Computers in Human Behavior 145 (2023) 107776

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

What is the reported relationship between self-esteem and gaming disorder? A systematic review and meta-analysis

Michael Kavanagh *, Charlotte Brett, Catherine Brignell

Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, SO17 1BJ, UK

ARTICLE INFO

Keywords:
Gaming disorder
Self-esteem
Systematic review
Meta-analysis

ABSTRACT

The relationship between self-esteem and gaming disorder has recently captured the interests of researchers. The aim of this systematic review is to synthesise the self-esteem and gaming disorder literature and investigate any potential variance in the size of the correlation between the two variables.

The databases PsycINFO, MEDLINE, CINAHL and Web of Science as well as grey literature was searched with backwards citation chaining for studies published between database inception and October 11, 2021. All studies that assessed the relationship between self-esteem and gaming disorder using recommended measures were included. A random effects meta-analysis for self-esteem and gaming disorder was conducted. Subgroup analysis investigating heterogeneity was conducted for: culture, age, measures of self-esteem and gaming disorder.

Of 2496 records identified, 37 were included in the review. The quality assessment indicated a moderate risk of bias. The meta-analysis indicated an association between low self-esteem and gaming disorder ($r = -0.269$ 95% CI -0.335 to −0.201). Subgroup analysis indicated this association appears to apply irrespective of culture or age. However, heterogeneity between studies can be partly explained by the self-esteem measure used, indicating investigating sub-components of self-esteem's relationship with gaming disorder would be worthwhile future research.

1. Introduction

Internet gaming disorder (IGD) was noted in the appendix of the latest Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as a condition requiring further investigation (American Psychiatric Association [APA], 2013). Since then, gaming disorder (GD) has been recognised as a diagnosis in the eleventh edition of the International Classification of Diseases (ICD-11; World Health Organisation [WHO], 2018). Whilst the DSM-5 and ICD-11 classify the disorder slightly differently, essentially both describe the disorder as: persistent gaming behaviour and impaired control over gaming, which effects social, occupational and personal functioning for a period of at least 12 months. The differences between the diagnostic categories has been described extensively elsewhere (Starcevic et al., 2020). Although this review includes studies using both definitions, from herein we will refer to the disorder as GD. A recent systematic review indicates the global prevalence rate of GD is 1.96% (Stevens et al., 2021), suggesting that for the vast majority, gaming provides positive experiences (Granic, Lobel, & Engels, 2014).

Self-esteem is an evaluation of one's self-concept by appraisals of adequacy, social comparisons and self-attributions (Rosenberg, 1989). Generally, research indicates that self-esteem is a risk factor for GD, with lower self-esteem associated with higher scores on GD measures (Leménager et al., 2020). The literature has offered theoretical explanations for self-esteem's relationship with GD using compensatory hypotheses. For example, the basic needs component of self-determination theory posits that people are motivated to engage in activities that fulfil the psychological needs of relatedness and competency which boosts self-esteem (Allen & Anderson, 2018; Bender & Gentile, 2020; Mills & Allen, 2020; Mills et al., 2018; Ryan & Deci, 2017; Ryan et al., 2006). For some people, the gaming-world potentially offers a means to satisfy these self-esteem needs when they are unsatisfied in real-life (Anderson et al., 2017). For example, in massively multiplayer online role-playing games (MMORPGs) a gamer can improve their unique avatar over time. As their avatar 'levels-up', a gamer's self-esteem may also concurrently 'level-up'/increase. With an improved avatar, the gamer can achieve better outcomes in the game and receive more self-esteem boosting praise and respect from other members of their 'guild', (a type of gaming

* Corresponding author.
E-mail addresses: m.kavanagh@soton.ac.uk (M. Kavanagh), c.brignell@soton.ac.uk (C. Brignell).

https://doi.org/10.1016/j.chb.2023.107776
Received 20 June 2022; Received in revised form 8 December 2022; Accepted 2 April 2023
Available online 3 April 2023
0747-5632/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

in-group; Zhang & Fung, 2014). If for example, a person is struggling to build self-esteem by levelling-up and achieving in their 'real-world' occupation, levelling-up via an avatar in the gaming-world may be able to compensate for this self-esteem need. Considering it can take years to achieve a promotion in one's occupation, levelling-up and boosting self-esteem in the gaming-world is likely a less strenuous and more enjoyable option for achieving self-esteem (McGonigal, 2011). Therefore, it is also likely to be a more reinforcing option for building self-esteem, making a person more suspectable to addictive tendencies associated with the behaviour.

Although most studies investigating the relationship between self-esteem and GD find a significant association, the reported effect size has varied between studies (Scerri et al., 2019; Zhou & Leung, 2012). An important goal of a synthesis is not just to report an overall effect size, but to investigate and quantify the extent of any effect size variance and consider the implications (Borenstein, 2009). For example, in the context of this review, investigating differences in the effect size of the GD and self-esteem relationship between studies can elucidate which populations are more at risk and thus interventions can be more efficiently targeted. Meta-analyses can provide an overall summary effect which helps to determine if self-esteem is a risk factor worth targeting in GD interventions (Borenstein, 2009). Importantly, a key strength of a meta-analysis is that it provides a full complement of statistical tools to assess the variance of effect sizes between studies (Borenstein & Higgins, 2013). By contrast, without such statistical analyses, a narrative reviewer is limited in their ability to accurately assess the pattern of variance and risks biases of interpretation through methods like p-value vote counting (Borenstein, 2009). Therefore, this study aimed to use the statistical tools of a meta-analysis to understand the summary effect, but importantly to also better quantify the extent of any effect size variance and consider the implications of this variance.

Thus far, no systematic review has been conducted with the primary aim of synthesising studies investigating the relationship between gaming disorder and self-esteem. Nor has a review investigated what factors might explain the variation in the association. Such an investigation could elucidate if self-esteem is a risk factor worth targeting in GD interventions, but also identify which populations are more susceptible to larger effect sizes so that interventions can be more efficiently targeted.

It is worth noting that Leménager et al. (2020) carried out a systematic review with the primary aim of understanding self-concept clarity's relationship with avatar identification. Within this, they included a secondary search of self-esteem and GD. However, the self-esteem and GD search strategy was limited to data bases of PubMed and PsycINFO which missed several key articles. Additionally, despite noting variability in effect sizes reported, the study did not investigate this dispersion through meta-analysis and subgroup analysis (although this is understandable considering it was not the primary aim of their review). Furthermore, Green et al. (2020) carried out a review synthesising avatar and self-related processes in relation to GD symptomatology. However, similarly to Leménager et al. (2020), investigating the relationship between GD and self-esteem was not the primary focus of the review as GD and self-esteem articles were excluded if they did not make reference to game avatars. Similarly, again, their review did not include a meta-analysis or subgroup analysis investigating variance in effect sizes.

Some research suggests that players of massively multiplayer online role-playing games (MMORPGs) are particularly vulnerable to GD (King et al., 2019). Additionally, studies have found that MMORPG players report lower self-esteem scores compared to players of other game genres (Stetina et al., 2011). Therefore, game genre may explain variability in effect sizes between GD and self-esteem, thus, this review aims to compare game genre subgroups. Considering studies indicate players of MMORPG games report both lower self-esteem and are more vulnerable to GD; it may be that the effect size in the relationship between the self-esteem and GD is also larger for MMORPG players.

The review also aims to compare effect size scores in Eastern vs Western countries as gaming disorder has been reported as a particular concern in Asian countries (Zastrow, 2017).

National prevalence rates of GD are estimated to be 10%–15% in Eastern countries and 1%–10% in Western countries (King et al., 2012; Saunders et al., 2017; Stevens et al., 2021). Researchers have hypothesised that the reason for prevalence differences could be due to how Eastern and Western cultures address GD, with Eastern countries more commonly viewing GD as a public health concern and employing regulations to reduce GD compared to Western countries (Király et al., 2018). Despite regulatory initiatives, gaming is more popular in Eastern countries (Buchholz, 2021; Stevens et al., 2021). Due to these Eastern vs Western differences in their recent systematic review of global prevalence rates of GD, Stevens et al. (2021, p. 565) concluded that "there is a need for cross-cultural research on GD, comparing Eastern and Western countries". Additionally, the self-esteem literature reports that Western countries score higher for self-esteem than Eastern countries (Errasti et al., 2018; Hofstede, 1984). It has been proposed that Westerners assign more importance to competence-based self-esteem (feeling that one is capable and efficacious; Tafarodi & Milne, 2002) whereas those in the East, assign more value to modesty (Chiu & Hong, 2006), being relationally skilful and accepted by others (Sedikides et al., 2015). Considering these differences for both self-esteem and gaming disorder between East vs West; it may be that the effect size in the relationship between the two variables also differs between East vs West.

The review will also compare individualistic and collectivistic countries. Although there is an overlap in comparing Eastern vs Western countries and collectivist vs individualistic countries, some studies have suggested that there are no differences in global self-esteem scores between collectivist vs individualistic countries (Schmitt & Allik, 2005).

Resilience research indicates that high self-esteem buffers against a person's vulnerability to stress (Thoits, 1994) substance abuse (Mann et al., 2004) and internet addiction more generally (Zhang et al., 2015). Therefore, high self-esteem may buffer against risk factors for GD. Again, understanding these moderating relationships could improve the efficiency of GD interventions by potentially elucidating a need to target self-esteem. Self-esteem may also be a mediating variable in explaining other risk factor's relationship with GD. Outside of the GD literature there is evidence that supports a "vulnerability hypothesis" where low self-esteem contributes to anxiety and depression (Orth & Robins, 2013). Considering these points, the review also aims to synthesise research that investigates mediating and moderating factors of self-esteem's relationship with GD.

1.1. Aims of the current review

The aim of this review is to synthesise the grey and published literature that has examined the relationship between self-esteem and GD. The primary research question is: what is the reported relationship between self-esteem and GD? Secondary research questions include: (i) does the relationship between self-esteem and GD vary based on game genre? (ii) Does the relationship between self-esteem and GD vary between Eastern vs Western and collectivist vs individualistic countries? (iii) What variables moderate and mediate the relationship between self-esteem and GD?

2. Methods

2.1. Protocol and registration

The review was conducted in line with the principles recommended by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses updated in 2020 (PRISMA; Page et al., 2021). The review protocol was registered on the international prospective register of systematic reviews (PROSPERO) in August 2021 post-initial scoping search and prior to the systematic search. This protocol can be accessed

at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=268744 registration number: CRD42021268744.

2.2. Eligibility and inclusion criteria

There were six inclusion or exclusion criteria, two regarding study type, two regarding measures needing to be included, one related to statistical results reported and one distinguishing gaming from gambling. Studies were eligible for inclusion if they were quantitative and original empirical studies. There were no exclusion criteria for age and gender as the criteria for GD applies equally to all gender and age groups. Studies of the general population, clinical samples and groups of specific gamers were all included. However, gambling as a specific game genre was excluded. There were no limits set on publication date. Studies were required to include a self-esteem measure.

Since the inclusion of IGD in the appendix of the DSM-5, research in the GD field accelerated and many new GD measurement tools have been developed (King, Russell, et al., 2020). Previous reviews on GD instrumentation have reported various inconsistencies and psychometric weaknesses (Griffiths et al., 2014; King, Russell, et al., 2020). Many tools have demonstrated inconsistent cut-off scores and symptom coverage, as well as inadequate data on predictive validity and inter-rater reliability (King, Russell, et al., 2020). In this review, only studies using the five tools that according to King et al.'s (2020) systematic review of assessment tools for GD have the best support for their psychometric properties (plus five others) were included: Assessment of Internet and Computer Addiction Scale-Gaming (AICA-Sgaming; Wölfling et al., 2012), Game Addiction Scale 7-items (GAS-7; Lemmens et al., 2009), Internet Gaming Disorder Test 10-items (IGD-10; Király et al., 2017), Nine-Item Internet Gaming Disorder Scale–Short Form (IGDS9-SF; Pontes & Griffiths, 2015), Internet Gaming Disorder Scale 9-items (IGD-9, Lemmens et al., 2015). Two of the five additional tools included were the long form versions of the already included GAS-7 (The Game Addiction Scale 21; GAS-21; Lemmens et al., 2009), and Lemmens IGD-9 (Internet Gaming Disorder Scale-27; IGD-27; Lemmens et al., 2015). The Petry Internet Gaming Disorder Tool (Petry-10; Petry et al., 2014) was also included. The Petry-10 has been administered to a large number of participants and provides coverage of all DSM-5 and ICD-11 criteria with good support for its psychometric properties. The Young Internet Addiction Test (YIAT; Young, 1998) and the Young Diagnostic Questionnaire (YDQ; Young, 1998) were also included even though they precede the recognition of IGD in the DSM-5 (2013). The YIAT and YDQ were included for greater cultural representation. King, Chamberlain, et al. (2020) noted the YIAT and YDQ to be the most cited tools in the GD literature and most commonly used in East Asian studies. It is important to note that the YIAT and YDQ were originally designed to measure internet addiction rather than GD. Therefore, this review will only consider studies that adapt the YIAT and YDQ for use in gaming samples. Studies that measure GD based on a clinical interview with a mental health professional were also included.

Finally, to be included, the study needed to report a Pearson's *r* statistic between self-esteem and GD or report analysis that could be converted to a Pearson's *r*.

2.3. Adaptions to the registered protocol

At the protocol stage, the plan was to limit the search to English-language studies. However, as there were a number of studies retrieved at both the abstract and full-text screening stage that were published in a language other than English, this decision was reviewed. Cochrane advocate that searches are not restricted by language (Higgins et al., 2019) and the Campbell Collaboration states: "ideally no language restrictions should be included in the search strategy" (Kugley et al., 2016, p. 28). Previously, Balk et al. (2013) reported that using Google Translate can be an effective approach to reducing language bias, but also concluded that reviewers should be cautious using data from Google translated articles. However, since Balk et al.'s (2013) study, the accuracy of machine translators has improved due to a greater quantity of data and developments in artificial intelligence (Walpole, 2019). In a more recent study, Jackson et al. (2019) using Google Translate for systematic reviews concluded that: Google Translate is a viable, accurate tool for translating non–English-language studies for the purpose of conducting systematic reviews and excluding such studies may lead to substantial bias, particularly if non–English-language studies tend to have a higher proportion of nonsignificant outcomes (p 679). Since then, other reviews have registered their protocols planning to translate non-English-language articles using Google Translate (Clephas et al., 2021; McGuier et al., 2021) and the same was decided for this review. Where an article was published in a non-English-language, an e-mail request for a possible English-language version of the article was sent to the author. However, no author who responded had an English-language version of their paper.

The original protocol also excluded offline gaming. However, it was subsequently decided to include both offline and online gaming studies for two reasons. Firstly, although the disorder is labelled "internet" gaming disorder in the DSM-5, the DSM-5 also states that "internet gaming disorder most often involves specific internet games, but it could involve non-internet computerized games as well, although these have been less researched" (APA, 2013, p. 796). Additionally, the ICD-11 includes both offline and online gaming in their description of GD (WHO, 2018). Secondly, during the systematic search, it became clear that many studies did not clearly specify whether they assessed online or offline gaming only.

2.4. Search strategy

An initial scoping search was performed in July 2021, prior to the finalisation of the review question and protocol. The scoping search was performed using the search terms "gaming disorder" and "self-esteem" conducted via PROSPERO, Cochrane Library, Google Scholar (first 100 pages), Trip medical data base, Thesis Commons, ProQuest Dissertations and Theses, OpenGrey, MEDLINE, PsychINFO, Web of Science and CINAHL. The search indicated the breath of the literature, identified existing reviews and protocols and informed the final search strategy.

A systematic literature search was then conducted through: PsycINFO, MEDLINE CINAHL Plus with Full Text (all accessed via EBSCO), Web of Science Clarivate (via Web of science core Collection) and PubMed. A systematic grey literature search was also conducted via Thesis Common, ProQuest Dissertation and Thesis Global, OpenGrey and Google Scholar. The search dates, subject heading and syntax used in each platform and data base is attached in the supplementary materials (table A). Both MEDLINE and PubMed were used as PubMed publishes articles earlier than MEDLINE. Only the first 100 articles of Google Scholar listed as "sort by relevance" were included. A similar process is used in other published systematic reviews due to Google Scholar's propensity to list a substantial number of articles (Figueroa et al., 2018). In line with the peer-review of electronic search strategy guidelines (McGowan et al., 2016) a librarian was consulted on the use of Boolean operators, truncation and to refine the search strategy. No limiters (field, language or date restrictions) were applied to the search strategy.

2.5. Screening process

Studies returned from the searches were exported to the reference management software, EndNote. Following removal of duplicates, the remaining studies were exported to Rayyan, a free web tool designed to help researchers screen studies during systematic reviews. The abstracts were screened independently by two researchers against the inclusion and exclusion criteria. For full-text screening, one reviewer screened all eligible studies and a second reviewer independently screened four (11%) randomly selected studies due to the large volume of eligible articles, as recommended by Boland et al. (2017).

2.6. Data extraction

One reviewer independently extracted relevant data from each study including: the study design, country the study was conducted in, where participants were recruited from, participant characteristics (age, gender, socio-economic status, education, ethnicity and religion), game type played, measures for GD and self-esteem, language the study was published in, self-esteem and GD correlation statistic, mediation or moderation analysis, longitudinal results and specific game type results. Due to the large volume of studies identified, a second reviewer independently extracted data from four (11%) randomly selected eligible studies (Boland et al., 2017).

2.7. Quality assessment for risk of bias

As this review investigated the correlation between self-esteem and GD, a quality assessment of the included studies was conducted using the National Institute of Clinical Excellence (NICE) Quality appraisal checklist for quantitative studies reporting correlations and associations (NICE, 2012). The checklist consists of 17 questions relating to the external and internal validity of a study. The checklist was adapted for use in the current review where one question 'is the setting applicable to the UK?' was deemed irrelevant and removed. Each item on the checklist is scored as high ("for that particular aspect of study design, the study has been conducted/designed to minimise the risk of bias"), medium ("the answer to the checklist question is not clear or the study may not have addressed potential bias for that aspect of the study design") or low ("reserved for aspects of the study design where significant bias may persist"). As the scoring guidance does not concretely describe what should be considered a high, medium or low quality score, if there was uncertainty whether a score fell between "high" or "medium" for example, it was reported as "medium/high". A similar approach has been used in other published studies (Dixon et al., 2018). The tool finally reports a summary score for both internal and external validity. Again, due to the volume of studies identified, one reviewer independently assessed all eligible studies and a second reviewer independently

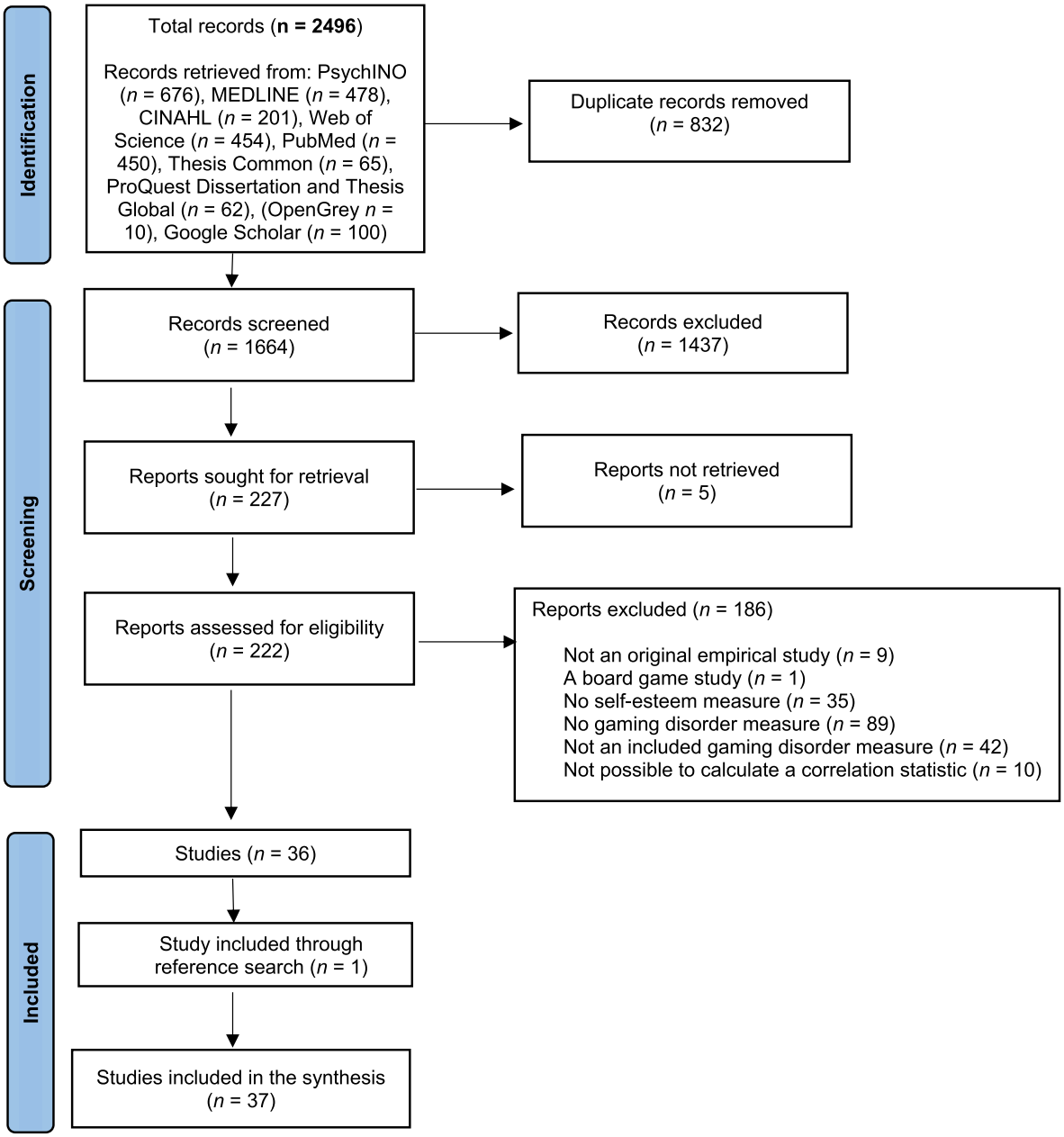

Fig. 1. PRISMA flow diagram depicting the process of study selection (Page et al., 2021).

assessed the quality of four (11%) randomly selected eligible studies.

2.8. Methods of data synthesis and analysis

As no study scored low for internal or external validity, all studies were included in the review. Study results were narratively synthesised. The software Comprehensive Meta-Analysis (version 3.0, Borenstein et al., 2013) was used to conduct a meta-analysis. This software as well as the guidance for converting effect sizes by Borenstein et al. (2009) was used to transform effect sizes not reported as a Pearson's *r*. As the studies included in the meta-analysis were likely heterogeneous, varying due to the nature of the sample and measures used, a random-effects model was adopted. A random-effects model can indicate an overall mean effect but also provide detail on the variation in effects.

3. Results

3.1. Screening

Fig. 1 displays the PRISMA flow diagram of the study selection process. A total of 2496 studies were returned from the searches and 1664 remained following the removal of duplicates. At the abstract screening stage, there was disagreement between the two reviewers on 198 studies (12%), which is a reasonable disagreement rate according to Polanin et al. (2019). The reason for the majority of disagreement was that one reviewer included internet addiction abstracts even if GD was not referred to. The reviewer's rationale was that GD fell under the heading of internet addiction in some databases and therefore a GD measure may have been included in the full-text article. The other reviewer agreed with this rationale post-conflict discussion. Post-abstract screening, 227 studies were eligible for full-text screening. There were no cases of disagreement between the two reviewers at this stage. It was not possible to obtain the full-text articles of Wei and Luo (2017), Cho and Kwon (2016), Kadam and Himanshi (2020), Liang (2011) and Yu (2020). Thirty-six studies met the full inclusion criteria. A backwards citation search of the reference lists for all 36 eligible studies was conducted which led to one additional study being included (Park et al., 2007). It was not possible to obtain the full-text articles of Cho and Lim (2010), Ko (2008), Nagygyörgy et al. (2012, pp. 242–248) and Peng and Li (2009) during the backwards citation search.

Studies that included a self-esteem and GD measure but reported their analysis in a way that meant it was not possible to convert the data to a correlation statistic were contacted and asked to provide such data. One study's author replied with the requested data (Fernandes et al., 2021).

3.2. Reasons certain studies were excluded

It is recommended by PRISMA to specify why certain studies that may appear to meet the inclusion criteria were excluded (Page et al., 2021). King and Delfabbro (2016) developed and used the 24-item Internet Gaming Cognition Scale in their study which includes two items assessing gaming cognition self-esteem. However, as these two items measure self-esteem relative to gaming only, studies that used this scale to measure self-esteem (King & Delfabbro, 2016; Moudiab & Spada, 2019; Yu et al., 2020) were excluded as this review aimed to assess a person's global self-esteem and not self-esteem contingent on one niche area of life. For the same reason, studies that used Beard and Wickham's (2016) Gaming Contingent Self-Worth scale as a measure of self-esteem were excluded (King, Russell, et al., 2020). Throuvala et al. (2019) used the Core Self-Evaluations Scale (CSES; Judge et al., 2003) to measure self-esteem. However, the CSES groups self-esteem, generalized self-efficacy, locus of control and neuroticism together and reports one grouped scale score rather than a sub-scale score of self-esteem. Similarly, The Flourishing Scale (Diener et al., 2010) used by Molinos (2016) groups items of success in relationships, self-esteem, purpose and optimism rather than reporting sub-scale scores.

3.3. Data extracted

A summary of the study characteristics is available in the supplementary materials (table B) and Table 1 below displays the study results. There were no discrepancies in the data extracted between the two reviewers.

3.4. Study characteristics

The 37 studies were published between 2007 and 2021. Other than one unpublished thesis (Law, 2018), all studies were peer-reviewed published articles. All studies used a cross-sectional design. The studies were conducted in a variety of countries: seven in South Korea, six in Germany, four in China, four in Turkey, three in the Netherlands, two in France and one in each of: Norway, Australia, India, Hong Kong, Spain, Croatia, Canada and the United States. One paper reported conducting the study across UK, Mexico, India, Philippines and Malaysia (Fernandes et al., 2020), another conducting the study across India, Mexico, Philippines and Turkey (Fernandes et al., 2021) and finally one recruiting in Asian countries (Aggarwal et al., 2020).

The recruitment setting(s) differed between the studies: 10 studies recruited students from schools, six recruited from cohort studies, four from universities, two from online gaming forums, one recruited medical interns on placement, one recruited from social media and one recruited through a marketing company. Three studies recruited from both online gaming forums and social media sites, two studies recruited from both social media sites and student groups, two studies recruited from online gaming forums and counselling centres, one study recruited from Facebook and a university, one recruited from gaming forums and gaming sites, one recruited from gaming forums and a university and one recruited from universities, social media, schools, Christian churches and a tutoring centre. One study did not indicate where they recruited their participants from (Leménager et al., 2013). All studies recruited a non-clinical sample, with the caveat that Leménager et al. (2013) may have recruited a clinical sample.

Some of the included studies overlapped, publishing data on the same set of participants. All four studies of Wartberg and Kammerl (2020), Wartberg et al. (2017), Wartberg et al. (2019) and Wartberg et al. (2021) included the same set of participants as did the two studies by Baysak et al. (2018) and Baysak et al. (2020) and the two studies Kim (2019) and Kim and Ko (2020). Not double counting the overlapping participants, the total sample size of the included studies was $N = 22,142$. The number of participants in each study varied between 37 (Aggarwal et al., 2020) and 2894 (van Rooij et al., 2012). Not double counting the overlapping participants and not including Bargeron and Hormes (2017) who did not report gender data or Fernandes et al. (2021) and Jeong et al. (2018) who both reported gender data but not based on the participants included the analysis; across the studies there were 11,721 males, 8985 females and five gender not reported participants. Not double counting the overlapping participants, a pooled mean age of 18.29 ($SD = 3.43$) was calculated which excluded studies that did not report data on age (Bilic & Golub, 2011; Jeong et al., 2018; Park et al., 2007) that did not report a *SD* (Aggarwal et al., 2020; Biegun et al., 2021; Kim et al., 2017; Law, 2018; Raja et al., 2020) or reported age range only (Teng et al., 2020; You et al., 2017; Zhou & Leung, 2012).

Nine studies did not report data on education status (Aggarwal et al., 2020; Biegun et al., 2021; Kim, 2019; Kim & Ko, 2020; Kircaburun et al., 2019; Laconi et al., 2015; Lemmens et al., 2015; Scerri et al., 2019; Von der Heiden et al., 2019). Ten studies reported data on socio-economic status or employment (Baysak et al., 2018, 2020; Bilic & Golub, 2011; Buiza-Aguado et al., 2018; Fernandes et al., 2021; Kim et al., 2017; Laconi et al., 2017; Law, 2018; Leménager et al., 2013; Wichstrøm et al., 2019). Two studies reported data on religion and a breakdown of religious subgroups (Fernandes et al., 2021; Law, 2018). Four studies

Table 1
Summary description of study results.

Study (Year)	Effect Size Need Transforming	Effect Size (p) and other relevant findings	Specific Gaming Genre Effect Size	Mediators and Moderators	Quality Assessment Score for Internal Validity	Quality Assessment Score for External Validity
Aggarwal et al. (2020)	No	$r = -.107$ ($p = .54$)	MOBA, $r = -.107$ ($p = .54$)		+	+
Bargeron and Hormes (2017)	Yes, converted from an ANOVA where time spent gaming was controlled ($\eta p^2 = 0.01$)	$r = -.1$ ($p = .11$)			-/+	-/+
Baysak et al. (2018)	No	($r = -.333$, $p < .01$). Hierarchical linear modeling indicated that self-esteem was not significantly associated with lowered GD scores over two years	(strategy game) ($r = -.333$, $p < .01$)		+	++
Baysak et al. (2020)	No	$r = -.333$, ($p < .01$)	(strategy game) ($r = -.333$, $p < .01$)		+	+
Biegun et al. (2021)	No	$r = -.204$ ($p < .01$)			+	-/+
Bilic and Golub (2011)	Yes, pooled the effect size for male and female r	$r = -.266$ ($p < .001$). Male: $r = -.310$ ($p < .001$), Female: $r = -.221$ ($p < .01$)		Educational environment moderates the relationship between self-esteem and GD. In an ANOVA of gender x self-esteem x school type, found that males in grammar schools with low self-esteem are more likely to have higher GD scores $F = 4.284$, $p = .039$ $\eta p^2 = 0.013$	++	+/++
Buiza-Aguado et al. (2018)	Yes, converted from Cohen's d comparing IGD vs non-IGD cut off scores ($d = .33$)	$r = -.163$ ($p < .001$)			+	+/++
Ekşi et al. (2020)	No	$r = -.204$ ($p < .01$)		GD predicted indirectly predicted well-being through self-esteem and family harmony $\beta = -.22$, $p < .001$, 95% CI $= -.37$ to .07.	+	+
Fernandes et al. (2020)	No	$r = -.14$ ($p = .057$)			+/++	+
Fernandes et al. (2021)	No	$r = -.092$, ($p = .002$)			+/++	+/++
Jeong et al. (2021)	No	$r = -.09$ ($p < .01$) Self-esteem at time-point two predicted negative pathological gaming at time-point 3 (6-month time difference) = $\beta = -.12$, 95%		Self-esteem's relationship with GD between timepoint two and three is significantly ($p < .001$) influenced by parental environment ($\beta = .40$), school environment ($\beta = .28$) peer environment ($\beta = .20$)	-/+	-/+
Jeong et al. (2018)	Yes, converted from Cohen's d ($d = 1.042$)	$r = -.462$ ($p < .001$)			+	+
Kim and Ko (2020)	No	$r = -.31$ ($p < .01$)			+	+
Kim (2019)	No	$r = -.34$ ($p < .01$)			+	+
Kim et al. (2017)	Yes, pooled the effect size for male and female β	$r = -.255$ ($p = <.001$) Low self-esteem predicted GD for both males ($\beta = -.23$, $p < .001$) and females, ($\beta = -.18$, $p < .001$)		GD had a full mediating effect on the path from self-esteem to school violence perpetration	+/++	+
Kircaburun et al. (2019)	No	$r = -.26$ ($p < .001$)		Investigated how childhood emotional trauma predicts GD with multiple mediators of self-esteem, depression, body image satisfaction, loneliness and social anxiety. Self-esteem was not a significant mediator in the model, but depression was	+/++	+/++
Laconi et al. (2017)	No	$r = -.33$ ($p < .01$)	Mean self-esteem scores for: casual gamers = 29.33 (SD		-/+	+/++

(*continued on next page*)

Table 1 (*continued*)

Study (Year)	Effect Size Need Transforming	Effect Size (p) and other relevant findings	Specific Gaming Genre Effect Size	Mediators and Moderators	Quality Assessment Score for Internal Validity	Quality Assessment Score for External Validity
			= 4.3), MMORPG = 28.68 (*SD* = 4.7) MOBA = 28.68 (*SD* = 4.7), FPS = 28.61 (*SD* = 4.5), RTS = 28.23 (*SD* = 4.5), Other games = 28.14 (*SD* = 5.7), Action and adventure = 25.79 (*SD* = 6.1). ANOVA indicated the action and adventure genre players had significantly less self-esteem, $F = 2.384$ ($p = .028$)			
Laconi et al. (2015)	Yes, pooled the effect size for r of adults aged 18–30 ($r = -.26$ $n = 328$) and adults aged 31–65 ($r = -.49$ $n = 50$).	$r = -.38$ ($p < .001$) For adults aged 18–30 ($n = 328$) $r = -.26$ For adults aged 31–65 ($n = 50$) $r = -.49$ For females $r = -.34$, effect size not reported for males			+	-/+
Law (2018)	No	$r = -.253$, ($p < .001$), when controlled for age and gender $r = -.268$ ($p < .001$)			+	+
Leménager et al. (2013)	Yes, converted from Z statistic for comparisons of cut off scores for addicted and non-addicted gamers ($Z = -4.644$)	$r = -.69$ ($p < .001$)			-/+	-/+
Lemmens et al. (2015)	No	9-item Dichotomous scale r = -.20 ($p < .001$) 27-item Dichotomous scale $r = -.21$ ($p < .001$) 9-item Polytomous scale $r = -.1$ ($p = .002$) 27-item Polytomous scale $r = -.11$ ($p = <.001$)			++	+/++
Lemmens et al. (2011)	No	Wave 1 $r = -.13$ ($p < .01$), Wave 2 $r = -.17$ ($p < .001$), Autoregressive structural equation models indicated that self-esteem at wave 1 was a significant predictor GD at wave 2 (6-month time difference) ($\beta = .10$, $b = .14$, SE = .066, $p < .05$) and pathological gaming at wave 1 showed a non-significant effect on self-esteem in wave 2 ($\beta = .05$, $b = .05$, SE = .044, $p = .22$).			+	+
Raja et al. (2020)	Yes, transformed from $\beta = -.098$	$r = -.148$ ($p = .05$)			-/+	+
Park et al. (2007)	Yes, transformed from $t = -34.62$	$r = -.81$ ($p < .001$)			+/++	+
Scerri et al. (2019)	No	$r = -0.43$ ($p < .001$)		Self-esteem through depression, significantly mediated the effect deficits in needs satisfaction has in predicting GD. The total standardised indirect effect was portioned as: deficits in need-satisfaction to self-esteem (−0.76) to depression (−0.46) to GD (.29)	+	-/+
Teng et al. (2020)	No	T1 $r = -0.17$ ($p < .01$) T2 $r = -0.28$ ($p < .01$) T3 $r = -0.31$ ($p < .01$) Cross-lagged panel models indicated that IGD significantly predicts self-esteem from T1 to T2 (six-month time difference), ($\beta =$			+/++	-/+

(*continued on next page*)

Table 1 (*continued*)

Study (Year)	Effect Size Need Transforming	Effect Size (p) and other relevant findings	Specific Gaming Genre Effect Size	Mediators and Moderators	Quality Assessment Score for Internal Validity	Quality Assessment Score for External Validity
		−.06, 95% CIs [-.09, −.02] and T2 to T3, (β = −.06, 95% CIs [-.10, −.02]). However, self-esteem did not significantly predict IGD from T1 to T2, (β = .03, 95% CIs [-.06, .01]) or T2 to T3, (β = −.03, 95% CIs [-.06, .01]).				
van Rooij et al. (2012)	No	$r = -.18$ ($p < .001$)			-/+	+
Von der Heiden et al. (2019)	No	$r = -0.27$ ($p < .001$),	Partial correlations, controlling for sex and age, showed the following relationships for game type and self-esteem: role-play games = −.12 (<.001) and non-significant relationships for simulation = −.01, strategy = 0, Action = −.03 and Unclassified games = 0.		+/++	+/++
Wang et al. (2020)	Yes, transformed from Cohen's d ($d = -.62$)	$r = -.296$ ($p = .038$)			-/+	-/+
Wartberg and Kammerl (2020)	Yes, transformed from β = −.1	$r = -.15$ ($p < .001$)			+/++	+
Wartberg et al. (2017)	No	$r = -.39$ ($p < .01$)			+	-/+
Wartberg et al. (2019)	No	At time point 1, $r = -.39$ ($p < .01$), At time point 2 $r = -.38$ ($p < .01$). The structural equation model analysis found that self-esteem problems at time point 1 significantly predicted GD at time point 2 (6-month time difference) standardised beta = .06 ($p = <.05$). However, GD at time point 1 was not a significant predictor of self-esteem at time point 2			-/+	-/+
Wartberg et al. (2021)	No	At time point 1, $r = -.39$ ($p < .01$) At time point 2, $r = -.38$ ($p < .01$)			-/+	-/+
Wichstrøm et al. (2019)	Yes, transformed from β = 0	$r = +0.05$ ($p = .174$) Note the correlation statistic is self-esteem at aged 8 predicting GD at aged 10 adjusting for gender and gaming time at age 8.			++	+/++
Yang et al. (2021)	Yes, transformed from an Odds Ratio of .93	$r = -0.02$ ($p = .78$)			+	-/+
You et al. (2017)	No	$r = -.17$ ($p < .05$)		Investigating whether the relationship between psychosocial factors (including self-esteem, social skills and depression) and GD was mediated by avatar identification found that avatar identification did not mediate the relationship between self-esteem and GD.	+	-/+
Zhou and Leung (2012)	No	$r = +.08$ ($p = 140$)			++	+/++

Note: GD = Gaming Disorder, MMORPG: Massively Multiplayer Online Role Playing Games, MOBA: Multiplayer Online Battle Arena, RTS: Real Time Strategy, FPS: First Person Shooter, ANOVA: Analysis of Variance.

++ = high "for that particular aspect of study design, the study has been conducted/designed to minimise the risk of bias"; + "the answer to the checklist question is not clear or the study may not have addressed potential bias for that aspect of the study design"; - "reserved for aspects of the study design where significant bias may persist".

reported data on ethnicity, with two studies reporting a breakdown of ethnic subgroups (Fernandes et al., 2021; Wichstrøm et al., 2019), one study reporting data on Western European vs non-Western European participants (van Rooij et al., 2012) and one study noting that their participants were all of Han ethnicity (Yang et al., 2021). Eighteen studies collected data for online gamers only. Six studies included both online and offline gamers. Thirteen studies did not declare if they included online or offline gamers.

Eleven studies reported data on the types of games played by participants. Kircaburun et al. (2019), Laconi et al. (2017), Raja et al. (2020) and Von der Heiden et al. (2019) reported a breakdown of the different game genres their participants played. Buiza-Aguado et al. (2018) specifically reported the number of MMORPG players in their sample. Leménager et al. (2013) and You et al. (2017) participants were MMORPG players. Baysak et al. (2018), Baysak et al. (2018) and Zhou and Leung (2012) participants were strategy players. Aggarwal et al. (2020) and Wang et al. (2020) participants were MOBA players. Two studies collected data on participant's game genre but did not report the results (Jeong et al., 2018, van Rooij et al., 2012).

In measuring GD, eight studies used the GD version of the YIAT, seven studies used the GAS-7, six studies used the IGD-9, four the IGDS9-SF, four the Petry-10, three used a clinical interview, two studies used the GAS-21, two studies the IGD-10 and one study the AICA-Sgaming. In measuring self-esteem, 20 studies used the Rosenberg 10-item self-esteem scale (RSES; Rosenberg, 1965). Two studies adapted the RSES to 9-items, two adapted the scale to 5-items, two to 3-items and one to 6-items. However, only one study provided justification for adapting the RSES (Teng et al., 2020) noting that the item "I wish I could have more respect for myself'' was unsuitable for Chinese culture. Four studies used the Reynolds Adolescent Adjustment Screening Inventory self-esteem sub-scale (Reynolds, 2001), three studies used the Single-Item Self--Esteem Scale (Robins et al., 2001), one study used the State Self-Esteem Scale (Heatherton & Polivy, 1991) and one study used The Lifespan Self-Esteem Scale (Harris et al., 2018).

Thirty-four studies were published in English. Baysak et al. (2020) was published in Turkish, Bilic and Golub (2011) in Croatian and Park et al. (2007) in Korean.

3.5. Study results

Thirty-one studies reported a significant correlation between lower self-esteem and GD. Four studies reported a non-significant correlation between lower self-esteem and GD (Aggarwal et al., 2020; Bargeron & Hormes, 2017; Fernandes et al., 2020; Yang et al., 2021) and two studies reported a non-significant correlation between higher self-esteem and GD (Wichstrøm et al., 2019; Zhou & Leung, 2012). Data in 13 studies needed to be converted to a Pearson's *r* value. Two study's reported correlation statistic had other variables controlled for in the analysis. Bargeron and Hormes (2017) controlled for time spent gaming and Wichstrøm et al. (2019) controlled for gender and gaming time.

There were six longitudinal studies reported. Three studies indicated that lower self-esteem at time-point one predicted GD at time-point two (Jeong et al., 2021; Lemmens et al., 2011; Wartberg et al., 2019). Additionally, Lemmens et al. (2011) and Wartberg et al. (2019) found that GD at time-point one did not significantly predict low self-esteem at time-point two. However, in contrast, three studies (Baysak et al., 2018; Teng et al., 2020; Wichstrøm et al., 2019) found low self-esteem at time-point one did not significantly predict GD at time-point two. Additionally, Teng et al. (2020) found that GD at time-point one significantly predicted low self-esteem at time-point two.

Seven studies reported data on the relationship between self-esteem and GD for a specific game type. There was variation in effect size scores within the game genre studies. For example, Leménager et al. (2013), found a much larger negative effect (−.69) for their MMORPG sample than You et al. (2017) (−0.17). Aggarwal et al. (2020) found a smaller negative effect size (−0.11) for their MOBA sample compared to Wang et al. (2020) (−0.3). Finally, Zhou and Leung (2012) found a small non-significant positive effect size (0.08) for their online strategy sample compared to Baysak et al. (2018) and Baysak et al. (2020) (−0.33). Whilst five other studies reported game genre data, they did not report a correlation statistic between GD and self-esteem for any specific game genre (Buiza-Aguado et al., 2018; Kircaburun et al., 2019; Laconi et al., 2017; Raja et al., 2020; Von der Heiden et al., 2019). However, Laconi et al. (2017), reported a one-way analysis of variance (ANOVA) comparing self-esteem scores for game genres and found players of action and adventure games had significantly lower self-esteem scores (although the ANOVA had varying sample sizes per game genre group that may have violated ANOVA assumptions). Additionally, Von der Heiden et al. (2019) carried out a partial correlation analysis (controlling for sex and age) and found a significant relationship between negative self-esteem and role-playing games but no significant self-esteem relationship for other game genres.

Six studies reported a mediation or moderation analysis. Bilic and Golub (2011) found that educational environment and gender moderates the relationship between self-esteem and GD with males in grammar schools with low self-esteem scoring higher for GD. Ekşi et al. (2020) found self-esteem partially mediated GD's relationship with family harmony. Kim et al. (2017) reported that GD had a full mediating effect on self-esteem's relationship with school violence perpetration. In their multiple mediation model, Kircaburun et al. (2019) found that self-esteem was not a significant mediator of childhood emotional traumas relationship with GD. In another multiple mediation model Scerri et al. (2019) found that self-esteem through depression, significantly mediated the relationship between deficits in needs satisfaction and GD. You et al. (2017) in their multiple mediation model found that avatar identification did not mediate the relationship between self-esteem and GD.

3.6. Results of quality assessment

There was 73% agreement between the two reviewers on each item of the quality assessment tool and conflicts were discussed and settled. The main reason for disagreement was that one reviewer scored more conservatively than the other. For example, one reviewer more consistently scored items as either low, medium or high whereas the other reviewer scored items as low/medium or medium/high. The quality assessment score for each study is presented in the results table (Table 1). Overall, the methodological quality of the included studies varied between medium/low for both internal and external validity to high for internal validity and medium/high for external validity. One study was rated as having high external validity, nine were rated medium/high, 14 medium and 13 medium/low. Four studies were rated as having high internal validity, eight were rated medium/high, 16 medium and nine medium/low. Bias for external validity was mostly affected by 16 studies that failed to describe their source population clearly. This effected the quality assessment tool's scoring process, as because the source population was unclear, it was then unclear if the eligible population reflected the source population. An additional problem for the eligibility scoring was that only 12 studies reported clear inclusion or exclusion criteria. Altogether 29 studies failed to score medium/high or high for the eligible population item.

Bias for internal validity was mostly affected by the 26 studies failing to score at least medium/high for selection bias, mainly due to convenience sampling. There was also bias of measurement in nine studies, mainly due to shortening validated scales without a justification. Internal validity was also affected by 12 studies failing to report the number of participants that started and completed the study and five studies reporting a high non-completion rate. Six studies were underpowered. Another six studies compared groups of unequal sample sizes that would have tested the assumptions of the analysis.

3.7. Meta-analysis

A meta-analysis of 30 studies was conducted. Using multiple effect sizes from the same sample of participants as though they were independent biases the results of a meta-analysis as extra weight is given to one same sample of participants (Scammacca et al., 2014). As the studies Wartberg and Kammerl (2020), Wartberg et al. (2017), Wartberg et al. (2019) and Wartberg et al. (2021) included the same sample of participants measured at different time points, the effect size from Wartberg et al. (2019) was used as this was the first time-point in their longitudinal study (Cortese et al., 2016). The author of Kim and Ko (2020) and Kim (2019) was e-mailed and confirmed that the participants in the two studies overlapped. Therefore, only the larger sample size in the Kim (2019) study was included. As the studies Baysak et al. (2018) and Baysak et al. (2020) overlapped participants, the effect size reported at time-point one in the Baysak et al.'s (2020) longitudinal study was used. Wichstrøm et al.'s (2019) study was excluded in the meta-analysis as the study measured self-esteem and GD at two different time-points with a two-year gap which is inconsistent with the 30 studies included that measured self-esteem and GD at the same time-point. Similarly, Jeong et al. (2021) measured self-esteem and GD at two different timepoints with a six-month gap and was excluded. In Lemmens et al.'s (2015) study, two separate 9-item scales were used, one dichotomous and one polytomous GD scale. Because the 9-item dichotomous scale is the recommended version, only this score was included in the meta-analysis. Additionally, Lemmens et al. (2015) reported scores for both the 27-item and 9-item version of the scale in their study (from the same set of participants). We chose to include the 9-item version as this is the scale recommended by King, Chamberlain, et al. (2020). The 30 study meta-analysis revealed a negative mean association between low self-esteem and GD ($r = -0.269$ with a 95% confidence interval (CI) of -0.335 to -0.201, $Z = -7.439$, $p = <.001$; $\tau = 0.193$, $Tau^2 = 0.037$; $I^2 = 95.837\%$; $Q\ (29) = 696.53$, $p < .001$). The prediction interval suggests that the true effect size of 95% of all comparable populations falls in the interval of -0.59 to 0.12.

For transparency, Borenstein et al. (2009) recommends reporting whether grey literature effected the mean effect size. After removing the one grey literature study (Law, 2018) the meta-analysis showed a mean effect size of $r = -0.27$ (95% CI $= -0.337$ to -0.199, $Z = -7.277$, $p = <.001$; $\tau = 0.194$, $Tau^2 = 0.038$; $I^2 = 95.98\%$; $Q\ (28) = 696.508$, $p < .001$) with a prediction interval of -0.59 to 0.13.

The literature recommends including converted effect sizes rather than omitting studies that use an alternative metric to avoid the loss of information so as not to bias the sample of studies (Borenstein et al., 2009). However, when values are converted, assumptions are made about the nature of the underlying effects. As such, Borenstein et al. (2009) recommends reporting a sensitivity analysis with non-transformed effect sizes only. After removing the converted effect size studies, the remaining 19 studies showed a mean effect size of $r = -0.224$ (95% CI $= -0.272$ to -0.175, $Z = -8.707$, $p = <.001$; $\tau = 0.101$, $Tau^2 = 0.01$; $I^2 = 88.188\%$; $Q\ (18) = 152.385$, $p < .001$) with a prediction interval of -0.42 to -0.01.

Considering the original published protocol stated the review would include online gaming only, for transparency a meta-analysis was conducted for studies that investigated online gaming only. The included 16 studies showed a mean effect size of $r = -0.35$ (95% CI $= -0.468$ to -0.22, $Z = -5.053$, $p = <.001$; $\tau = 0.279$, $Tau^2 = 0.078$; $I^2 = 96.992\%$; $Q\ (15) = 498.637$, $p < .001$) with a prediction interval of -0.75 to 0.25. The original protocol also intended to include English-language studies only, for transparency a meta-analysis was conducted for studies published in English. The included 27 studies showed a mean effect size of $r = -0.23$ (95% CI $= -0.273$ to -0.186, $Z = -10.05$, $p = <.001$; $\tau = 0.105$, $Tau^2 = 0.011$; $I^2 = 87.335\%$; $Q\ (15) = 205.288$, $p < .001$) with a prediction interval of -0.43 to -0.01.

As Bargeron and Hormes (2017) was the only study included in the analysis that controlled for another variable (time spent gaming) a separate meta-analysis was conducted excluding Bargeron and Hormes (2017). The analysis showed a mean effect size of $r = -0.275$ (95% CI $= -0.342$ to -0.205, $Z = -7.456$, $p = <.001$; $\tau = 0.193$, $Tau^2 = 0.037$; $I^2 = 95.938\%$; $Q\ (15) = 689.295$, $p < .001$) with a prediction interval of -0.59 to 0.12.

Meta-analyses that focus on the mean effect can be misleading. A strength of a random-effects meta-analysis is its ability to analyse the heterogeneity of effects so that clinical interventions can be more efficiently targeted. The I^2 statistic indicates what proportion of the observed variance reflects true variance. Therefore, the higher the I^2 value (high in this review's analysis), the more accurately reasons for variation can be investigated and interpreted. The prediction interval and forest plot (Fig. 2) indicate there was heterogeneity across the studies. On visual inspection, two studies stand out as having a stronger negative effect size. Park et al. (2007) showed the largest negative effect size. The second largest negative effect, Leménager et al. (2013) was one of two studies included in the meta-analysis that used a clinical interview to measure gaming disorder. The study with the third largest effect (Jeong et al., 2018) also included a clinical interview. Therefore, to better understand the dispersion, a non *a priori* sensitivity analysis was conducted investigating the effects of firstly removing Park et al. (2007) from the meta-analysis and then secondly when removing Leménager et al. (2013) and Jeong et al. (2018) only. After removing Park et al. (2007) the meta-analysis showed a mean effect size of $r = -0.236$ (95% CI $= -0.276$ to -0.194, $Z = -10.827$, $p < .001$; $\tau = 0.104$, $Tau^2 = 0.011$; $I^2 = 86.941\%$; $Q\ (28) = 214.405$, $p < .001$) with a prediction interval of -0.43 to -0.02. After removing just Leménager et al. (2013) and Jeong et al. (2018) the meta-analysis showed a mean effect size of $r = -0.249$ (95% CI $= -0.317$ to -0.179, $Z = -6.733$, $p = <.001$; $\tau = 0.191$, $Tau^2 = 0.036$; $I^2 = 95.969\%$; $Q\ (27) = 669.834$, $p < .001$) with a prediction interval of -0.57 to 0.14.

As all studies were included in the analysis regardless of study quality, for transparency a cumulative analysis forest plot is reported below (Fig. 3), which illustrates how the mean effect size changes as studies of lower quality are cumulatively added to the meta-analysis. Considering the mean effect size moves both higher and lower as studies of lower quality are added, it is unlikely that the heterogeneity observed is due to study quality error variance.

3.8. Subgroup comparisons

As per *a priori* investigations, two subgroup analyses were conducted to better understand the dispersion of effect sizes comparing: studies from Eastern vs Western countries and collectivist vs individualistic countries. For the East vs West subgroup comparisons, Fernandes et al. (2020) and Fernandes et al. (2021) were excluded as their study's samples included participants from both Eastern and Western countries. Additionally, each of the Turkish studies Baysak et al. (2018), Ekşi et al. (2020) and Kircaburun et al. (2019) were excluded as Turkey is within both Europe and Asia. For the collectivist vs individualistic comparisons, Fernandes et al. (2020) study was excluded as it includes participants from both collectivist and individualistic countries. Countries were identified as collectivist or individualistic based on Hofstede's (1984) cultural dimensions theory and using the criteria from his company's website (Hofstede Insights, 2022). The results of the subgroup analysis are in Table 2.

An *a priori* subgroup analysis was planned based on game type. However, as only six studies (not including overlapping study samples) reported an association between GD and self-esteem for a specific game genre, it was deemed inappropriate to conduct the analysis based on Schwarzer et al.'s (2015) power recommendations that subgroup analysis is only appropriate if it contains over ten studies, particularly in a random-effects model.

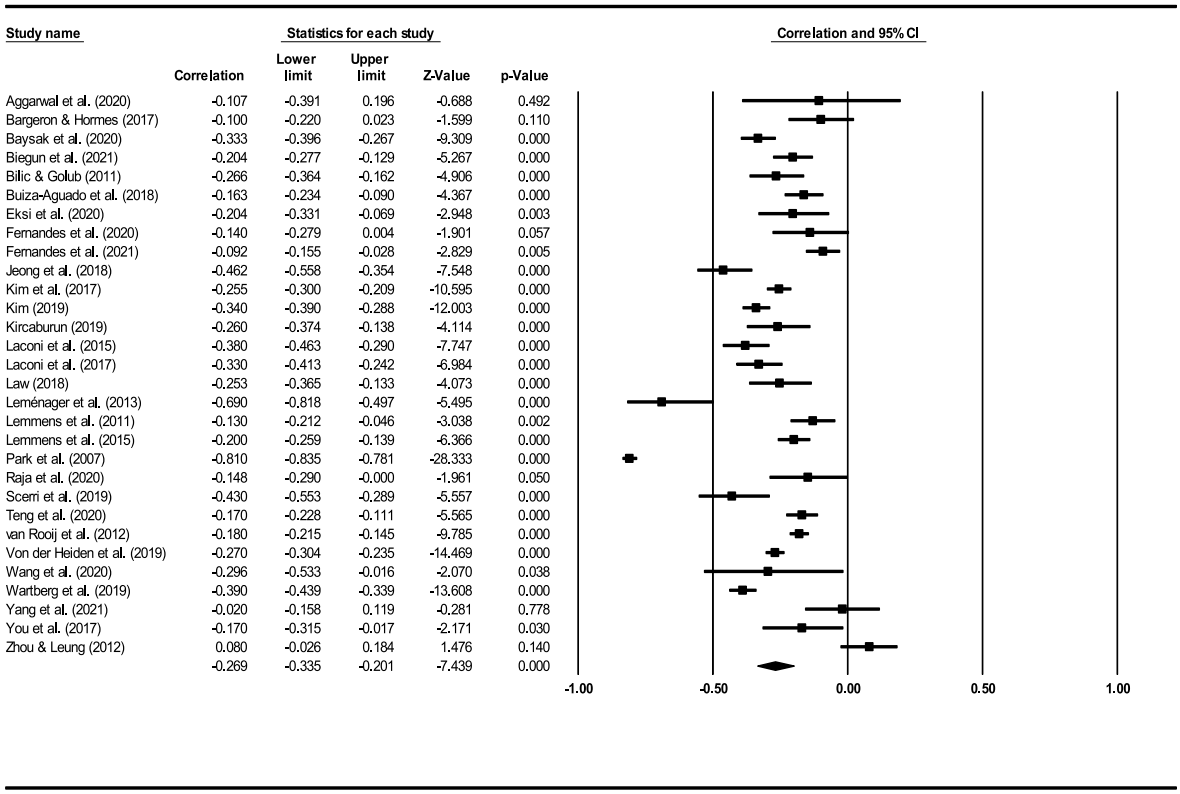

Fig. 2. Forest Plot for the 30 Study Meta-Analysis using a Random-Effects Model.

3.9. Exploratory subgroup comparisons

As there were no significant differences between any of the *a priori* subgroup comparisons (Table 2), non *a priori* subgroup analyses were conducted to better understand study effect size dispersion. Firstly, from observing the results in Table 1 above there appeared to be a pattern of larger effect sizes for studies using the RSES to measure self-esteem compared to studies using other self-esteem measures or shortened versions of the RSES. Therefore, a subgroup analysis comparing studies using the RSES vs non-RSES measure or shortened RSES measures was conducted.

Considering the described psychometric weaknesses of certain self-report GD tools, a non *a priori* subgroup analysis comparing the YIAT (which has weaker support for its psychometric properties when measuring GD; King, Russell, et al., 2020) vs the other eight included self-report GD tools which have better support for their psychometric properties, was conducted.

A non *a priori* subgroup analysis for age was also conducted (high school students vs adults) as previous research has indicated that GD may be more prevalent among adolescents (Festl et al., 2013). Teng et al.'s (2020) sample consisted of 17–21-year-old university students which overlaps with high-school student's age and therefore was excluded from the subgroup analysis. Additionally, Kim et al.'s (2017), Yang et al.'s (2021) and Jeong et al.'s (2018) samples included students younger than high school age and were therefore excluded from the analysis. Studies were also excluded from the subgroup analysis if their age range was unclear or if they included both school students and adults.

A non *a priori* subgroup analysis of gender was considered as research indicates that GD is more prevalent among males (Wittek et al., 2016). However again it was not possible to meet Schwarzer et al.'s (2015) power recommendations. Only Bilic et al. (2011), Kim et al. (2017) and Teng et al. (2020) provided separate self-esteem and GD correlations for males and females, all of which showed a larger effect size for males. Laconi et al. (2015) provided an effect size for females but not males in their study. Von der Heiden et al. (2019) when controlling for gender and age, found that the relationship between self-esteem and GD showed a small negative increase from −.27 to −0.28. Law (2018) also found a small negative increase when controlling for age and gender from −0.304 to −0.309. Lemmens et al. (2011) found that gender did not have a significant moderating effect on the relationship between psychosocial variables (of which self-esteem was included) and GD.

A non *a priori* subgroup analysis of offline vs online gaming was also considered as research indicates that online gaming may have a stronger association with GD (Bodi et al., 2021). However, no studies clearly indicated that they investigated offline gaming only.

Finally, a non *a priori* subgroup analysis of studies who recruited their participants from gaming forums, websites or gaming social media groups vs non-specific gaming group websites was considered as researchers have argued that participants recruited from gaming forums may draw participants with higher levels of problematic gaming use (Oggins & Sammis, 2012). However, again the review lacked the appropriate analytical power. The results of the non *a priori* subgroup analysis are in Table 2.

Table 2 shows that the only significant between subgroup difference was for the RSES vs adapted or non-RSES groups, indicating that the full RSES measure produces larger negative associations with GD than non-RSES and adapted RSES measures.

3.10. Publication bias

To ensure the integrity of the meta-analysis a publication bias assessment was conducted. A funnel plot (Fig. 4) considering standard error (Sterne & Egger, 2001) elucidated potential publication bias with 15 studies above the centred mean line, eight of which fell outside of the funnel plot and 10 studies below the mean line, seven of which fell outside of the funnel plot. Considering a greater proportion of studies below the mean fell outside of the funnel plot compared to the

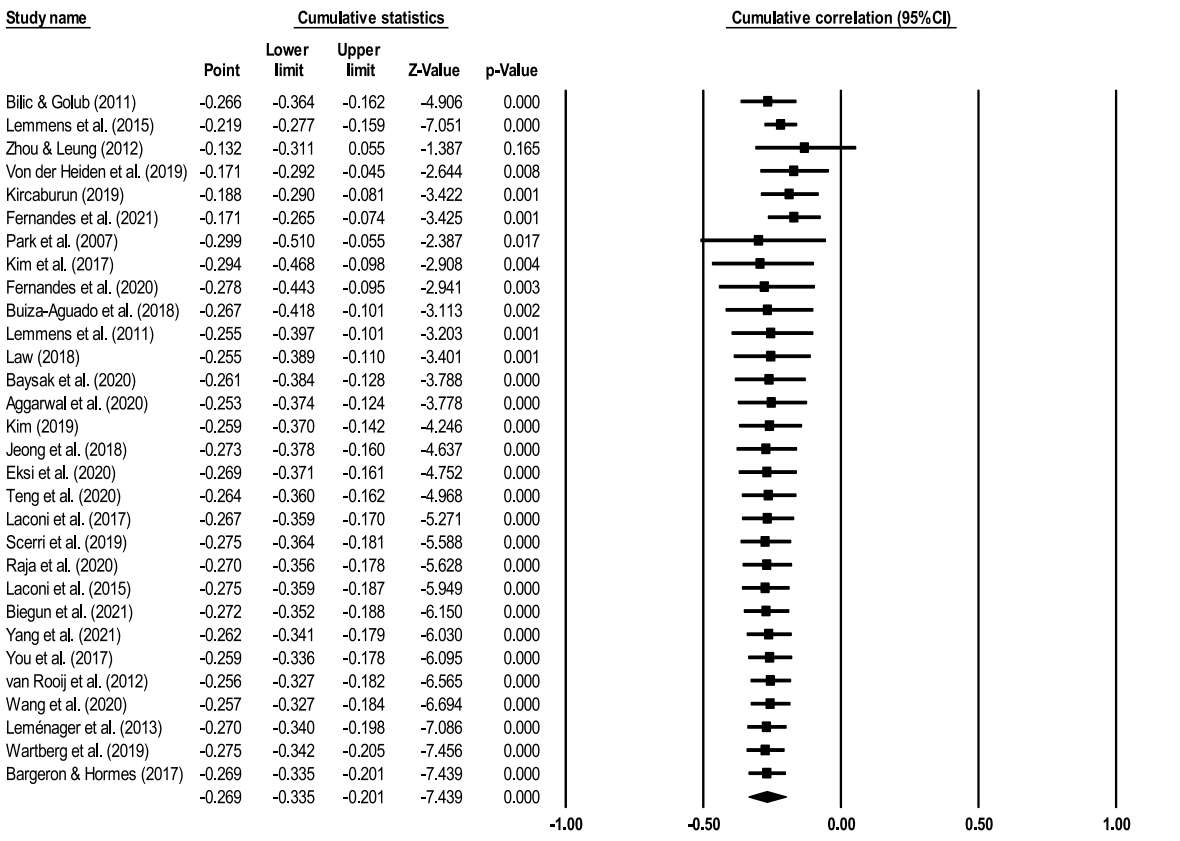

Fig. 3. Cumulative Analysis Forest Plot Illustrating how the Mean Effect Size Changes as Studies of Lower Quality are Added to the Meta-Analysis. *Note*: Moving down the forest plot, studies of lower quality are cumulatively added.

Table 2
Subgroup comparisons.

Subgroup analysis	Subgroup	Number of studies	*r* (*p*)	CI Lower limit/upper limit.	I^2	*P (Q)	Prediction interval	$p_{subgroup\ diff}$
East vs West								.934
	East	12	−.28 (.003)	−.44/−.09	97.99	.003	−.78/.44	
	West	13	−.27 (<.001)	−.33/−.21	87.93	<.001	−.46/−.06	
Collectivist vs Individualistic								.753
	Collectivist	18	−.26 (<.001)	−.37/−.14	97.14	<.001	−.69/.3	
	Individualistic	11	−.28 (<.001)	−.34/−.22	88.37	<.001	−.48/−.05	
YIAT vs non-YIAT								.442
	YIAT	7	−.32 (.013)	−.53/−.07	98.74	.013	−.86/.56	
	Non-YIAT	21	−.22 (<.001)	−.27/−.18	84.6	<.001	−.41/−.02	
Rosenberg vs Adapted/non-Rosenberg								.016
	Rosenberg	18	−.33 (<.001)	−.42/−.23	97.09	<.001	−.68/.14	
	Adapted/non-Rosenberg	12	−.18 (<.001)	−.24/−.12	84.11	<.001	−.39/.05	
Age								.31
	School student	9	−.315 (.002)	−.49/-.12	98.46	.002	−.8/.43	
	Adult	7	−.2 (.003)	−.33/−.07	88.79	.003	−.57/.24	

Note: *P values for Cochran's *Q* statistic (Cochran, 1937).
YIAT: The Young Internet Addiction Test (Young, 1998).

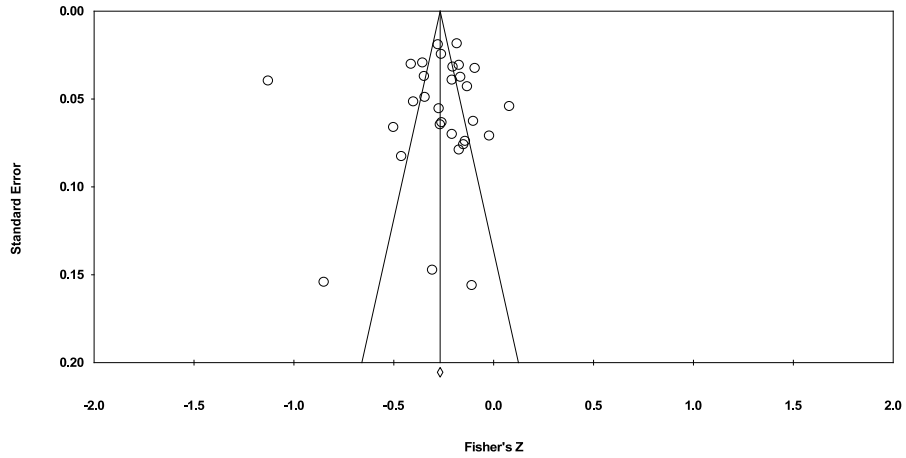

Fig. 4. Funnel plot of standard error by Fisher's Z, assessing for publication bias.

proportion of studies above the mean, there was potential bias of several studies with larger negative effect sizes missing. Duval and Tweedie's (2000) Trim and Fill analysis was used to assess for missing studies below the mean. This resulted in nine additional studies included with a point estimate of $r = -0.34$, (95% CI $= -0.401$ to -0.275. Q (38) $=$ 1044.92, $p < .001$). When the same analysis was used for studies above the mean, no additional studies were identified. The funnel plot with observed and Trim and Fill imputed studies is displayed in Fig. 5. The Trim and Fill analysis should be interpreted with caution as more commonly with publication bias, it is studies with smaller effect sizes that go unpublished rather than studies with larger effect sizes. As such, further publication bias analysis was carried out. The Egger, Smith, Schneider, and Minder (1997) tests indicated no publication bias (intercept $= -0.563$, $p = .387$). However, it is important to note that the power of Egger's test is lowered when there is heterogeneity (Simmonds, 2015). Rosenthal's (1979) Fail-safe N suggested that 8793 studies would need including before the cumulative effect became statistically non-significant. Orwin's (1983) Fail-safe N analysis allows a researcher to select a value to represent the smallest effect deemed to be of substantive importance in investigating how many missing studies it would take to bring the summary effect below this point. The analysis suggested that with a score of -0.1, $N = 51$ studies would be needed, and with a score of -0.05 $N = 131$ studies would be needed. It is unlikely the review missed 51 studies.

4. Discussion

This study used systematic review methodology and meta-analysis to investigate the association between self-esteem and GD. In total, 37 studies met the inclusion criteria with 31 reporting a significant negative correlation between low self-esteem and GD. The dispersion of effect sizes in the meta-analysis indicates that the overall mean effect should be interpreted with caution. In the subgroup analysis, there were no significant differences found between Eastern vs Western countries or collectivist vs individualistic countries. In the three non *a priori* subgroup analyses, a significant difference was found only for the self-esteem measure comparison, indicating the full-scale RSES shows a stronger relationship with GD than other self-esteem measures. Despite the heterogeneity in the studies reported, generally the findings support literature which indicates low self-esteem is a risk factor for GD (Lemenager et al., 2020). Additionally, this finding appears to apply irrespective of culture or age. Considering this, self-esteem may be a risk factor worth targeting in GD interventions. Both cognitive-behavioural therapy (CBT) and compassion focused therapy (CFT) have been found to be effective in treating self-esteem (Kolubinski et al., 2018; Thomason & Moghaddam, 2021). Potentially then, these models could be effective in treating self-esteem as part of a GD intervention.

The longitudinal studies that attempted to infer direction of causality produced contrasting results. Therefore, it is not possible to infer support for or against compensatory hypotheses of self-esteem's relationship

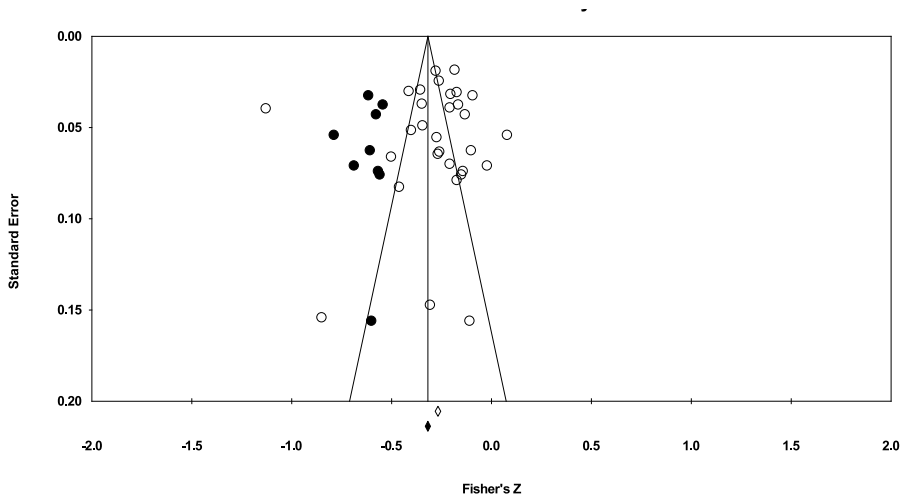

Fig. 5. Trim and fill analysis of standard error by Fisher's Z.

with GD. One explanation for the contrasting longitudinal results could be because the participants were not drawn from clinical samples and therefore changes over time in self-esteem or GD scores may have been more heavily influenced by other confounding variables. Other than the small number of longitudinal studies attempting to infer causality, no studies experimentally investigated the causal direction of the relationship between GD and self-esteem. Experimental design studies would offer more credibility in supporting or refuting the compensatory hypotheses of self-esteem and GD and therefore are an important area for future research. It would also give clinicians more confidence in understanding the maintenance factors in GD and improve GD interventions. A potential study design for initial causality exploration could be brief self-esteem experimental manipulations (Williams & Jarvis, 2006). For example, a researcher could investigate whether people scoring higher for GD are more motivated to play their game after their self-esteem is experimentally lowered. If their motivation to play did increase, it would indicate that lowered self-esteem motivates people scoring higher for GD to game. The researcher could then investigate if the self-esteem of those same gamers increases when they are given the opportunity to game post-experimentally lowered self-esteem vs those who are not given the opportunity to game post-experimentally lowered self-esteem. If those who have the opportunity to game experience an increase in self-esteem vs those who do not have an opportunity to game, this would indicate that compensating for lower self-esteem is a motivation to game; inferring support for the compensatory hypothesis.

Other than the comparison of self-esteem measures, there were no significant differences in the subgroup analyses. This indicates that low self-esteem is a risk factor for GD regardless of culture. However, interestingly there were observable differences in the prediction intervals between Western vs Eastern countries. The much wider prediction interval for Eastern counties could indicate that there were large within group differences which could explain a proportion of this study's overall heterogeneity. Due to power, it was not possible to compare Eastern countries against each other. However, the results table indicates that studies conducted in China showed much lower effect sizes. Two of the four Chinese studies showed a non-significant effect and one a very low effect size. It is not clear why Chinese studies found lower effect sizes. One potential explanation might be the government restrictions on gaming over the last 22 years (Xiao, 2021). For example, China banned the production, import and sale of videogame consoles from 2000 to 2015 (Xiao, 2021). With restrictions, less problematic gaming mediums and genres grew in popularity, with the mobile device gaming market growing 400% between 2013 and 2018 whilst riskier gaming mediums and genres stagnated (King et al., 2019; Lozic et al., 2018, pp. 208–216). Recently the government restricted the amount of time minors can play video games to weekends only between 20:00 and 21:00 (Xiao, 2021). Compared to other countries included in the review it might be that through restricted access to gaming, Chinese people have had less of an opportunity to compensate for low self-esteem by gaming. However, it may also indicate that Chinese people have had their self-esteem protected by reduced exposure to gaming. This interpretation is also supported by the finding that the only longitudinal Chinese study included in the review (Teng et al., 2020) found that GD at time-point one significantly predicted low self-esteem at time-point two. This interpretation would challenge the compensatory hypothesis and instead indicate that gaming causes low self-esteem. This tentative hypothesis again highlights the need for experimental studies to investigate the direction of causality between the variables self-esteem and GD. If it is the case that restrictions have mitigated against problematic gaming and low self-esteem the clinical implication is that restricting access to gaming could be one of the more effective treatment mechanisms for GD. Of course, such a draconian sounding intervention would need to be collaboratively negotiated with clients. The intervention could be a fruitful area for future research. It may also point to responsibilities needing to be taken by gaming companies and regulators.

One explanation for a non-significant difference in effect size between countries and cultures could be due to how self-esteem was measured. Schmitt and Allik (2005) in one of the largest studies measuring self-esteem across 53 countries, contrary to their expectations did not find a difference in global self-esteem between individualistic or collectivist countries (it is important to note that other studies have found a difference: Brown et al., 2009; Cai et al., 2007). However, Schmitt and Allik (2005) did report differences in sub-components of self-esteem between the countries and cultures. They found collectivist countries reported lower perception of competences (feeling one is capable and efficacious) but higher self-liking (feeling one is good, socially relevant, and contributes to group harmony) than participants from individualist countries. Therefore, there is a possibility that there are cross-cultural differences in the relationship between GD and self-esteem but that these differences exist within unassessed sub-components of self-esteem. This might also explain why the only significant subgroup difference in the analysis was between the self-esteem scales. Unfortunately, as most studies did not report which items were removed from the RSES before administration it was not possible to make inferences about which subcomponents of self-esteem better explain its relationship with GD. Therefore, future research using the RSES should investigate which sub-components of self-esteem are a greater risk factor for GD and if this differs across cultures. These subcomponents could then be more efficiently targeted in GD interventions.

The review included five measures of GD which according to King, Chamberlain, et al. (2020) have the best support for their psychometric properties plus five additional measures. However, two of the additional measures included for the purpose of cultural representation (YIAT and YDQ, although none of the eligible studies used the YDQ) have weaker support for their psychometric properties when measuring GD. In the subgroup comparisons for measurement of GD, the prediction interval for the YIAT was much larger than that of the better supported GD measures indicating credibility for their precision in measuring GD, thus supporting King et al.'s (2020) GD tool measurement recommendations. It is important to note the better supported GD measures prediction interval was also narrower than the full 30 study meta-analysis prediction interval indicating that the prediction interval was not narrower due to the number of studies included in the non-YIAT vs YIAT comparison, adding further credibility to the psychometric properties of these tools. The wide prediction interval for the YIAT could also suggest this less reliable and valid measure of GD added error heterogeneity to the analysis. Furthermore, as Park et al. (2007) used the YIAT it could explain their large effect size. Park et al.'s (2007) large effect size may also be explained by publication date, as the study was published before the replication crisis movement in psychology when stronger effect sizes were more likely to be published (Pashler & Wagenmakers, 2012). In comparison, the 2nd earliest publication date in this review is 2011. The large effect size is also likely explained by the study comparing those scoring 20–40 on the YIAT vs those above 50 and therefore missing a middle 10 points of scores inflating the differences between the two groups in a *t*-test. Another reason for the large effect size as well as heterogeneity in the meta-analysis, could be due to converted effect sizes being added to the analysis. Considering a larger number of studies should increase precision in the prediction interval, it is notable that removing the 12 converted effect size studies actually narrowed the prediction interval. Therefore, the transformed effect size studies potentially added more error heterogeneity to the meta-analysis than true heterogeneity. Transformed effect sizes might also explain why the two clinical interview studies showed larger effect sizes. However, their effect sizes could also indicate that low self-esteem is a bigger risk factor for problem gaming than previous self-report studies have found. It could be that gamers and clinicians perceive problem gaming differently, with problematic gamers underreporting problems or clinicians overreporting problems.

Due to power, it was not possible to explore whether the relationship between self-esteem and GD varied based on game genre. However, the

variation in effect size scores within each of the specific game genre groups could indicate that game genre is not a strong explanatory variable for heterogeneity. Although, this within group variance could be explained by other factors. It might be that Leménager et al. (2013) found a larger effect size for MMORPG participants than You et al. (2017) because they measured GD via clinical interview. Aggarwal et al. (2020) potentially found a smaller effect size for an MOBA sample than Wang et al. (2020) because they used the single Robbins self-esteem scale. Compared to Baysak et al. (2018), Zhou and Leung (2012) may have found a non-significant effect size for strategy games because they used a Chinese sample. Therefore, it is important for future reviews to re-visit this game genre question.

Due to power, it was not possible to complete a gender subgroup analysis. However, all three studies that included gender correlations found larger effect sizes for males compared to females. Although, when two different studies controlled for gender, the relationship between self-esteem and GD showed only a small negative increase. Therefore, it is difficult to conclude if gender explains true heterogeneity. However, if it does, it is more likely that males have larger effect sizes. An explanation for this potential gender difference may be due to how males and females value different criteria for self-evaluation. For example, Josephs et al. (1992) found that male self-esteem is associated with personal achievements whereas female's evaluation stems from their attachments to important others. Potentially then, males with low self-esteem might be more drawn to boosting their self-esteem in the gaming world through levelling up their status and experiencing personal achievement, which, if satisfied and compensated for, may maintain their drive to game for further self-esteem boosts. Of course, that does not mean attachment is not an important explanatory variable too. Future research could investigate whether the sub-components of self-esteem differ in their association with GD by gender. Such findings again could help inform more efficiently targeted GD interventions.

As the six studies that carried out a moderation or mediation analysis tested uniquely different models, it is not possible to draw any aggregated conclusions about the variables that moderate or mediate the relationship between self-esteem and GD.

Considering the limitations of the included studies, overall, the methodological quality of the included papers was moderate. In the cumulative meta-analysis for study quality, the overall effect size did not shift predominately lower or higher but instead moved slightly in both directions as lower quality studies were added. This potentially indicates that the heterogeneity observed in the review was not due to study quality error variance. Considering the results of the quality assessment, to reduce bias in future reviews, research should carefully consider how participants are recruited and whether the included participants are representative of the population the study aims to extrapolate their results to. Clear inclusion and exclusion criteria need to be stated in this process. If shortening a self-esteem measure, this needs to be justified with a clear description of the items excluded in the measure.

4.1. Limitations

There are a number of limitations in this review. Firstly, there were two changes made to the original published PROSPERO protocol. However, both changes were justified and likely added strength to the review. The review also transparently reported the meta-analysis results that would have been found if the changes were not made. Secondly, although recent evidence was used to justify using Google Translate, the accuracy of the translation would have been more robust if the machine translations were cross checked by a qualified translator. Thirdly, three non *a priori* sub-group analyses were conducted. However, these were justified in line with Borenstein's (2009) suggestions that the purpose of a random effects analysis is to understand the dispersion of study effect sizes and the *a priori* sub-group analysis did not offer a statistical explanation. Fourthly, sub-group analyses are purely observational (Borenstein & Higgins, 2013).

Fifthly, the review adhered to Schwarzer et al.'s (2015) subgroup power recommendations. However, detecting effect size differences between subgroups in a meta-analysis is difficult because differences are usually smaller and therefore more statistical power is needed (Hedges & Pigott, 2004). Consequently, to conclude that certain subgroups in this review have equivalent effects could be a misinterpretation (Hedges & Pigott, 2004; Borenstein, 2009). For example, there was a notable difference in effect size for the age sub-group analysis, however this was a non-significant effect. There is a possibility that in the future with more studies conducted, reviews will detect a significant difference with greater power. Sixthly, considering the concerns regarding statistical power in the sub-group analysis, it could be argued that it would have been more appropriate to include all studies assessing GD in the review regardless of the GD measure's psychometric properties. However, this would have increased bias, making it more difficult to interpret what might be true or error heterogeneity and sacrificed the validity and reliability of the review's results. Seventhly, as the six studies that included a moderation or mediation analysis tested uniquely different models, it was not possible to aggregate studies and investigate appropriately the review's secondary research question 'what variables moderate and mediate the relationship between self-esteem and GD?

Eighthly, a significant limitation of the review was that there were no clinical samples included. Therefore, making generalisations to the crucial population that most GD research aims to serve is limited. It is important to note that the three studies with the largest effect sizes in the meta-analysis compared people who met a cut-off score for GD vs those that did not. Therefore, it may be that the effect size in clinical samples is much larger. The lack of clinical samples also raises questions around the methodology of published clinical trials for GD. Why is self-esteem (a potentially crucial risk factor for GD) not being targeted or measured for change in GD clinical trials?

Finally, there was a questionable quality assessment agreement rate between the researchers. However, this might be explained by the number of scale point response options. The researchers scoring bias would likely have had a higher agreement level if there were just two options "high" or "low" bias, rather than the actual 5-point response. Consider that most disagreements were due to one researcher scoring an item one scale point away from the other e.g., "medium" vs "medium/high".

Despite the limitations, there were strengths to this review. Notably, the review adhered to the updated PRISMA principles (Page et al., 2021), it was pre-registered, non-English-language studies were included, a wide variety of data-bases and grey literature was searched as well as backwards citation searching of the reference lists of all included studies, which allows confidence in the conclusion that all relevant research was included in the review and that conclusions arising from the review can be based on the synthesis of all available evidence.

4.2. Conclusions

Despite the heterogeneity in the studies reported, generally the findings indicate low self-esteem is a risk factor for GD. Heterogeneity can be partly explained by the various self-esteem measures used. Tentative explanations for other sources of the heterogeneity are suggested, including differences within Eastern countries and gender but more power is needed to investigate these interpretations. Additionally, as more research is generated, it would be important to complete sub-group analyses comparing game type and offline vs online gaming. Tentative clinical recommendations are made such as restricting access to gaming. Recommendations for future research include: using an experimental design to understand the direction of causality between self-esteem and GD; to use the RSES to investigate which sub-components of self-esteem are a greater risk factor for GD and how these sub-component associations differ by gender; measuring and targeting self-esteem in clinical samples/interventions. Such research could

improve the efficiency of GD interventions.

Declaration of competing interest

This study was conducted as part of the first author's doctorate in clinical psychology and supported by a research training grant from the University of Southampton. The funding source had no involvement in the project nor the decision to submit the study for publication. There are no other interests to declare.

Data availability

The data that has been used is confidential.

Acknowledgements

This study was conducted as part of the first author's doctorate in clinical psychology and supported by a research training grant from the University of Southampton. The funding source had no involvement in the project nor the decision to submit the study for publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chb.2023.107776.

References

Aggarwal, S., Saluja, S., Gambhir, V., Gupta, S., & Satia, S. P. S. (2020). Predicting likelihood of psychological disorders in PlayerUnknown's Battlegrounds (PUBG) players from Asian countries using supervised machine learning. *Addictive Behaviors, 101*, 1–8. https://doi.org/10.1016/j.addbeh.2019.106132. Article 106132.

Allen, J. J., & Anderson, C. A. (2018). Satisfaction and frustration of basic psychological needs in the real world and in video games predict internet gaming disorder scores and well-being. *Computers in Human Behavior, 84*, 220–229. https://doi.org/10.1016/j.chb.2018.02.034

American Psychiatric Association. (2013). *Anxiety disorders. In Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596.dsm05

Anderson, E. L., Steen, E., & Stavropoulos, V. (2017). Internet use and problematic internet use: A systematic review of longitudinal research trends in adolescence and emergent adulthood. *International Journal of Adolescence and Youth, 22*(4), 430–454. https://doi.org/10.1080/02673843.2016.1227716

Balk, E. M., Chung, M., Chen, M. L., & Trikalinos, T. A. (2013). *Assessing the accuracy of Google translate to allow data extraction from trials published in non-English languages* (Vol. 12). Report prepared for *Agency for Healthcare Research and Quality* U.S. Department of Health and Human Services [Publication No (13)-EHC145-EF].

Bargeron, A. H., & Hormes, J. M. (2017). Psychosocial correlates of internet gaming disorder: Psychopathology, life satisfaction, and impulsivity. *Computers in Human Behavior, 68*, 388–394. https://doi.org/10.1016/j.chb.2016.11.029

Baysak, E., Kaya Yertutanol, F. D., Şahiner, İ. V., & Candansayar, S. (2020). Sociodemographic characteristics of online game players and the relationship of game addiction with some psychosocial factors. *Journal of Clinical Psychiatry, 23*(2), 196–203. https://doi.org/10.5505/kpd.2020.10337

Baysak, E., Yertutanol, F., Dalgar, I., & Candansayar, S. (2018). How game addiction rates and related psychosocial risk factors change within 2-years: A follow-up study. *Psychiatry investigation, 15*(10), 984–990. https://doi.org/10.30773/pi.2018.08.16

Beard, C. L., & Wickham, R. E. (2016). Gaming-contingent self-worth, gaming motivation, and internet gaming disorder. *Computers in Human Behavior, 61*, 507–515. https://doi.org/10.1016/j.chb.2016.03.046

Bender, P. K., & Gentile, D. A. (2020). Internet gaming disorder: Relations between needs satisfaction in-game and in life in general. *Psychology of Popular Media, 9*, 266–278. https://doi.org/10.1037/ppm0000227

Biegun, J., Edgerton, J. D., & Roberts, L. W. (2021). Measuring problem online video gaming and its association with problem gambling and suspected motivational, mental health, and behavioral risk factors in a sample of university students. *Games and Culture, 16*(4), 434–456. https://doi.org/10.1177/1555412019897524

Bilic, E., & Golub, T. L. (2011). Pathological videogame use: The role of gender, selfesteem and educational context/patolosko igranje videoigara: Uloga spola, samopostovanja i edukacijske sredine. *Hrvatska Revija za Rehabilitacijska Istrazivanja, 47*(2), 1–14.

Bodi, G., Maintenant, C., & Pennequin, V. (2021). The role of maladaptive cognitions in gaming disorder: Differences between online and offline gaming types. *Addictive Behaviors, 112*, 106595 1–5. https://doi.org/10.1016/j.addbeh.2020.106595

Boland, A., Cherry, G., & Dickson, R. (2017). *Doing a systematic review: A student's guide* (2nd ed.). Sage.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons Ltd.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). *Comprehensive MetaAnalysis (version 3). Biostat.* www.meta-analysis.com/pages/new_v3.php.

Borenstein, M., & Higgins, J. (2013). Meta-analysis and subgroups. *Prevention Science, 14*(2), 134–143. https://doi.org/10.1007/s11121-013-0377-7

Brown, J. D., Cai, H., Oakes, M. A., & Deng, C. (2009). Cultural similarities in self-esteem functioning: East is East and West is West, but sometimes the twain do meet. *Journal of Cross-Cultural Psychology, 40*(1), 140–157. https://doi.org/10.1177/0022022108326280

Buchholz, K. (2021). *Asian countries make up 40% of the world's top 10 video gaming markets*. World Economic Forum. https://www.weforum.org/agenda/2021/08/which-countries-have-the-largest-video-gaming-markets/.

Buiza-Aguado, C., Alonso-Canovas, A., Conde-Mateos, C., Buiza-Navarrete, J. J., & Gentile, D. (2018). Problematic video gaming in a young Spanish population: Association with psychosocial health. *Cyberpsychology, Behavior, and Social Networking, 21*(6), 388–394. https://doi.org/10.1089/cyber.2017.0599

Cai, H., Brown, J. D., Deng, C., & Oakes, M. A. (2007). Self-esteem and culture: Differences in cognitive self-evaluations or affective self-regard? *Asian Journal of Social Psychology, 10*(3), 162–170. https://doi.org/10.1111/j.1467-839X.2007.00222.x

Chiu, C.-Y., & Hong, Y.-Y. (2006). *Social psychology of culture*. Psychology Press.

Cho, S. H., & Kwon, J. H. (2016). OR-10: Self-esteem fluctuation and internet game use. *Journal of Behavioral Addictions, 5*(S1), 7–8.

Cho, C. B., & Lim, J. S. (2010). Parent-child interaction and internet games addiction and delinquency in adolescence. *Korean Journal of Youth Studies, 17*, 197–226.

Clephas, P. R. D., Hoeks, S. E., Trivella, M., Guay, C. S., Singh, P. M., Klimek, M., & Heesen, M. (2021). Prognostic factors for chronic post-surgical pain after lung or pleural surgery: A protocol for a systematic review and meta-analysis. *British Medical Journal Open, 11*(6), 1–5. https://doi.org/10.1136/bmjopen-2021-051554

Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society, 4*(1), 102–118. https://doi.org/10.2307/2984123

Cortese, S., Moreira-Maia, C. R., St Fleur, D., Morcillo-Peñalver, C., Rohde, L. A., & Faraone, S. V. (2016). Association between ADHD and obesity: A systematic review and meta-analysis. *American Journal of Psychiatry, 173*(1), 34–43. https://doi.org/10.1176/appi.ajp.2015.15020266

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y

Dixon, J., Karagiannidou, M., & Knapp, M. (2018). The effectiveness of advance care planning in improving end-of-life outcomes for people with dementia and their carers: A systematic review and critical discussion. *Journal of Pain and Symptom Management, 55*(1), 132–150. https://doi.org/10.1016/j.jpainsymman.2017.04.009

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ekşi, F., Demirci, I., & Tanyeri, H. (2020). Problematic technology use and well-being in adolescence: The personal and relational effects of technology. *Addicta: The Turkish Journal on Addictions, 7*(2), 107–121. https://doi.org/10.5152/ADDICTA.2020.19077

Errasti, J. M., Vázquez, I., Fernández, J. M., & Fernández, J. (2018). Differences between individualist and collectivist cultures in emotional Facebook usage: Relationship with empathy, self-esteem, and narcissism. *Psicothema, 30*(4), 376–381. https://doi.org/10.7334/psicothema2018.101

Fernandes, B., Biswas, U. N., Mansukhani, R. T., Casarín, A. V., & Essau, C. A. (2020). The impact of COVID-19 lockdown on internet use and escapism in adolescents. *Revista de psicología clínica con niños y adolescentes, 7*(3), 59–65. https://doi.org/10.21134/rpcna.2020.mon.2056

Fernandes, B., Uzun, B., Aydin, C., Tan-Mansukhani, R., Vallejo, A., Saldaña-Gutierrez, A., … Essau, C. A. (2021). Internet use during COVID-19 lockdown among young people in low-and middle-income countries: Role of psychological wellbeing. *Addictive Behaviors Reports, 14*, 1–6. https://doi.org/10.1016/j.abrep.2021.100379. Article 100379.

Festl, R., Scharkow, M., & Quandt, T. (2013). Problematic computer game use among adolescents, younger and older adults. *Addiction, 108*(3), 592–599. https://doi.org/10.1111/add.12016

Figueroa, C., Johnson, C., Ford, N., Sands, A., Dalal, S., Meurant, R., … Baggaley, R. (2018). Reliability of HIV rapid diagnostic tests for self-testing compared with testing by health-care workers: A systematic review and meta-analysis. *The lancet HIV, 5*(6), 277–290. https://doi.org/10.1016/S2352-3018(18)30044-4

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

Green, R., Delfabbro, P. H., & King, D. L. (2020). Avatar-and self-related processes and problematic gaming: A systematic review. *Addictive Behaviors, 108*, 1–11. https://doi.org/10.1016/j.addbeh.2020.106461. Article 106461.

Griffiths, M., King, D., & Demetrovics, Z. (2014). DSM-5 internet gaming disorder needs a unified approach to assessment. *Neuropsychiatry, 4*(1), 1–4. https://doi.org/10.2217/npy.13.82

Harris, M. A., Donnellan, M. B., & Trzesniewski, K. H. (2018). The Lifespan Self-Esteem Scale: Initial validation of a new measure of global self-esteem. *Journal of Personality Assessment, 100*(1), 84–95. https://doi.org/10.1080/00223891.2016.1278380

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*(6), 895–910. https://doi.org/10.1037/0022-3514.60.6.895

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426–445. https://doi.org/10.1037/1082-989X.9.4.426

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). John Wiley & Sons.

Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). Sage.

Insights, Hofstede (2022). *Country comparison*. www.hofstede-insights.com/country-comparison//.

Jackson, J. L., Kuriyama, A., Anton, A., Choi, A., Fournier, J. P., Geier, A. K., … Sun, R. (2019). The accuracy of Google Translate for abstracting data from non–English-language trials for systematic reviews. *Annals of Internal Medicine, 171*(9), 677–679. https://doi.org/10.7326/M19-0891

Jeong, E. J., Lee, D., Lee, S. J., & Kong, J. (2021). Pathological gaming in South Korean adolescents from the perspectives of self-esteem and self-control. *Social Behavior and Personality: An International Journal, 49*(3), 1–11. https://doi.org/10.2224/sbp.9657

Jeong, H., Yim, H. W., Lee, S. Y., Lee, H. K., Potenza, M. N., Kwon, J. H., … Choi, J. S. (2018). Discordance between self-report and clinical diagnosis of Internet gaming disorder in adolescents. *Scientific Reports, 8*(1), 1–8. https://doi.org/10.1038/s41598-018-28478-8

Josephs, R. A., Markus, H. R., & Tafarodi, R. W. (1992). Gender and self-esteem. *Journal of Personality and Social Psychology, 63*(3), 391. https://doi.org/10.1037/0022-3514.63.3.391

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The core self-evaluations scale: Development of a measure. *Personnel Psychology, 56*(2), 303–331. https://doi.org/10.1111/j.1744-6570.2003.tb00152.x

Kadam, K. S., & Himanshi. (2020). A cross sectional study of prevalence of gaming disorder and correlation of self-esteem, parenting style, perceived stress and sleep quality with gaming in college students. *Indian Journal of Psychiatry, 62*, S25.

Kim, B. N. (2019). Korean validation of the internet gaming disorder-20 test. *Cyberpsychology, Behavior, and Social Networking, 22*(4), 271–276. https://doi.org/10.1089/cyber.2018.0096

Kim, B. N., & Ko, H. (2020). Psychometric properties of the nine-item Korean internet gaming disorder scale: Short form. *Cyberpsychology, Behavior, and Social Networking, 23*(12), 854–859. https://doi.org/10.1089/cyber.2020.0227

Kim, J. Y., Lee, J. S., & Oh, S. (2017). A path model of school violence perpetration: Introducing online game addiction as a new risk factor. *Journal of Interpersonal Violence, 32*(21), 3205–3225. https://doi.org/10.1177/0886260515597435

King, D. L., Chamberlain, S. R., Carragher, N., Billieux, J., Stein, D., Mueller, K., … Delfabbro, P. H. (2020). Screening and assessment tools for gaming disorder: A comprehensive systematic review. *Clinical Psychology Review, 77*, 1–20. https://doi.org/10.1016/j.cpr.2020.101831. Article 101831.

King, D. L., & Delfabbro, P. H. (2016). The cognitive psychopathology of Internet gaming disorder in adolescence. *Journal of Abnormal Child Psychology, 44*(8), 1635–1645. https://doi.org/10.1007/s10802-016-0135-y

King, D. L., Delfabbro, P. H., & Griffiths, M. D. (2012). Clinical interventions for technology-based problems: Excessive Internet and video game use. *Journal of Cognitive Psychotherapy, 26*(1), 43. https://doi.org/10.1891/0889-8391.26.1.43

King, D. L., Delfabbro, P. H., Perales, J. C., Deleuze, J., Király, O., Krossbakken, E., & Billieux, J. (2019). Maladaptive player-game relationships in problematic gaming and gaming disorder: A systematic review. *Clinical Psychology Review, 73*, 1–13. https://doi.org/10.1016/j.cpr.2019.101777. Article 101777.

King, D. L., Russell, A. M., Delfabbro, P. H., & Polisena, D. (2020). Fortnite microtransaction spending was associated with peers' purchasing behaviors but not gaming disorder symptoms. *Addictive Behaviors, 104*, 1–7. https://doi.org/10.1016/j.addbeh.2020.106311. Article 106311.

Király, O., Griffiths, M. D., King, D. L., Lee, H. K., Lee, S. Y., Bányai, F., … Demetrovics, Z. (2018). Policy responses to problematic video game use: A systematic review of current measures and future possibilities. *Journal of Behavioral Addictions, 7*(3), 503–517. https://doi.org/10.1556/2006.6.2017.050

Király, O., Tóth, D., Urbán, R., Demetrovics, Z., & Maraz, A. (2017). Intense video gaming is not essentially problematic. *Psychology of Addictive Behaviors, 31*(7), 807–817. https://doi.org/10.1037/adb0000316

Kircaburun, K., Griffiths, M. D., & Billieux, J. (2019). Psychosocial factors mediating the relationship between childhood emotional trauma and internet gaming disorder: A pilot study. *European Journal of Psychotraumatology, 10*(1), 1–10. https://doi.org/10.1080/20008198.2018.1565031

Ko, T. G. (2008). Research on the adolescents' undesirable leisure activities playing with the violent games in the internet. *Journal of Leisure and Recreation Studies, 32*(3), 109–119.

Kolubinski, D. C., Frings, D., Nikčević, A. V., Lawrence, J. A., & Spada, M. M. (2018). A systematic review and meta-analysis of CBT interventions based on the Fennell model of low self-esteem. *Psychiatry Research, 267*, 296–305. https://doi.org/10.1016/j.psychres.2018.06.025

Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. M. K., Hammerstrøm, K., & Sathe, N. (2016). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Methods Guides, 3*(1), 1–73. https://doi.org/10.4073/cmg.2016.1

Laconi, S., Pirès, S., & Chabrol, H. (2017). Internet gaming disorder, motives, game genres and psychopathology. *Computers in Human Behavior, 75*, 652–659. https://doi.org/10.1016/j.chb.2017.06.0120747-5632

Laconi, S., Tricard, N., & Chabrol, H. (2015). Differences between specific and generalized problematic Internet uses according to gender, age, time spent online and psychopathological symptoms. *Computers in Human Behavior, 48*, 236–244. https://doi.org/10.1016/j.chb.2015.02.006

Law, C. (2018). *Pathological technology use among young adolescents [Doctoral thesis*. Alliant International University]. ProQuest One Academic. www.proquest.com/dissertations-theses/pathological-technology-use-among-young/docview/2103854461/se-2?accountid=13963.

Leménager, T., Gwodz, A., Richter, A., Reinhard, I., Kämmerer, N., Sell, M., & Mann, K. (2013). Self-concept deficits in massively multiplayer online role-playing games addiction. *European Addiction Research, 19*(5), 227–234. https://doi.org/10.1159/000345458

Leménager, T., Neissner, M., Sabo, T., Mann, K., & Kiefer, F. (2020). Who Am I" and "how should I Be": A systematic review on self-concept and avatar identification in gaming disorder. *Current Addiction Reports, 7*(2), 166–193. https://doi.org/10.1007/s40429-020-00307-x

Lemmens, J. S., Valkenburg, P. M., & Gentile, D. A. (2015). The Internet gaming disorder scale. *Psychological Assessment, 27*(2), 567. https://doi.org/10.1037/pas0000062

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2009). Development and validation of a game addiction scale for adolescents. *Media Psychology, 12*(1), 77–95. https://doi.org/10.1080/15213260802669458

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2011). Psychosocial causes and consequences of pathological gaming. *Computers in Human Behavior, 27*(1), 144–152. https://doi.org/10.1016/j.chb.2010.07.015

Liang, H. Y. (2011). *Verifying mmorpg as cultural worldview with mortality salient hypothesis [Masters thesis*. Sun Yat-Sen University] ProQuest One Academic. www.proquest.com/dissertations-theses/verifying-mmorpg-as-cultural-worldview-with/docview/1873866658/se-2?accountid=13963.

Lozic, J., Milkovic, M., & Lozic, I. (2018). *Correlation between mobile games market and the development of video game market in China.36th international scientific Conference on Economic and social development – "building resilient society*. www.proquest.com/conference-papers-proceedings/correlation-between-mobile-games-market/docview/2176198473/se-2?accountid=13963.

Mann, M. M., Hosman, C. M., Schaalma, H. P., & De Vries, N. K. (2004). Self-esteem in a broad-spectrum approach for mental health promotion. *Health Education Research, 19*(4), 357–372. https://doi.org/10.1093/her/cyg041

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin.

McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS peer review of electronic search strategies: 2015 guideline statement. *Journal of Clinical Epidemiology, 75*, 40–46. https://doi.org/10.1016/j.jclinepi.2016.01.021

McGuier, E. A., Kolko, D. J., Klem, M. L., Feldman, J., Kinkler, G., Diabes, M. A., … Wolk, C. B. (2021). Team functioning and implementation of innovations in healthcare and human service settings: A systematic review protocol. *Systematic Reviews, 10*(1), 1–7. https://doi.org/10.1186/s13643-021-01747-w

Mills, D. J., & Allen, J. J. (2020). Self-determination theory, internet gaming disorder, and the mediating role of self-control. *Computers in Human Behavior, 105*, Article 106209. https://doi.org/10.1016/j.chb.2019.106209

Mills, D. J., Milyavskaya, M., Mettler, J., & Heath, N. L. (2018). Exploring the pull and push underlying problem video game use: A self-determination theory approach. *Personality and Individual Differences, 135*, 176–181. https://doi.org/10.1016/j.paid.2018.07.007

Molinos, M. (2016). *The relationship between video game use, internet use, addiction, and subjective well-being [Masters thesis*. California State University]. ProQuest One Academic. www.proquest.com/dissertations-theses/relationship-between-video-game-use-internet/docview/1834503373/se-2.

Moudiab, S., & Spada, M. M. (2019). The relative contribution of motives and maladaptive cognitions to levels of Internet Gaming Disorder. *Addictive Behaviors Reports, 9*, 1–5. https://doi.org/10.1016/j.abrep.2019.100160. Article 100160.

Nagygyörgy, K., Mihalik, Á., & Demetrovics, Z. (2012). *Az online játékok pszichológiai vonatkozásai [Psychological aspects of online games] A média hatása a gyermekekre és fiatalokra*.

National Institute for Clinical Excellence. (2012). *Methods for the development of NICE public health guidance* (3rd ed.) [Process and methods, PMG4]. In: www.nice.org.uk/process/pmg4/chapter/appendix-g-quality-appraisal-checklist-quantitative-studies-reporting-correlations-and.

Oggins, J., & Sammis, J. (2012). Notions of video game addiction and their relation to self-reported addiction among players of World of Warcraft. *International Journal of Mental Health and Addiction, 10*(2), 210–230. https://doi.org/10.1007/s11469-010-9309-y

Orth, U., & Robins, R. W. (2013). Understanding the link between low self-esteem and depression. *Current Directions in Psychological Science, 22*(6), 455–460. https://doi.org/10.1177/0963721413492763

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*(2), 157–159. https://doi.org/10.2307/1164923

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, Article 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Park, H. S., Kwon, Y. H., & Park, K. M. (2007). Factors on internet game addiction among adolescents. *Journal of Korean Academy of Nursing, 37*(5), 754–761. https://doi.org/10.4040/jkan.2007.37.5.754

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science, 7*(6), 528–530. https://doi.org/10.1177/1745691612465253

Peng, L. H., & Li, X. (2009). A survey of Chinese college students addicted to video games. *China Education Innovation Herald, 28*, 111–112.
Petry, N. M., Rehbein, F., Gentile, D. A., Lemmens, J. S., Rumpf, H. J., Mößle, T., … O'Brien, C. P. (2014). An international consensus for assessing internet gaming disorder using the new DSM-5 approach. *Addiction, 109*(9), 1399–1406. https://doi.org/10.1111/add.12457
Polanin, J. R., Pigott, T. D., Espelage, D. L., & Grotpeter, J. K. (2019). Best practice guidelines for abstract screening large-evidence systematic reviews and meta-analyses. *Research Synthesis Methods, 10*(3), 330–342. https://doi.org/10.1002/jrsm.1354
Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior, 45*, 137–143. https://doi.org/10.1016/j.chb.2014.12.006
Raja, H. H., Patel, V. K., Tiwari, D. S., Kadavala, N. C. B., & Patel, N. (2020). Gaming addiction: Study of gaming characteristics and personality traits among the health professional undergraduates. *Annals of Indian Psychiatry, 4*(2), 164. https://doi.org/10.4103/aip.aip_21_20
Reynolds, W. M. (2001). *Reynolds adolescent adjustment screening inventory: RAASI: Professional manual*. PAR.
Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*(2), 151–161. https://doi.org/10.1177/0146167201272002
van Rooij, A. J., Schoenmakers, T. M., Van den Eijnden, R. J., Vermulst, A. A., & van de Mheen, D. (2012). Video game addiction test: Validity and psychometric characteristics. *Cyberpsychology, Behavior, and Social Networking, 15*(9), 507–511. https://doi.org/10.1089/cyber.2012.0007
Rosenberg, M. (1965). *Society and adolescent child*. Princeton University Press. https://doi.org/10.1126/science.148.3671.804
Rosenberg, M. (1989). Self-esteem and adolescent problems: Modeling reciprocal effects. *American Sociological Review, 54*(6), 1004–1018. https://doi.org/10.2307/2095720
Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638. https://doi.org/10.1037/0033-2909.86.3.638
Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications. https://doi.org/10.1521/978.14625/28806
Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*(4), 344–360. https://doi.org/10.1007/s11031-006-9051-8
Saunders, J. B., Hao, W., Long, J., King, D. L., Mann, K., Fauth-Bühler, M., … Poznyak, V. (2017). Gaming disorder: Its delineation as an important condition for diagnosis, management, and prevention. *Journal of Behavioral Addictions, 6*(3), 271–279. https://doi.org/10.1556/2006.6.2017.039
Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84*(3), 328–364. https://doi.org/10.3102/0034654313500826
Scerri, M., Anderson, A., Stavropoulos, V., & Hu, E. (2019). Need fulfilment and internet gaming disorder: A preliminary integrative model. *Addictive Behaviors Reports, 9*, 1–7. https://doi.org/10.1016/j.abrep.2018.100144. Article 100144.
Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology, 89*(4), 623. https://doi.org/10.1037/0022-3514.89.4.623
Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R* (Vol. 4784). Springer. https://doi.org/10.1007/978-3-319-21416-0
Sedikides, C., Gaertner, L., & Cai, H. (2015). On the panculturality of self-enhancement and self-protection motivation: The case for the universality of self-esteem. *InAdvances in motivation science, 2*, 185–241 (Elsevier).
Simmonds, M. (2015). Quantifying the risk of error when interpreting funnel plots. *Systematic Reviews, 4*(1), 1–7. https://doi.org/10.1186/s13643-015-0004-8
Starcevic, V., Choi, T. Y., Kim, T. H., Yoo, S. K., Bae, S., Choi, B. S., & Han, D. H. (2020). Internet gaming disorder and gaming disorder in the context of seeking and not seeking treatment for video-gaming. *Journal of Psychiatric Research, 129*, 31–39. https://doi.org/10.1016/j.jpsychires.2020.06.007
Sterne, J. A., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology, 54*(10), 1046–1055. https://doi.org/10.1016/S0895-4356(01)00377-8
Stetina, B. U., Kothgassner, O. D., Lehenbauer, M., & Kryspin-Exner, I. (2011). Beyond the fascination of online-games: Probing addictive behavior and depression in the world of online-gaming. *Computers in Human Behavior, 27*(1), 473–479. https://doi.org/10.1016/j.chb.2010.09.015
Stevens, M. W., Dorstyn, D., Delfabbro, P. H., & King, D. L. (2021). Global prevalence of gaming disorder: A systematic review and meta-analysis. *Australian and New Zealand Journal of Psychiatry, 55*(6), 553–568. https://doi.org/10.1177/0004867420962851
Tafarodi, R. W., & Milne, A. B. (2002). Decomposing global self-esteem. *Journal of Personality, 70*(4), 443–483. https://doi.org/10.1111/1467-6494.05017
Teng, Z., Pontes, H. M., Nie, Q., Xiang, G., Griffiths, M. D., & Guo, C. (2020). Internet gaming disorder and psychosocial well-being: A longitudinal study of older-aged adolescents and emerging adults. *Addictive Behaviors, 110*, 1–7. https://doi.org/10.1016/j.addbeh.2020.106530. Article 106530.
Thoits, P. A. (1994). Stressors and problem-solving: The individual as psychological activist. *Journal of Health and Social Behavior*, 143–160. https://doi.org/10.2307/2137362
Thomason, S., & Moghaddam, N. (2021). Compassion-focused therapies for self-esteem: A systematic review and meta-analysis. *Psychology and Psychotherapy: Theory, Research and Practice, 94*(3), 737–759. https://doi.org/10.1111/papt.12319
Throuvala, M. A., Janikian, M., Griffiths, M. D., Rennoldson, M., & Kuss, D. J. (2019). The role of family and personality traits in internet gaming disorder: A mediation model combining cognitive and attachment perspectives. *Journal of Behavioral Addictions, 8*(1), 48–62. https://doi.org/10.1556/2006.8.2019.05
Von der Heiden, J. M., Braun, B., Müller, K. W., & Egloff, B. (2019). The association between video gaming and psychological functioning. *Frontiers in Psychology, 10*, 1–11. https://doi.org/10.3389/fpsyg.2019.01731. Article 1713.
Walpole, S. C. (2019). Including papers in languages other than English in systematic reviews: Important, feasible, yet often omitted. *Journal of Clinical Epidemiology, 111*, 127–134. https://doi.org/10.1016/j.jclinepi.2019.03.004
Wang, Y., Zheng, L., Wang, C., & Guo, X. (2020). Attenuated self-serving bias in people with internet gaming disorder is related to altered neural activity in subcortical-cortical midline structures. *BMC Psychiatry, 20*(1), 1–12. https://doi.org/10.1186/s12888-020-02914-4
Wartberg, L., & Kammerl, R. (2020). Empirical relationships between problematic alcohol use and a problematic use of video games, social media and the internet and their associations to mental health in adolescence. *International Journal of Environmental Research and Public Health, 17*(17), 1–11.
Wartberg, L., Kriston, L., Kramer, M., Schwedler, A., Lincoln, T. M., & Kammerl, R. (2017). Internet gaming disorder in early adolescence: Associations with parental and adolescent mental health. *European Psychiatry, 43*, 14–18. https://doi.org/10.1016/j.eurpsy.2016.12.013
Wartberg, L., Kriston, L., Zieglmeier, M., Lincoln, T., & Kammerl, R. (2019). A longitudinal study on psychosocial causes and consequences of Internet gaming disorder in adolescence. *Psychological Medicine, 49*(2), 287–294. https://doi.org/10.1017/S003329171800082X
Wartberg, L., Zieglmeier, M., & Kammerl, R. (2021). An empirical exploration of longitudinal predictors for problematic internet use and problematic gaming behavior. *Psychological Reports, 124*(2), 543–554. https://doi.org/10.1177/0033294120913488
Wei, C., & Luo, Q. (2017). Stressful life events and internet gaming disorder among LeftBehind children: The moderating effect of self-esteem. *Educational Measurement and Evaluation, 6*, 45–51. https://doi.org/10.16518/j.cnki.emae.2017.06.007
Wichstrøm, L., Stenseng, F., Belsky, J., von Soest, T., & Hygen, B. W. (2019). Symptoms of internet gaming disorder in youth: Predictors and comorbidity. *Journal of Abnormal Child Psychology, 47*(1), 71–83. https://doi.org/10.1007/s10802-018-0422-x
Williams, K. D., & Jarvis, B. (2006). Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behavior Research Methods, 38*(1), 174–180. https://doi.org/10.3758/BF03192765
Wittek, C. T., Finserås, T. R., Pallesen, S., Mentzoni, R. A., Hanss, D., Griffiths, M. D., & Molde, H. (2016). Prevalence and predictors of video game addiction: A study based on a national representative sample of gamers. *International Journal of Mental Health and Addiction, 14*(5), 672–686. https://doi.org/10.1007/s11469-015-9592-8
Wölfling, K., Beutel, M. E., & Müller, K. W. (2012). Construction of a standardized clinical interview to assess internet addiction: First findings regarding the usefulness of AICA-C. *Journal of Addiction Research & Therapy, 6*, 1–7. https://doi.org/10.4172/2155-6105.S6-003
World Health Organization. (2018). *International classification of diseases for mortality and morbidity statistics (11th Revision)*. Retrieved from: icd.who.int/browse11/l-m/en.
Xiao, L. Y. (2021). People's Republic of China legal update: The notice on further strictly regulating and effectively preventing online video gaming addiction in minors (published August 30, 2021, effective september 1, 2021). *Gaming Law Review, 25*(9), 379–382. https://doi.org/10.1089/glr2.2021.0026
Yang, X., Huang, B., & Wong, K. M. (2021). Prevalence and socio-demographic, anthropometric, and cognitive correlates of internet gaming disorder among children in China. *Children and Youth Services Review, 122*, 1–6. https://doi.org/10.1016/j.childyouth.2020.105893
You, S., Kim, E., & Lee, D. (2017). Virtually real: Exploring avatar identification in game addiction among massively multiplayer online role-playing games (MMORPG) players. *Games and Culture, 12*(1), 56–71. https://doi.org/10.1177/1555412015581087
Young, K. S. (1998). *Caught in the net: How to recognize the signs of internet addiction–and a winning strategy for recovery*. John Wiley & Sons.
Yu, Y. (2020). *The roles of maladaptive cognitions in understanding internet gaming disorder among adolescents in two Chinese cities: Risk, mediation, and moderation effects [Doctoral thesis, the Chinese University of Hong Kong]*. ProQuest One Academic. www.proquest.com/dissertations-theses/roles-maladaptive-cognitions-understanding/docview/2531569044/se-2?accountid=13963.
Yu, Y., Mo, P. K. H., Zhang, J., Li, J., & Lau, J. T. F. (2020). Validation of the Chinese version of the revised internet gaming cognition scale among adolescents in China: Maladaptive cognitions as potential determinants of internet gaming disorder. *International Journal of Environmental Research and Public Health, 17*(1), 290. https://doi.org/10.3390/ijerph17010290
Zastrow, M. (2017). News Feature: Is video game addiction really an addiction? *Proceedings of the National Academy of Sciences, 114*(17), 4268–4272. https://doi.org/10.1073/pnas.1705077114

Zhang, L., & Fung, A. Y. (2014). Working as playing? Consumer labor, guild and the secondary industry of online gaming in China. *New Media & Society, 16*(1), 38–54. https://doi.org/10.1177/1461444813477077

Zhang, Y., Mei, S., Li, L., Chai, J., Li, J., & Du, H. (2015). The relationship between impulsivity and internet addiction in Chinese college students: A moderated mediation analysis of meaning in life and self-esteem. *PLoS One, 10*(7), Article e0131597. https://doi.org/10.1371/journal.pone.0131597

Zhou, S. X., & Leung, L. (2012). Gratification, loneliness, leisure boredom, and self-esteem as predictors of SNS-game addiction and usage pattern among Chinese college students. *International Journal of Cyber Behavior, Psychology and Learning, 2* (4), 34–48. https://doi.org/10.4018/ijcbpl.2012100103