Loot boxes, problem gambling, and problem video gaming: a systematic review and meta-synthesis

NB: This is the author’s accepted manuscript version of this paper.

Abstract
Loot boxes are video game-related purchases with a chance-based outcome. Due to similarities with gambling, they have come under increasing scrutiny from media, academics and policymakers alike. Initial evidence suggested that loot box (LB) engagement might be associated with both problem gambling (PG) and problem video gaming (PVG). We therefore conducted a systematic review of the evidence for associations between LB purchasing, PG and PVG. For LB/PG, 12 of 13 publications reported a positive relationship, with a moderately sized mean effect of $r = 0.27$. For LB/PVG, the mean effect was $r = 0.40$, although this finding was drawn from only six surveys in total. For PG/PVG, the mean effect was $r = 0.21$, with only 11 of 20 studies reporting significant effects. Whilst further evidence is required to determine direction of causality, the strength of relationships suggests that policy action on loot boxes may have benefits for harm minimisation.
Background

Loot boxes (LB) are video game-related purchases with a chance-based outcome. They come in various configurations and guises, and might be called boxes, crates, cases, card packs, shards or eggs. Due to these complexities, there is often confusion about what, exactly, constitutes a loot box. Our definition encapsulates any purchasable game content with a randomised outcome. The contents within loot boxes might include ‘pay to win’ features such as improved weapons or power ups, or instead, they might be purely cosmetic upgrades and customisation. Loot boxes may also be purchased with in-game currency, or they might be opened using purchasable ‘keys’, such as the ‘crate and key’ mechanics found in games like Counter Strike: Global Offensive or the similar ‘eggs and incubator’ mechanics found in Pokémon Go. Furthermore, ‘free’ loot boxes are offered in some games – but these are also always available as paid-for content (i.e. or else they would not be classified as loot boxes under our definition, because there is no opportunity for monetary exchange). These sorts of free giveaways are often aimed at encouraging future loot box purchasing (Zendle, Meyer and Over, 2019).

Loot boxes have become increasingly prevalent over the last decade (Zendle, Meyer, Cairns, et al., 2019), are available in the majority of video games across various formats (Zendle, Meyer, Cairns, et al., 2019), and have become increasingly sophisticated (Koeder et al., 2018). An estimated 44-78% of gamers are thought to have purchased loot boxes (Brooks and Clark, 2019; Li et al., 2019; Zendle and Cairns, 2018). Widely available to children (Zendle, Meyer, Cairns, et al., 2019), they have come under increasing scrutiny from academics, policymakers, the media (Drummond, Sauer, Ferguson, et al., 2020), and the gaming community itself (Allan, 2018), due to structural and psychological similarities with gambling (Drummond and Sauer, 2018).

The first academic studies on loot boxes were published in 2018 (Drummond and Sauer, 2018; Zendle and Cairns, 2018, 2019), where the first survey of loot box purchasing appeared to confirm such concerns: establishing a significant, moderate sized correlation between problem gambling severity and loot box purchasing (Zindle and Cairns, 2018). Due to the correlational nature of the evidence, the direction of the relationship is unknown, but there are three possibilities. Either (a) problem gamblers purchase more loot boxes; (b) loot box purchasers are more likely to start gambling – i.e. via ‘gateway effects’; or (c) there is a complex, dynamic relationship between the two behaviours.

There has been considerable debate about the extent to which loot boxes constitute gambling. Some senior gaming executives have argued that since loot boxes always contain an item, they are the digital analogue of baseball cards or Kinder Eggs (Digital, Culture, Media and Sport Committee, 2019). However, analysis of game-related marketplaces such as the Steam Community Market (where digital assets obtained in loot boxes can be bought and sold) reveals that the items obtained within loot boxes usually constitute a monetary loss (Drummond, Sauer, Hall, et al., 2020). This type of ‘loss disguised as a win’ has parallels with traditional gambling games such as online slots, where ‘wins’ can actually be lower than the stake wagered. Some academics have therefore argued that loot boxes could be regulated under existing gambling laws (Drummond, Sauer, Hall, et al., 2020).

Whilst loot boxes are a relatively new innovation in video gaming, controversies around links between video games and gambling predate the emergence of loot boxes by several decades (Iain
and Brown, 1989). It has been hypothesised that both problematic video gaming (PVG) and problem gambling (PG) may be driven by shared cognitive biases or distortion (Fisher, 1993; Gupta and Derevensky, 1996; Iain and Brown, 1989). Whilst there is some debate about classifying problem video gaming as a psychological disorder (Aarseth et al., 2017; Kuss et al., 2017), Gaming Disorder is included in the ICD-11 (World Health Organisation, 2018), and the DSM-5 (American Psychiatric Association, 2013) recommends Internet Gaming Disorder as a condition requiring further research.

Similar to gambling, video gaming often operates on chance-based principles, reinforced by the physiological arousal of aesthetic elements such as lights, sounds and graphics. It has already been suggested that habituation and repeated exposure to video games (i.e. regardless of loot boxes) could lead adolescents, via gateway effects, to develop later gambling problems (Fisher, 1993; Gupta and Derevensky, 1996; Iain and Brown, 1989).

Early research appeared to support such conclusions (Gupta and Derevensky, 1996; Ladouceur and Dubé, 1995). However, these studies were conducted during the era of “coin-op[erated]” arcade games, often co-located with similar-looking gambling devices. Today is an era of increasing domestic proximity between gaming and gambling (i.e. both activities are increasingly available within the home). Loot boxes are one aspect of this ongoing technological and cultural gambling-gaming convergence (Brooks and Clark, 2019; King et al., 2015), with increasing prevalence of in-game casinos, free-to-play ‘social’ casinos (Kim et al., 2015), eSports betting (Macey and Hamari, 2019a), crypto-games (Scholten et al., 2019), and game-related wagering (Abarbanel and Macey, 2019).

Loot boxes provide a well-defined opportunity to understand how such gambling-related features interact with problematic gaming and gambling; and longer-terms harms associated with these behaviours. We conducted an independent, systematic literature review, to assess the consistency, robustness, effect sizes and quality of evidence (Moher et al., 2009) for associations between the three constructs of interest: loot boxes and problem gambling (LB/PG); loot boxes and problem video gaming (LB/PVG); and problem gambling and problem video gaming (PG/PVG).

**Methods**

A systematic review and, where possible, a meta-analysis was conducted using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA (Moher et al., 2009)) to account for survey heterogeneity, infrequent reporting of standard deviation/error and cross-sectional nature of evidence (see below).

**Literature searches**

We searched literature databases for publications related to loot boxes (Search 1), and also for links between PG and PVG (Search 2). See Box 1 for full search terms. Initial searches were conducted prior to 28th March 2020. However, further/subsequent papers were included via expert knowledge, publication alerts, and a snowballing approach (i.e. via the references of primary articles). Non-English articles were excluded and duplicates were removed. For Search 1, further papers were
excluded if they were not relevant to loot boxes, only discussed loot boxes as a peripheral subject, or were publications of a non-empirical nature (e.g. reviews, commentaries and book chapters).

For Search 2, we targeted papers investigating associations between gambling and video gaming. Studies were excluded if they: did not measure both gambling and gaming; were pre-year 2000 (i.e. generally pre-internet, and often “coin-op”); were specific to a certain gaming context (i.e. social casinos, which are free-to-play structural homologues of online casinos and known to have high migration rates into online gambling (Gainsbury et al., 2016; Kim et al., 2015)); or failed to investigate direct correlations between gaming and gambling. Screening was divided between SGS, LLN, JC and JL; with at least two researchers assessing all primary research.

Results synthesis

For surveys of gamers reporting associations between our three constructs of interest, we extracted key statistical results. Due to heterogeneity of reporting, it was often necessary to convert results into a standardised Pearson’s r value, using (where possible) the method specified by Lenhard and Lenhard (2016) or otherwise alternate conversions (Gilpin, 1993; Lin, 2020). Where publications reported multiple statistical tests (i.e. bivariate correlations and path analysis, e.g. Molde et al. (2019)), we extracted bivariate Pearson’s r values, as the most commonly (and directly comparable) statistic in the literature.

We transformed each r value using Fisher’s z, and then calculated a weighted mean z value (using the Hunter-Schmidt method (Field and Gillett, 2010)) for the three associations of interest (LB/PG, LB/PVG and PG/PVG). These mean z values were then reverse transformed to provide the final weighted mean r values. Stem and leaf plots were generated as a visual proxy of the consistency/reproducibility of studies (Field and Gillett, 2010). Other meta-analytical approaches (i.e. confidence intervals; forest plots; funnel plots) could not be calculated due lack of available data: standard deviation/errors were often not reported in the literature, and not sufficiently available when contacting authors directly (only 33% authors responded with necessary data).

Quality appraisal

We utilised the Mixed Methods Appraisal Tool (MMAT (Hong et al., 2018)) to assess the quality of evidence, using items relating to quantitative descriptive studies: 4.1 = appropriate sampling strategy; 4.2 = representative population; 4.3 = appropriate measures; 4.4 = low nonresponse bias; 4.5 = appropriate statistical approach. For item 4.3, we scored all methods of LB measurement positively (as yet, there is no ‘gold standard’ approach), and for gambling/gaming measures, we scored positively for all previously validated scales (e.g. PGSI (Problem Gambling Severity Index), IGD (Internet Gaming Disorder)) and negatively for one-off, non-validated scales or measures of gaming/gambling participation. Additionally, we assessed whether publications were peer-reviewed (due to existence of pre-print papers; item 4.6), if publications were pre-registered (or a replication study; item 4.7), if data was open access (item 4.8), and if there was a low risk of cohort response bias (item 4.9). This final assessment judged from the perspective of our review questions (i.e. associations between LB/PG/PVG in the general population). Here, some studies had a relevant sampling strategy for the goals of the primary study (e.g. elite athletes (Håkansson et al., 2018)) and therefore scored appropriately for item 4.1 on the MMAT, but were limited/biased cohorts for
investigating such population-wide associations. Our full checklist was assessed by two researchers (JC and SS), with any discrepancies discussed and resolved.

**Results & Discussion**

**Literature searches**

For Search 1 (LB related searches), we identified 256 publications, although only 23 of these were primary research. See Figure 1 for identification, screening, and inclusion of publications. Of the 23 empirical publications, 18 were surveys of gamers, including some measure of loot box behaviour. In addition to the survey data, there were three scoping papers (Drummond and Sauer, 2018; Kaneko et al., 2019; Zendle, Meyer, Cairns, et al., 2019). These investigated issues such as the size and scope of loot boxes in gaming. Also, two preliminary lab-based experimental studies (Brady and Prentice, n.d.; Larche et al., 2019), observed gamers opening loot boxes. The remaining 233 publications were non-empirical, consisting of commentaries, editorials and informal reviews.

From the 18 surveys, 14 reported associations between some combination of LB, PG and/or PVG. Of the remaining four papers, three did not report associations with PG or PVG (Kristiansen and Severin, 2019; Macey and Hamari, 2019b; Seo et al., 2019), and one did not clearly differentiate between randomised and non-randomised in-game purchases (Shibuya et al., 2019). For publications that reported correlations, these included LB/PG (13 studies), LB/PVG (6 studies) and PG/PVG (3 studies) with a small number of studies also reporting measures of wellbeing, impulsivity and motivations (Drummond, Sauer, Ferguson, et al., 2020; Li et al., 2019; Zendle, Meyer and Over, 2019).

For Search 2, we identified an additional 17 surveys that explored associations between PG/PVG. The relationships between these constructs are summarised in Figure 2. An overview of publications and the results of associations between these three constructs (i.e. LB/PG, LB/PVG and PG/PVG) is in Table 1.

**Evidence synthesis**

The 13 publications (14 studies) investigating LB/PG (total n = 25,334; n usable for meta-synthesis = 19,334) were all cross-sectional, apart from one longitudinal survey (Zendle, 2019c). Thirteen (93%) studies established positive associations, with a moderate weighted effect size of $r = 0.27$. There was a heterogeneity of measurement approach (e.g. measuring loot box spend, engagement or frequency). Whilst PGSI was used to measure PG in all but two studies (Kristiansen and Severin, 2020; Zendle, Meyer and Over, 2019), studies varied in using numerical PGSI scores or categories (i.e. problem gambling status).

The six studies investigating LB/PVG (total n = 4941; meta-synthesis n = 3,433) established a larger weighted effect size ($r = 0.40$), but with only 57% of studies revealing positive associations. The studies were relatively heterogeneous, although all were cross-sectional, and PVG was measured
using the IGD in all but one (von Meduna et al., 2020). One study \( n = 1508 \) (von Meduna et al., 2019) with mixed results was omitted from our analysis, because it was not possible to generate a standardised \( r \) value.

The 20 studies investigating PG/PVG (total \( n = 24,514 \); meta-synthesis \( n = 23,794 \)) established a moderate weighted effect size of \( r = 0.21 \). Being drawn from a longer time period (i.e. last 20 years, rather than only the last 3 years for LB literature), these studies were the most methodologically heterogeneous, using PVG measures such as IGD, GAS (Gaming Addiction Scale) and GAIT (Gaming Addiction Identification Test); and PG measures such as PGSI, CPGI (Canadian Problem Gambling Index) and SOGS (South Oaks Gambling Screen). Eight studies used PGSI, while five used IGD. All but two studies were cross-sectional. Over half (57\%) of the PG/PVG studies established positive associations. For two further studies, it was not possible to generate a standardised \( r \) value. See Figure 2 and Table 1 for further details.

Across all studies, there was a large variation in cohorts. Samples sizes ranged from 112 to 7,422, sourced from a variety of forums and participation pools, with a large geographic spread (including Europe, America and Australia). Extended details of surveys are provided in Supplementary Table 1. A summary of the associations is provided in Table 2.

The stem and leaf plots of individual associations (Figure 3), illustrate the reason for apparent discrepancies between effect sizes and reproducibility. Whilst the LB/PG associations (which were consistently reported) cluster around the mean value, the associations between LB/PVG and PVG/PG exhibit a broader spread of results (i.e. a greater number of higher and lower values). Potential reasons for this are explored in the discussion below.

**Evidence quality**

According to the MMAT, the quality of the evidence (see Table 1) provided by publications on loot boxes was of moderate quality, but improving over time. They currently include pre-registered/replication studies (4 publications), open-access data (7 studies) and representative samples (4 studies); including a nationally representative sample of around 1,100 Danish adolescents (Kristiansen and Severin, 2020); and a nationally representative sample of around 1,100 UK adults (Zendle, 2019a).

Publications derived from the secondary searches (i.e. non-loot box papers, focusing on PG/PVG interactions) covered a broader span of time, and were generally of lower quality than the loot box publications, with no studies being pre-registered or open access, although five publications did include nationally representative samples.

**Loot Boxes and Problem Gambling**

These publications found positive associations, apart from one with mixed findings (a positive effect in a first Study of 144 adults, but no evidence of an association in a second (relatively small sample) study of 113 Canadian undergraduate students (Brooks and Clark, 2019)). Replications have been
demonstrated across various cohorts, nationalities (Europe, North America and Australasia) and age groups (including younger adolescents (Kristiansen and Severin, 2020), older adolescents (Zendle, Meyer and Over, 2019) and adults). The magnitude of the relationship is not significantly altered after controlling for age and gender (Drummond, Sauer, Ferguson, et al., 2020; Kristiansen and Severin, 2020; Li et al., 2019).

Specific findings of these papers further support the conclusion that loot box purchasing and gambling are related behaviours. Firstly, loot box purchasing is predicted by gambling related cognitions (e.g. illusions of control (Raylu and Oei, 2004)); implicated as an etiological factor in gambling disorder (Brooks and Clark, 2019)). Secondly, associations with problem gambling are weakest with free loot boxes (Kristiansen and Severin, 2020; Zendle and Cairns, 2018), slightly stronger for loot boxes that use in-game currency or show near misses (Zendle et al., 2020), and stronger when players are actively engaged in selling loot boxes on secondary marketplaces (Kristiansen and Severin, 2020; Larche et al., 2019; Zendle et al., 2020). This suggests that loot box purchasing has a closer link with gambling, when the contents have potential monetary value. Thirdly, in a naturalistic follow-up study, the removal of loot boxes (in the game Heroes of the Storm) resulted in problem gamblers spending significantly less money; although this was a small cohort study, with longitudinal data giving mixed, model dependent results (Zendle, 2019c). Fourthly, the strength of association is only moderately altered by the various available configurations of loot boxes, including options such as showing near misses, use of in-game currencies, and cosmetic versus pay2win content (Kristiansen and Severin, 2020; Zendle and Cairns, 2018, 2019).

Observable differences in survey results (Figure 3) were likely related to heterogeneity of: samples (different nationalities and ages, ranging from convenience gamer samples to demographically representative cohorts); instruments (e.g. to measure gaming and gambling); units of comparison (e.g. scale data versus categorical ‘problem’ status); loot box measurements (e.g. spend, engagement, or the Risky Loot Box Index scale of problematic loot box behaviour (Brooks and Clark, 2019)); analytical methods (e.g. Pearson/Spearman correlations, Kruskal Wallis H Test, Kendall’s tau, ANOVA, Mann-Whitney U test, bivariate and multiple regressions). However, there was less survey heterogeneity for LB/PG than with PG/PVG (see below).

Loot Boxes and Problem Video Gaming

The LB/PVG results are more mixed than for LB/PG (see Table 1 and Figure 3), an issue that is likely related to both a greater survey heterogeneity (see Supplementary Table 1), and a smaller number of published surveys. Moreover, the larger effect size is influenced by the presence of two larger cohort studies (see Table 1 (Drummond, Sauer, Ferguson, et al., 2020; Zendle, 2019a)) – although one of these studies was conducted with a nationally representative sample (Zendle, 2019a). Furthermore, another (n = 1508) survey with mixed results (von Meduna et al., 2019) was excluded from our analysis, because it was not possible to produce an r statistic from unadjusted (regression) coefficients using available methods (Gilpin, 1993; Lenhard and Lenhard, 2016; Lin, 2020). Further loot box studies (measuring PVG) are needed, to confirm whether such associations are reproducibly larger than those with PG - it is important to know which of these problematic behaviours is most predictive of loot box purchasing, in order to develop future interventions accordingly.
The survey with mixed results (von Meduna et al., 2019), which used a cohort of German pay2win players, illustrates how unit of analysis (i.e. categorical versus scale data) influences results. Here, high risk gaming (rather than gambling) was more strongly associated with the initial decision to purchase a loot box (i.e. a binary yes/no). However, with loot box purchasing frequency, the finding was reversed. Instead, there was no association with gaming, but there was a significant association with high risk gambling. In other words, whilst problem gaming influenced the decision to purchase loot boxes, once this threshold had been crossed, problem gambling predicted the frequency of purchases.

A further study, of spending in Fortnite (before loot boxes were removed from this game) (King et al., 2020), found no significant association between loot box spend and gaming disorder status, and instead found that spending was driven by social influences. Conflicting with the other studies, this result (along with (von Meduna et al., 2019)) highlights the context-dependent nature of relationships between loot boxes, gaming and gambling, mediated via the multiway interactions between these similar underlying constructs, alongside additional, external motivations (e.g. social influences).

**Problem Video Gaming and Problem Gambling**

There are several possible reasons for the less consistent results for PG/PVG than LB/PG (Table 1 and Figure 3). First, some results are from limited cohorts (e.g. sexual minorities (Broman and Hakansson, 2018) or elite athletes (Håkansson et al., 2018)), which may not be generalizable. Second, in some studies, the association between PG/PVG was of negligible magnitude after controlling for age and gender (Biegun et al., 2020; Forrest et al., 2016), or when considered alongside other factors, such as time spent gaming, gaming motivations and feelings of social alienation (Biegun et al., 2020). Such results suggest that uncontrolled, bivariate correlations may be overstating the links between video games and gambling – and in some studies reporting positive associations the effect size was small (Fu and Yu, 2015; Wood et al., 2004). In contrast, associations between loot boxes and problem gambling are not substantially altered after controlling for gender and age (Brooks and Clark, 2019; Drummond, Sauer, Ferguson, et al., 2020; Kristiansen and Severin, 2020; Li et al., 2019).

Third, these studies were even more heterogeneous than LB studies, and conducted over a larger timeframe. The two longitudinal studies provide a good example. In one, a study of Swedish adolescents, PVG was not a likely predictor of later PG (Vadlin et al., 2018)). Conversely, a longitudinal study using a representative Norwegian cohort, established a relationship between PVG and later PG, but no evidence of the reverse relationship (Molde et al., 2019). Without further data, explaining these disparate conclusions is speculative, but may be due to the younger cohort of the Swedish sample (with limited access to gambling), or the Norwegian cohort (who had already started gambling).

Finally, in some samples, gaming and gambling may be subject to exclusivity effects. In a study of Swedish eSports enthusiasts, PVG and PG had an inverse relationship (Macey and Hamari, 2018). This may have been due to the “dedicated” gamer cohort being more sceptical of the chance-based, uncontrolled features in games. Such exclusivity effects may explain some non-significant associations, where one Australian survey established that gambling was a generally unpopular
activity amongst video gamers (Forrest et al., 2016). Another study established that gamers reported lower enjoyment of gambling, and (contrary to gamblers) were less likely to attribute gambling wins to illusion of control or superstitious explanatory factor (King et al., 2012).

Such findings illustrate that the interactions between gaming and gambling may be operating in both asymmetrical and non-linear fashions. Heavy gaming (in certain contexts) may have exclusivity effects on gambling. More research is needed to determine whether the inverse may be true. Furthermore, these relationships may not scale linearly – i.e. “dedicated” gamers and gamblers may behave differently to “recreational” players. Here, the majority of studies with positive (or mixed) associations used either convenience or nationally representative samples (13/15), whereas studies establishing no (or negative) links between the two activities used either dedicated gamer or gambler cohorts (4/5 studies; Table 1). Such complex interactions, combined with variability in cohorts, measurement and study design, may explain the inconsistently reproduced relationships between video gaming and gambling.

**Loot Boxes, Problem Gambling and Problem Video Gaming**

In an attempt to better understand the complex interactions between these behaviours, three surveys have recently investigated LB purchasing within the broader context of both PVG and PG (Brooks and Clark, 2019; Drummond, Sauer, Ferguson, et al., 2020; Li et al., 2019). All three studies established significant correlations, of moderate effect size, between all three behaviours (see Table 1). Furthermore, the one pre-registered, demographically representative sample (Drummond, Sauer, Ferguson, et al., 2020) confirmed that a combination of excessive gaming and problem gambling interact additively to drive even greater increases of LB purchasing. Similarly, a path analysis (Li et al., 2019) revealed that LB purchasing was independently related to both PVG and PG severity, whilst also indirectly influenced by the increased gaming associated with PG.

**Other risk factors**

Beyond PG, risk factors for increased LB purchasing may include male sex (Kristiansen and Severin, 2020), impulsivity - although evidence is currently weak (Zendle, Meyer and Over, 2019) or negative (King et al., 2020), and gambling-related cognitions (Brooks and Clark, 2019). Furthermore, there may be an inverse relationship with age. Two surveys of adolescent loot box behaviour established links with PG (Kristiansen and Severin, 2020; Zendle, Meyer and Over, 2019), and in one, the link was more than twice as strong as with adults (Zendle, Meyer and Over, 2019). There has been much commentary around the potential dangers of loot boxes for young people, who are more vulnerable (Fisher, 1993; Griffiths, 1999) and have a higher prevalence of problem gambling (Calado et al., 2017; Forrest and McHale, 2018; Splevins et al., 2010).

Any dangers posed by loot box engagement could be a particular issue for young people, to whom they are widely available (Drummond and Sauer, 2018; Zendle, Meyer, Waters, et al., 2019), with around 25%-40% of UK children and adolescents purchasing them (Gambling Commission, 2018; Ipsos / Gambling Commission, 2019; Royal Society for Public Health, 2019). However, with a lack of longitudinal studies, any gateway effects remain entirely hypothetical. Moreover, any associations between LB and PG may be explained by the shared structural mechanics (Gainsbury, 2019), rather
than any direct causation. These potential shared mechanics might be mediated by underlying extraneous variables such as impulsivity (Cosenza et al., 2019), although further research is needed.

Conclusions

Associations between LB/PG have been repeated across a variety of cohorts, nationalities, and age groups, with improving study designs (including open-science, pre-registered studies and nationally representative cohorts). Whilst associations between LB/PVG and PG/PVG are of similar or larger magnitude, the results are more mixed. This likely reflects the greater heterogeneity of study methods. Nonetheless, the magnitude of all three associations are of statistical and practical significance (Ferguson, 2009). These relationships are often stronger than relationships between PG and well-established co-morbidities, including depression ($\rho = 0.10$) and major drug problems ($r = 0.12$) (Feigelman et al., 1995; Zendle and Cairns, 2018), and comparable to the relationship between problem gambling and current alcohol dependence (Welte et al., 2006).

The results suggest that PVG, PG and high LB engagement are related behaviours, comprised of a mix of overlapping (and divergent) drivers, pathways and possible dangers. The three publications that have investigated all three behaviours highlight a complex interaction of direct and indirect effects. Furthermore, preliminary evidence has linked such associations with psychological harm, albeit a finding that is indirect (Li et al., 2019) and cautiously interpreted by the authors (Drummond, Sauer, Ferguson, et al., 2020). Further research into downstream harms is required.

Furthermore, the cross-sectional nature of all these surveys means results are purely correlational, and there is no way to distinguish between the various alternatives. Problem gambling might drive increased loot box purchasing. Alternatively, loot box purchasing – via the exposure and habituation of ‘gateway effects’ – might lead to a higher incidence of future gambling problems. Or, perhaps more likely, the associations might be the result of bidirectional links between all these related behaviours – where, for example, we already know that there are complex relationships between gambling and other risky behaviours (Derevensky et al., 1996; Forrest and McHale, 2018).

Strengths, limitations and future research recommendations

We could not conduct more sophisticated meta-analytical approaches, due to lack of data availability (see methods). Furthermore, evidence was largely correlational, and many researchers have emphasised that longitudinal studies are required (Kristiansen and Severin, 2020; Li et al., 2019; Zendle and Cairns, 2019). However, gateway effects are difficult to prove, even with longitudinal studies: temporality, just like correlation, does not prove causation (Kandel, 2002).

Complex, dynamic processes drive interactions between gaming and gambling. Future studies will need to control for demographic, psychological, gambling and gaming profiles, and utilise open science approaches, validated scales, and agreed standards for measuring loot box engagement. Studies with representative samples should investigate influences on wellbeing, financial harms, and long-term outcomes (Li et al., 2019; Zendle, 2019a), especially in children and adolescents. At present, only six studies have investigated LB/PVG; and only three have investigated LB/PG/PVG.
Surveys including all three constructs are required, along with attention to units of analysis and cohorts. Such steps may allow future research to confirm potential explanations of the existing data. For example, it may be that (a) whilst problem video gaming drives the initial decision to purchase loot boxes, problem gambling drives the level of engagement; and (b) that there are exclusivity effects in “dedicated” gamer cohorts, rendering them less liable to problem gambling. Future studies should allow for a clearer understanding of these processes, whilst unpicking some of the inconsistencies in the literature.

Some limitations of survey data could be overcome with other research methods, including lab-based experimental work, objective records of expenditure, and qualitative research. For example, extended qualitative interviews would enable exploration of gamers’ perceptions of the gaming/gambling interface, along with perceptions of causality and mechanisms of impact, and exploration of meanings of value beyond current legal standards of “money’s worth” (Digital, Culture, Media and Sport Committee, 2019).

Finally, loot boxes represent only one category of game monetisation, alongside items including downloadable content, season passes, events, skins and weapons. Other forms of gaming-gambling convergence overlap with this monetisation ecosystem (e.g. in-game casinos, eSports betting, crypto-games, and game-related wagering).

**Policy implications**

Numerous academic researchers (Drummond et al., 2019; Zendle, Meyer and Over, 2019), public organisations (Office of the Children’s Commissioner, 2019; Royal Society for Public Health, 2019), charities (A Parent Zone Report, 2019) and even senior games developers (Perez, 2020) have called for industry and government action on loot boxes. Responses have included legislation in Belgium, The Netherlands and Denmark, with several other nations (including the UK, Australia, Sweden and the USA) now also proposing legislation (McCaffrey, 2019). Our findings – establishing reproducible links between loot box purchasing, and both problem gambling and problem video gaming – suggest that such policy may be warranted for harm minimisation purposes. Whilst the cross-sectional nature of the evidence means that directions of causality cannot be established, from the perspectives of harm-minimisation, such questions are something of a moot point: either loot boxes cause problem gambling, or at-risk individuals (such as problem gamblers and gamers) are more engaged with loot boxes. Either way, policy action may have utility for harm minimisation.

This policy need not extend to offline ‘lucky dip’ activities; where there is evidence to distinguish loot boxes from the traditional ‘surprise mechanics’ of Kinder Eggs or trading card games. One study investigating players of an offline collectible card game, for example, could not reproduce any associations with problem gambling (Zendle, 2019b), suggesting that loot boxes qualitatively differ from these traditional contexts. These differences include the greater scale, scope, availability and technological sophistication of loot boxes (Zendle, 2019b), leading to a ‘continuous play’ effect that is not seen with traditional ‘surprise’ games. Policy on loot boxes need not extend to all offline activities.

Loot boxes, represent the most obvious face of an increasing convergence of gambling and gaming (Derevensky and Griffiths, 2019; Johnson and Brock, 2019). For example, an analysis of patents (King
et al., 2019) has revealed that some game designers are engaging in practices that specifically target psychological tendencies with so-called ‘dark-nudges’ (Ekre, 2015; Harris, 2016; Hodent, 2019; Jernström, 2016), and players can also be manipulated in ways that would be considered illegal in the context of traditional gambling (King et al., 2019; King and Delfabbro, 2018). In the dynamic and rapidly evolving world of video games – and with large financial incentives for potentially unscrupulous developers – any legislation against loot boxes is in danger of quickly being rendered anachronistic. Longer term mitigation of risk will require provisions for ongoing research, development of child-focused data protection (such as the Information Commissioner’s Office ‘Age Appropriate Design Code’ (Information Commissioner’s Office, 2020)), and, finally, educational approaches designed to curb the exploitation of psychological nudges and biases (A Parent Zone Report, 2019; King and Delfabbro, 2018).
Figures and Tables

**Box 1.** Terms used for literature searches. Search 1 shows different combinations of searches relating to loot boxes. Search 2 shows searches relating to problem gaming and problem gambling. Whilst the syntax in different databases varies, the logic of searches is consistent with details below.

**Search 1: loot boxes**

“loot box” OR “loot boxes”

(microtransactions OR microtransaction) AND “chance based” AND games AND (video OR computer OR online OR mobile)

(microtransactions OR microtransaction) AND “chance based” AND gaming AND (video OR computer OR online OR mobile)

gaming AND gambling AND “reward schedule” AND variable AND online (video OR computer OR mobile)

**Search 2: problem gaming and problem gambling**

(“problem gambling” or “gambling addiction”) AND (“problem video gaming” OR “problem gaming” OR “gaming addiction”)

---

**Figure 2.** Surveys reporting associations between Loot Boxes, Problem Gambling and Problem Video Gaming. The proportions (in red) at the intersection of each circle represent the number of papers reporting significant correlations between these two constructs; with n the total number of participants. The number at the central intersection is the publications reporting all three associations (i.e. LBs v PG, LBs v PVG and PG v PVG)
Figure 3. Stem and leaf plots of the three key associations. These represent the distribution of converted effect sizes, rounded to two decimal places. The values to the left of each line represent where each value sits to the first decimal place, while the values of the right of the line represent the second decimal place of each value e.g. the bottom value for LB/PG is \( r = 0.49 \). Statistically significant results are shown in green, null results in red, mixed results in yellow, and reverse correlations in deep red (i.e. statistically significant, but negatively correlated).
Table 1. Primary publications on LB/PG, LB/PVG and PG/PVG, including standardised $r$ values (with method of conversion for associations), and MMAT Quality Analysis. Standardised $r$ values are listed in the columns LB/PG, LB/PVG and PG/PVG, where statistically significant results are shown in green, null results in red, mixed results in yellow, and reverse correlations in deep red (i.e. statistically significant, but negatively correlated). For the MMAT Quality analysis, green = positive for this checklist item; red = negative. The checklist items are as follows: 4.1 = appropriate sampling strategy; 4.2 = representative population; 4.3 = appropriate measures; 4.4 = low nonresponse bias; 4.5 = appropriate statistical approach. We also included an additional 4 items on the checklist, see methods, where: 4.6 = peer-reviewed; 4.7 = pre-registered/replication study; 4.8 = open access data; 4.9 = low risk of cohort response bias. See Supplementary Table 1 for full details.
<table>
<thead>
<tr>
<th>Author and date</th>
<th>Design</th>
<th>n</th>
<th>LB/PG</th>
<th>LB/PVG</th>
<th>PVG/PG</th>
<th>Conversion Method</th>
<th>4.1</th>
<th>4.2</th>
<th>4.3</th>
<th>4.4</th>
<th>4.5</th>
<th>4.6</th>
<th>4.7</th>
<th>4.8</th>
<th>4.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kristiansen and Severin (2020)</td>
<td>Cross-sectional</td>
<td>1,137</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macey and Hamari (2019a)</td>
<td>Cross-sectional</td>
<td>582</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wardle and Zendle (2020)</td>
<td>Cross-sectional</td>
<td>3,549</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zendle (2019c)</td>
<td>Longitudinal</td>
<td>112</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zendle and Cairns (2018)</td>
<td>Cross-sectional</td>
<td>7,422</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zendle and Cairns (2019)</td>
<td>Cross-sectional</td>
<td>1,172</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zendle et al. (2019)</td>
<td>Cross-sectional</td>
<td>1,155</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zendle et al. (2020)</td>
<td>Cross-sectional</td>
<td>1,200</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Von Meduna et al. (2019)</td>
<td>Cross-sectional</td>
<td>1,508</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zendle (2019a)</td>
<td>Cross-sectional</td>
<td>1,081</td>
<td>0.14</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brooks and Clark (2019)</td>
<td>Cross-sectional</td>
<td>144</td>
<td>0.23</td>
<td>0.18</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drummond et al. (2020)</td>
<td>Cross-sectional</td>
<td>1,049</td>
<td>0.34</td>
<td>0.6</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al. (2019)</td>
<td>Cross-sectional</td>
<td>618</td>
<td>0.19</td>
<td>0.27</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>King et al. (2020)</td>
<td>Cross-sectional</td>
<td>428</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biegun et al. (2020)</td>
<td>Cross-sectional</td>
<td>651</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broman and Håkansson (2018)</td>
<td>Cross-sectional</td>
<td>605</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delfabbro et al. (2009)</td>
<td>Cross-sectional</td>
<td>2,669</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ford and Håkansson (2020)</td>
<td>Cross-sectional</td>
<td>2,038</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forrest et al. (2016)</td>
<td>Cross-sectional</td>
<td>485</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fu &amp; Yu (2015)</td>
<td>Cross-sectional</td>
<td>700</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Håkansson et al. (2018)</td>
<td>Cross-sectional</td>
<td>352</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jimenez-Murcia et al. (2014)</td>
<td>Cross-sectional</td>
<td>193</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karlsson et al. (2019)</td>
<td>Cross-sectional</td>
<td>1,593</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>King et al. (2012)</td>
<td>Cross-sectional</td>
<td>115</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macey and Hamari (2018)</td>
<td>Cross-sectional</td>
<td>613</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McBride and Derevensky (2016)</td>
<td>Cross-sectional</td>
<td>1,229</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mills et al. (2020)</td>
<td>Cross-sectional</td>
<td>1,621</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molde et al. (2019)</td>
<td>Longitudinal</td>
<td>4,601</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vadlin et al. (2018)</td>
<td>Longitudinal</td>
<td>1,576</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walther et al. (2012)</td>
<td>Cross-sectional</td>
<td>2,553</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wood et al. (2004)</td>
<td>Cross-sectional</td>
<td>996</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Mean r values for each of the three key relationships. “W-Mean” is transformed and weighted according to the sample in each study. “Meta n” refers to the n used for statistical conversion - see Methods and Results Synthesis.

<table>
<thead>
<tr>
<th>Association</th>
<th>Meta n</th>
<th>Mean r</th>
<th>W-Mean r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loot Boxes and problem gambling</td>
<td>19,334</td>
<td>0.25</td>
<td>0.27</td>
</tr>
<tr>
<td>Loot boxes and problem video gaming</td>
<td>3433</td>
<td>0.26</td>
<td>0.40</td>
</tr>
<tr>
<td>Problem gambling and problem video gaming</td>
<td>23794</td>
<td>0.21</td>
<td>0.21</td>
</tr>
</tbody>
</table>
References


Johnson MR and Brock T (2019) What do Game Developers Think about Loot Boxes?


Zendle D (2019a) Gambling-like video game practices: Links with problem gambling and disordered gaming in a nationally representative sample. Available at: https://doi.org/10.31234/osf.io/fh3vx MLA.

Zendle D (2019b) Links between problem gambling and spending on booster packs in collectible card games: A conceptual replication of research on loot boxes.

Zendle D (2019c) Problem gamblers spend less money when loot boxes are removed from a game: a before and after study of Heroes of the Storm. *PeerJ* 7: e7700.


Zendle D, Meyer R, Waters S, et al. (2019) The prevalence of loot boxes means that content descriptors are just as important as limit setting when it comes to minimising potential harm: A reply to Drummond et al. *Unpublished working paper*.