. I easily lose confidence.			
I am kind to younger children.			
I am often accused of lying or cheating.			
Other children or young people pick on me or bully me.			
I often volunteer to help others (parents, teachers, children).			
I think before I do things.			
I take things that are not mine from home, school or elsewhere.			
I get on better with adults than with people my own age.			
I have many fears, I am easily scared.			
I finish the work I am doing. My attention is good.			

Appendix C. Outcome Measure for Empirical Study 3

Intrinsic Motivation Inventory (IMI)

	Not at all True			Somewhat True			Very True
	1	2	3	4	5	6	7
I enjoyed doing this activity very much.							
I believe this activity could be of some value to me.							
This activity was fun to do.							
I would be willing to do this again because it has some value to me.							
I thought this was a boring activity.							
I believe doing this activity could be beneficial to me.							
This activity did not hold my attention at all.							
I think this is an important activity.							
I would describe this activity as very interesting.							
I think that doing this activity is useful for my wellbeing.							
I thought this activity was quite enjoyable.							
I think this is important to do because it can teach me skills.							
While I was doing this activity, I was thinking about how much I enjoyed it.							
I think doing this activity could help me look after my mental health.							

Appendix D. Outcome Measure for Empirical Study 4

Warwick Edinburgh Mental Wellbeing Scale (WEMWBS)

Below are some statements about feelings and thoughts. Please select the answer that best describes your experience of each over the last 2 weeks.

	None of the Time	Rarely	Some of the Time	Often	All of the Time
I've been feeling optimistic about the future.	1	2	3	4	5
I've been feeling useful.	1	2	3	4	5
I've been feeling relaxed.	1	2	3	4	5
I've been feeling interested in other people.	1	2	3	4	5
I've had energy to spare.	1	2	3	4	5
I've been dealing with problems well.	1	2	3	4	5
I've been thinking clearly.	1	2	3	4	5
I've been feeling good about myself.	1	2	3	4	5
I've been feeling close to other people.	1	2	3	4	5
I've been feeling confident.	1	2	3	4	5
I've been able to make up my own mind about things.	1	2	3	4	5
I've been feeling loved.	1	2	3	4	5
I've been interested in new things.	1	2	3	4	5
I've been feeling cheerful.	1	2	3	4	5

Appendix E. Content of Wellbeing Challenges
Content of Wellbeing Challenges

Wellbeing Domain	Topics and Skills Covered
Cognitive	• CBT based approaches to understanding thoughts and feelings • Recognizing the effects of others' judgements on self-confidence and self-esteem • Recognizing the effects of media on body image • Promoting cognitive flexibility by generating alternative explanations for a series of situations (e.g. friend said they would call but didn't) • Perceptions- focus on what you have rather than what you don't have • Perceptions- how do others see you? • Reframing and cognitive reappraisal • Antecedent, Behavior, Consequence • Reflecting on actions and alternative ways of behaving • Ambiguous stimulus- identifying the dragon's thoughts, feelings and behaviors
Emotional	• Recognizing emotions (in self and others focusing on range of cues including facial expressions, body language etc.) • Labelling emotions • Being able to describe emotions to others • Learning how to manage emotions • Strategies for managing stress, anxiety and depression
Social	• Understanding the difference between constructive feedback and unhelpful criticism • The effects of peer pressure and ways of resisting it • Harassment and ways of managing it • Bullying and cyberbullying • Consent and boundaries • What healthy and unhealthy relationships look like • Ways of dealing with an issue online (e.g. sexting) • Ways of preventing online issues • Becoming a better listener • Ways of questions • Open vs closed questions
Physical	• Nutrition • Exercise and its physical and mental benefits • Wellbeing in relation to place and nature • Body image

- Puberty and changing bodies
- The effects of smoking (first and second hand), understanding the terms “habit”, “dependence” and “addiction”
- Alcohol and drugs
- Screen time and screen addiction
- Gambling and gaming

Appendix F. Recruitment Email Template for Schools

Dear SENCO/ Wellbeing Lead,

We would like to invite you to take part in an exciting opportunity to support young people's wellbeing. Our research team, consisting of clinical psychology experts and educational games developers, have created an innovative board game using augmented reality technology. The novel game aims to provide adolescents with valuable skills to improve their wellbeing in schools in an effective and engaging way. Please find the attached leaflet for information on our trial for this wellbeing game.

Please do not hesitate to contact us if you have any questions about the wellbeing game trial. We look forward to hearing from you.

Idil Kilinc

Trainee Clinical Psychologist
Royal Holloway University of London

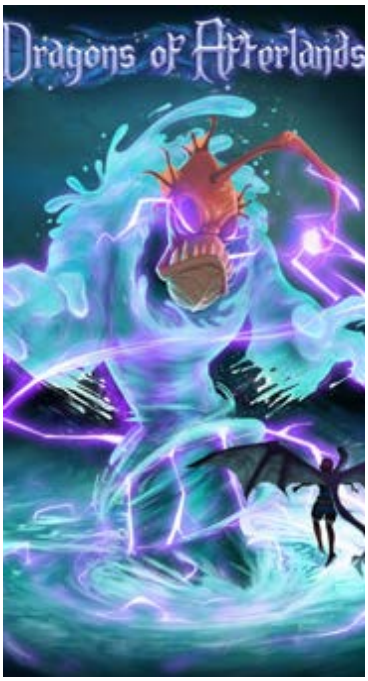
Appendix G. Recruitment Flyer

AUGMENTED REALITY BOARD GAME TO IMPROVE ADOLESCENT WELLBEING

The Proposed Wellbeing Game Trial

Providing adolescents with knowledge and skills to improve their wellbeing is important, as it can promote better mental health and prevent psychological problems. Research shows that playing games is good for wellbeing (Ruiz-Ariza et al., 2018; Koivisto et al., 2019).

We designed a board game for adolescents, using augmented reality technology to create an immersive and engaging experience. Students play the game with their peers and complete various tasks to progress in the game. The game tasks are based on PSHE wellbeing curriculum and are grounded in well-established psychological models to promote learning. In this proposed wellbeing game trial, we aim to evaluate if this game is successful in improving wellbeing in adolescents in a school setting.



Benefits of Participation

The game intervention offers a novel method of encouraging students to participate in wellbeing skills development during their PSHE lectures. This can potentially make teaching both more fun and effective. School based mental health interventions can play a preventative role and increase students' ability to function better academically and socially in schools, decreasing negative outcomes, such as bullying or non-attendance (Lindsey, 2017). Participating schools will have the opportunity to offer a unique learning experience to their students and contribute to research on improving adolescent mental health. Participating students will be included in a prize draw (one 50 pound voucher, two 25 pound vouchers).

Who are we?

The research team includes Idil Kilinc (trainee clinical psychologist at RHUL), Helen Pote (clinical psychologist with 20+ years of experience of working with young people & a professor at RHUL) and Sarah Campbell (founder & director of Play Well for Life, which develops educational games).

Who are we recruiting?

We will recruit students between the ages of 13 and 18. We are looking to recruit 10-15 students.

What does it involve for the school?

- The wellbeing game trial will run over 6 weeks.
- The trial will take place during regular PSHE hour.
- We will meet with teachers and students in the first session to introduce them to the game and to facilitate the completion of baseline questionnaires.
- Students will play the game with their peer groups during PSHE hour for 6 weeks.
- Students will complete questionnaires at the end of 6 weeks and one month after that. These will be facilitated by us.
- A control group will run alongside the game intervention group. The control group will be attending their regular PSHE classes and will be asked to complete questionnaires at the same time points outlined above.
- We will be contactable throughout the trial should any questions or difficulties arise.
- All necessary equipment will be provided.

We asking interested schools to please contact Idil Kilinc (Trainee Clinical Psychologist) (idil.kilinc.2019@live.rhul.ac.uk).

Appendix H. Information Sheet and Consent Form for Parents

INFORMATION SHEET FOR RESEARCH STUDY

DRAGONS OF AFTERLANDS

AUGMENTED REALITY BOARD GAME TO IMPROVE WELLBEING

The Wellbeing Game Trial

Providing teenagers with knowledge and skills to improve their wellbeing is important, as it can promote better mental health and prevent psychological problems. Research shows that playing games is good for wellbeing (Ruiz-Ariza et al., 2018; Koivisto et al., 2019). We designed a board game specifically to support your teenager's wellbeing skills in a fun and engaging way using augmented reality technology. Teens will play the game with their friends and complete various tasks that improve their learning. The game tasks are based on wellbeing curriculum and are designed by clinical psychologists with over 20 years of experience in adolescent psychology, in collaboration with a company designing educational games to improve mental health. In this trial, we aim to evaluate if this game is successful in improving wellbeing of teenagers in a school setting.

Participation in the Wellbeing Game Trial

The wellbeing game trial will take place during regular class hour at school and will run over six weeks. Participating students will be divided into two groups, with one group attending their regular lessons and the other group playing the game in groups of four during their lecture hour. Both groups will be asked to complete questionnaires about their wellbeing at the beginning and end of the study and one month later. This will help us track any changes in their wellbeing over time to test if the game is effective.

Participation in this wellbeing game trial is entirely voluntary and those who choose to not participate will not be affected by their decision. Participants may choose to stop their participation at any time point. They do not need to provide an explanation for this and will not be affected negatively in any way. We will provide all necessary equipment for your child's participation in the study.

Confidentiality and Use of Collected Information

All data collected throughout the trial will be kept confidential and will only be visible to the research team. Information will be stored anonymously on secure, password protected systems for up to 4 years and cannot be linked back to the participants. Information collected in this trial will only be used for the analysis of this project. Findings of this trial will be disseminated in scientific publications.

Who are we?

This wellbeing game trial is run by Idil Kilinc (trainee clinical psychologist at RHUL- idil.kilinc.2019@live.rhul.ac.uk) under the supervision of Prof. Helen Pote (clinical psychologist with 20+ years of experience of working with young people & professor at RHUL- h.pote@rhul.ac.uk) and Dr. Sarah Campbell (founder & director of Play Well for Life, which develops educational games- sarah@playwellforlife.com). Please feel free to contact us if you have any questions about the trial. This trial has been reviewed and approved by the College Ethics Committee at Royal Holloway University of London.

Please tear off and hand to the class teacher if you do not want your child to participate in this study.

Participation in this trial is voluntary and participants are free to withdraw from the project at any time. Participating teenagers will be asked to fill out questionnaires, which will be stored anonymously on secure systems for up to 4 years and will be used to analyse the trial.

I do not want my child to participate in this study.

Name of Student:.....

Signature of Parent:..... Date:.....

Appendix I. Information Sheet and Consent Form for Adolescents (Participants)

AUGMENTED REALITY BOARD GAME TO IMPROVE WELLBEING

DRAGONS OF AFTERLANDS

WHAT IS DRAGONS OF AFTERLANDS?

Learning about ways of taking care of our mental health is important, as it can help us live better lives and achieve our goals. We wanted to explore whether it is possible for people to develop these skills through playing a game with their friends. We came up with a board game called Dragons of Afterlands, that uses augmented reality technology. In this wellbeing game trial, we want to test if playing this game can improve wellbeing in schools.

PARTICIPATING IN THE GAME TRIAL

We are looking for young people between the ages of 13 and 18 to join our trial. The game trial will take place at school. Half of the participating students will attend their regular classes, whereas the other half will play the game during their class hour for six weeks. We will ask everyone to complete questionnaires about their wellbeing.

Participating in this project is optional and you can decide if you would like to participate or not. If you decide to participate, you can choose to stop at any time as well.

WHAT HAPPENS TO QUESTIONNAIRES?

We will keep the information from your questionnaires safe and private for up to 4 years. Only the researchers will see your answers. You will fill out the questionnaires anonymously, meaning that your information will not be linked to your name.

WHO ARE WE?

Idil Kilinc is a trainee clinical psychologist from Royal Holloway University of London (RHUL). Helen Pote is a clinical psychologist and a professor at RHUL. Dr Sarah Campbell is CEO of educational games company, Play Well For Life.

Please complete the following section if you would like to participate in this study.

Yes, I want to participate in this study. I understand that my participation is optional and that I can stop participating at any time. I agree that my questionnaires can be stored anonymously for up to 4 years and that this information can be used to draw conclusions about the game trial.

Name:..... *Date:.....*

Appendix J. Ethics Application

Ethics Review Details

You have chosen to submit your project to the REC for review:	
Name:	Kilinc, Idil (2019)
Email:	NHUT017@live.rhul.ac.uk
Title of research project or grant:	A cluster randomised controlled trial to evaluate the effectiveness of an augmented reality (AR) board game for adolescent wellbeing
Project type:	Royal Holloway postgraduate research project/grant
Department:	Psychology
Academic supervisor:	Prof Helen Pote
Email address of Academic Supervisor:	h.pote@rhul.ac.uk
Funding Body Category:	No external funder
Funding Body:	
Start date:	29/04/2021
End date:	29/07/2022

Research question summary:

There is emerging evidence with non-clinical populations that digital technologies can have a positive impact on mental health and wellbeing (Collins et al., 2019). Studies have shown that playing Augmented Reality (AR) games have a positive impact on psychological, physical and social wellbeing outcomes (Ruiz-Ariza et al., 2018; Koivisto et al., 2019). However, these effects may not remain after participants stop playing the games, as it is the experience, rather than the game content that contributes to wellbeing in these contexts. To our knowledge, no study has tested the impact of psychologically informed AR games on youth wellbeing. AR games, based on evidence-based psychological principles, have the potential to provide psychoeducation and aid in the development of knowledge and skills to improve wellbeing within an immersive format. Immersive games are likely to increase engagement with the content of the game, leading to increased learning and longer lasting effects than the immediate effects of playing games. The developed game is novel in that it adopts a preventative approach to mental health by drawing on cognitive behavioural, systemic and narrative approaches, within a social, hybrid experience.

Aim: To test if an augmented reality (AR) game can improve wellbeing in adolescents when compared to existing interventions to improve wellbeing.

Hypothesis 1: Adolescents playing the AR game, will report higher wellbeing on standardised wellbeing outcome measures at the end of the intervention and at one month follow up compared to their baseline scores.

Hypothesis 2: Adolescents, who play the AR game, will report equivalent or higher wellbeing compared to those who attend regular PSHE lessons at the end of the intervention and at one month follow up.

Hypothesis 3: Adolescents, who play the AR game, will report higher engagement with the process on standardised outcome measures, compared to those, who attend regular PSHE classes.

Research method summary:

Design

The study will follow a cluster randomised controlled design. The intervention arm of the trial will involve adolescents playing the newly developed AR board game during PHSE hour and the control arm will involve adolescents attending their usual PSHE classes at school. The trial will span over six weeks and a follow up assessment will be conducted one month after the end of the trial.

Participants

Participants will include 86 adolescents, between the ages of 13 and 16. Blindness, deafness and cognitive impairment will be exclusion criteria.

Intervention

The game intervention was developed using cognitive behavioural and systemic narrative approaches, and will be presented as an AR

board game to create a multisensory immersive serious game experience. Participants will complete various tasks to progress in the game. Tasks promote learning and skills development in the four domains of wellbeing in adolescents; physical, social, cognitive and emotional. Adolescents will play the game in groups of 4.

Outcomes

Demographics will be collected at baseline for descriptive purposes. Participants will complete the Strengths and Difficulties questionnaire (emotional and behavioural screening tool) as a baseline measure to characterise the sample. Completion of wellbeing tasks within the game will be monitored through in-game measurements collected throughout the intervention. Participants will complete wellbeing questionnaires at baseline, end of intervention and at one month follow up. These will include WHO-5 (measure of overall subjective wellbeing) and Warwick Edinburgh Mental Wellbeing Scale (positive mental wellbeing scale including affective-emotional aspects, cognitive and psychological functioning). Participants will complete the Intrinsic Motivation Inventory after the first and sixth session as a measure of their engagement.

Analysis

We will use 2-way mixed ANOVAs to assess changes over time in wellbeing and engagement outcome measures.

Risks to participants

Does your research involve any of the below?

Children (under the age of 16),

Yes

Participants with cognitive or physical impairment that may render them unable to give informed consent,

No

Participants who may be vulnerable for personal, emotional, psychological or other reasons,

No

Participants who may become vulnerable as a result of the conduct of the study (e.g. because it raises sensitive issues) or as a result of what is revealed in the study (e.g. criminal behaviour, or behaviour which is culturally or socially questionable),

No

Participants in unequal power relations (e.g. groups that you teach or work with, in which participants may feel coerced or unable to withdraw),

No

Participants who are likely to suffer negative consequences if identified (e.g. professional censure, exposure to stigma or abuse, damage to professional or social standing),

No

Details,

Study will take place in schools and will involve participants between the ages of 13 and 16. Informed consent will be obtained from the participants and their parents. Participants and parents will be informed that they can withdraw from the study at any time point if they wish to do so. Game tasks are depersonalised and are unlikely to cause distress to the participants. Participants will be advised to contact researchers if they find themselves getting distressed about any of the tasks. If participants feel distressed following their participation in the study or if they score clinically severe on outcome measures, pastoral care services will be provided within schools. Participants will be informed of these during recruitment. Participating schools' safeguarding policies will be adhered to if any information gathered during the study leads to safeguarding concerns regarding individual participants.

Design and Data

Does your study include any of the following?

Will it be necessary for participants to take part in the study without their knowledge and/or informed consent at the time?,
No

Is there a risk that participants may be or become identifiable?,
No

Is pain or discomfort likely to result from the study?,
No

Could the study induce psychological stress or anxiety, or cause harm or negative consequences beyond the risks encountered in normal life?,
No

Does this research require approval from the NHS?,
No

If so what is the NHS Approval number,

Are drugs, placebos or other substances to be administered to the study participants, or will the study involve invasive, intrusive or potentially harmful procedures of any kind?,
No

Will human tissue including blood, saliva, urine, faeces, sperm or eggs be collected or used in the project?,
No

Will the research involve the use of administrative or secure data that requires permission from the appropriate authorities before use?,
No

Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?,
Yes

Is there a risk that any of the material, data, or outcomes to be used in this study has been derived from ethically-unsound procedures?,
No

Details,
Participants will be entered into a prize draw (1x50 pounds, 2x25 pounds).
Minor changes may be made by schools to the information and consent form (attached) to reflect their local procedures but key information will not be altered.

Risks to the Environment / Society

Will the conduct of the research pose risks to the environment, site, society, or artifacts?,
No

Will the research be undertaken on private or government property without permission?,
No

Will geological or sedimentological samples be removed without permission?,
No

Will cultural or archaeological artifacts be removed without permission?,
No

Details,

Risks to Researchers/Institution

Does your research present any of the following risks to researchers or to the institution?

Is there a possibility that the researcher could be placed in a vulnerable situation either emotionally or physically (e.g. by being alone with vulnerable, or potentially aggressive participants, by entering an unsafe environment, or by working in countries in which there is unrest)?,
No

Is the topic of the research sensitive or controversial such that the researcher could be ethically or legally compromised (e.g. as a result of disclosures made during the research)?,
No

Will the research involve the investigation or observation of illegal practices, or the participation in illegal practices?,
No

Could any aspects of the research mean that the University has failed in its duty to care for researchers, participants, or the environment / society?,
No

Is there any reputational risk concerning the source of your funding?,
No

Is there any other ethical issue that may arise during the conduct of this study that could bring the institution into disrepute?,
No

Details,

Declaration
By submitting this form, I declare that the questions above have been answered truthfully and to the best of my knowledge and belief, and that I take full responsibility for these responses. I undertake to observe ethical principles throughout the research project and to report any changes that affect the ethics of the project to the University Research Ethics Committee for review.

Certificate produced for user ID, NHUT017

Date:	17/03/2021 17:03
Signed by:	Kilinc, Idil (2019)
Digital Signature:	Idil Kilinc
Certificate dated:	17/03/2021
Files uploaded:	Adolescent info sheet_consent form.pdf Parent information sheet_consent form.pdf research-ethics-risk-checklistv3.pdf Full-Review-2566-2021-03-17-17-29-NHJT017.pdf schools leaflet.pdf Full-Review-2566-2021-03-17-17-32-.pdf

Appendix K. Ethics Application Approval

ES

Ethics Application System <ethics@rhul.ac.uk>

To: Kilinc, Idil (2019); Pote, H; Ethics

Mon 2021-04-19 11:45 PM

PI: Prof Helen Pote
Project title: A cluster randomised controlled trial to evaluate the effectiveness of an augmented reality (AR) board game for adolescent wellbeing

REC ProjectID: 2566

Your application has been approved by the Research Ethics Committee.
Please report any subsequent changes that affect the ethics of the project to the University Research Ethics Committee ethics@rhul.ac.uk

This email, its contents and any attachments are intended solely for the addressee and may contain confidential information. In certain circumstances, it may also be subject to legal privilege. Any unauthorised use, disclosure, or copying is not permitted. If you have received this email in error, please notify us and immediately and permanently delete it. Any views or opinions expressed in personal emails are solely those of the author and do not necessarily represent those of Royal Holloway, University of London. It is your responsibility to ensure that this email and any attachments are virus free.

Appendix L. Ethics Application Amendment

Amendment request form

Name: Idil Kilinc, PI: Prof. Helen Pote **Department:** Psychology

Project ID: 2566 **School**: Life Sciences and the Environment

Project Title: A cluster randomised controlled trial to evaluate the effectiveness of an augmented reality (AR) board game for adolescent wellbeing

Amendment request:

☐Extension request

☐Change in team members

☒Change to participant groups

☐Change to research methods

☐Change to research summary

☐Change to data collection

☐Change to participant documents (e.g. recruitment documents, information sheet, consent form or debrief form)

Other Click here to enter text.

Details of Amendment:

(List each proposed change and its reference in the original application)

The original ethics application stated that we were intending to recruit 13-16 year olds. We would now like to extend this to include 16-18 year olds as well (13-18 year olds overall).

Please provide an explanation for the requested amendment:

Our initial proposal of recruiting from 13-16 year old students did not yield sufficient numbers for our study. We have been made aware that some A-level students may be interested in participating in research. We want to extend our inclusion criteria to include students between the ages of 16-18 to be able to contact them regarding our study.

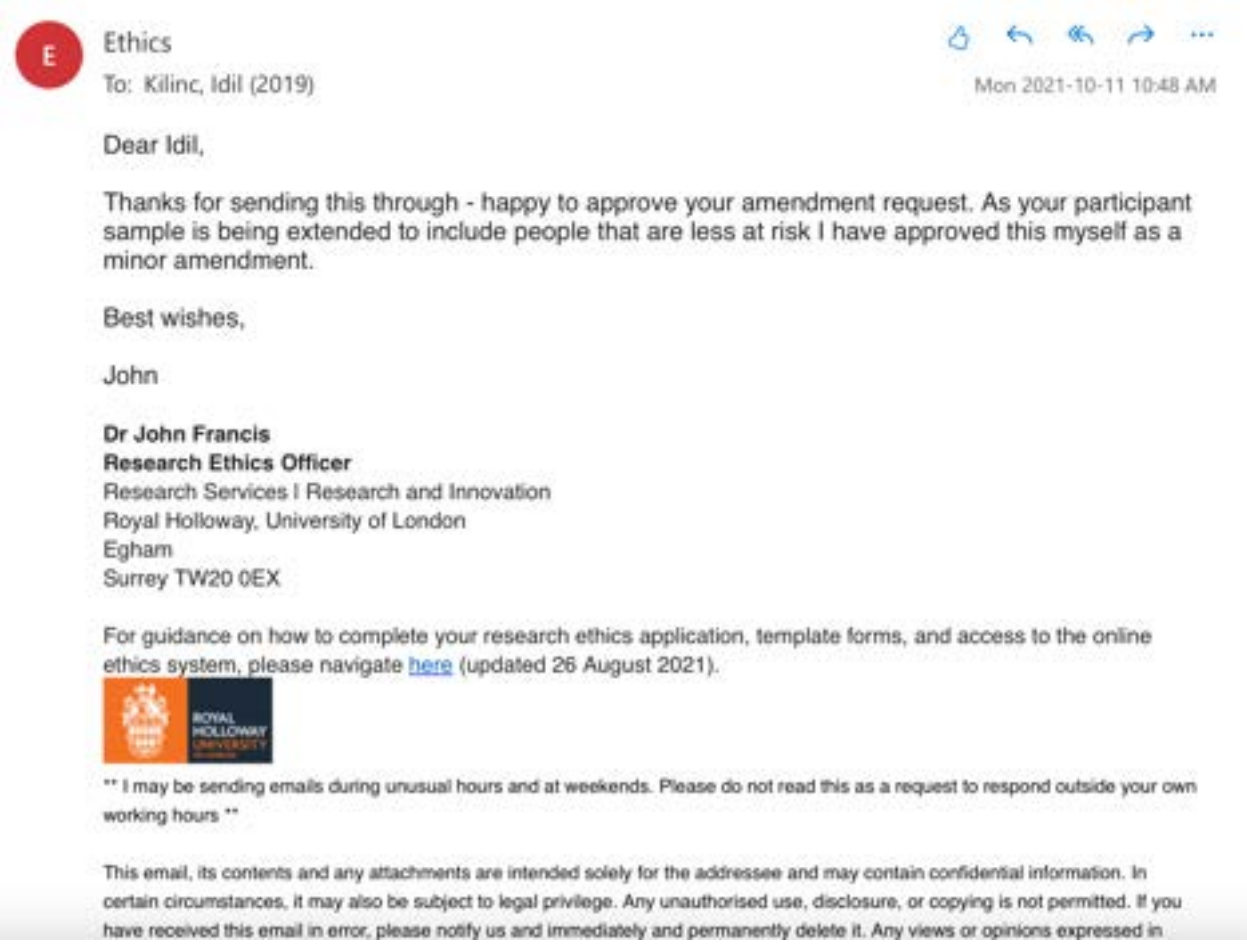
Additional information:

(Please list and attach tracked copies of amended documents)

Click here to enter text.

Signed Idil Kilinc **Date:** 08/10/2021

Appendix M. Ethics Amendment Approval



Ethics

To: Kilinc, Idil (2019)

Mon 2021-10-11 10:48 AM

Dear Idil,

Thanks for sending this through - happy to approve your amendment request. As your participant sample is being extended to include people that are less at risk I have approved this myself as a minor amendment.

Best wishes,

John

Dr John Francis
Research Ethics Officer
Research Services | Research and Innovation
Royal Holloway, University of London
Egham
Surrey TW20 0EX

For guidance on how to complete your research ethics application, template forms, and access to the online ethics system, please navigate here (updated 26 August 2021).

** I may be sending emails during unusual hours and at weekends. Please do not read this as a request to respond outside your own working hours **

This email, its contents and any attachments are intended solely for the addressee and may contain confidential information. In certain circumstances, it may also be subject to legal privilege. Any unauthorised use, disclosure, or copying is not permitted. If you have received this email in error, please notify us and immediately and permanently delete it. Any views or opinions expressed in

Appendix N. Parental Consent form for Youth Advisor Participation

Dragons of Afterlands Wellbeing Game Feedback Interview Consent Form

Dear Parent,

Your child has now completed their participation in the Dragons of Afterlands Wellbeing Game Trial during class hour. (S)he attended all sessions offered and engaged very well with the game. We would, therefore, like to invite her/him for a one-off feedback interview on their experience of participating in the project.

The Wellbeing Game Trial

We designed a board game specifically to support your teenager's wellbeing skills in a fun and engaging way using augmented reality technology. Teens played the game with their friends in class hour for 6 weeks and completed various tasks that aimed to improve their learning.

Participation in the Feedback Interview

We will be conducting one-off post-participation feedback interviews to obtain some students' views on the wellbeing game project. The interviews will be conducted verbally in person and will cover topics, such as how to improve the game, how to interpret some of the results we may find, how to disseminate the game/our findings and to discuss any future directions they think would be beneficial.

Participation in this feedback interview is entirely voluntary and those who choose to not participate will not be affected by their decision. Participants may choose to stop their participation at any time point. They do not need to provide an explanation for this and will not be affected negatively in any way. We will be offering your child an Amazon voucher worth 25 pounds for their participation in the feedback interview.

Confidentiality and Use of Collected Information

All data collected during the feedback interview will be kept confidential and will only be visible to the research team. Information will be stored anonymously on secure, password protected systems for up to 4 years. Information collected in the feedback interview will be used for the analysis and interpretation of this project as well as contribute to the development of the wellbeing game. Findings of this trial will be disseminated in scientific publications.

Who are we?

This wellbeing game trial is run by Idil Kilinc (trainee clinical psychologist at RHUL- idil.kilinc.2019@live.rhul.ac.uk) under the supervision of Prof. Helen Pote (clinical psychologist

with 20+ years of experience of working with young people & professor at RHUL- h.pote@rhul.ac.uk) and Dr. Sarah Campbell (founder & director of Play Well for Life, which develops educational games- sarah@playwellforlife.com). Please feel free to contact us if you have any questions about the feedback interviews or the trial. This trial has been reviewed and approved by the College Ethics Committee at Royal Holloway University of London.

Please tear off and hand to the class teacher if you do not want your child to participate in this feedback interview.

Participation in this feedback interview is voluntary. Participating students will be asked to verbally answer questions about their experience of the project. Their answers will be transcribed and stored anonymously on secure systems for up to 4 years and will be used to evaluate and improve the project.

I **do not want** my child to participate in this feedback interview.

Name of Student:.....

Signature of Parent:.....

Date:.....