Please cite the Published Version

Are videogames represented as violent in the press? Hypothesis testing using MD-CADS and corpora from UK newspapers in 2000 and 2020

Abstract

Previous academic research has argued that videogames are represented as violent in the media, and that these representations provide a justification to study videogames as a social object. However, little systematic research has been done to confirm whether such representations occur frequently, or whether this is a non-representative assumption based on a small number of news stories. Through using methods associated with Modern-Diachronic Corpus-Assisted Discourse Studies (MD-CADS), we explore two corpora of newspaper articles from the British press containing references to videogames: one contains data from 2000 and the other from 2020. We find that while newspapers do represent videogames as promoting violence, these representations occur much more frequently in 2000 than in 2020, and that in 2020 this idea is more regularly challenged. Furthermore, in 2020 there is a greater focus on the benefit of videogames for the economy compared to 2000. The research highlights the need to critically revaluate assumptions about how the media represents not only videogames, but also comparable topics.

1. Introduction

Speaking from a criminological perspective, Kelly et al. (2020) provide several examples where media outlets and politicians have correlated the violence in (some) videogames with the actions committed by violent offenders who play such videogames. The authors go on to reject such a correlation – arguing that the relationship between violence in videogames and the actions of violent offenders is far more complex than the media portrays. Similarly, Hart (2020) concurs with this idea and has conducted a close analysis of the media discussion around school shootings, such as the Columbine shooting in 1999. Hart argues that the media create moral panics around topics like videogames, pornography, and action movies, and that they place the blame for violent actions on violent media (see also Partington, 2010 for a mention of this with regards to the murder of James Bulger). Stemming from sociology, the term 'moral panic' refers to a suspected rise in activities perceived as immoral (Hunt, 1997), usually pejoratively attached to a certain social group, with the anticipated outcome of the immoral activities being the social breakdown of society (Thompson, 2013). They are typically seen as short term, socially constructed events (Goode & Ben-Yahuda 1994). Social concerns, on the other hand, have a less pejorative use, and have been defined by Agnew (2013) as a moral inclination to elevate the issues of others beyond the self. Throughout this paper, in line with Partington's (2010) findings of what is labelled a 'moral panic', we opt to

use the term 'social concern', as it is somewhat more reflective of issues from a diachronic perspective and for analysing objects (e.g., videogames) as opposed to groups.

Studies such as those conducted by Kelly et al. (2020) and Hart (2020) are somewhat exemplary of research into how the media constructs videogames: typically, these kinds of research papers argue that the media constructs videogames as violent, even if there are several non-violent videogames (see also, for example, Ferguson, 2008). However, a lot of previous research into the relationship between the media and violent videogames is predicated on the idea that there are social concerns around videogames – and that the media represents videogames in such a way. However, beyond anecdotal evidence, or somewhat cherry-picked data, this assumption has not yet been tested. This paper seeks to interrogate this underpinning assumption about how the media represents videogames, through analysing representative corpora of media reporting. We aim to provide foundational research to confirm/reject the hypothesis that videogames are represented as the cause of violence within the media (see Partington, 2010 for a discussion of hypothesis-driven corpus and discourse research).

In order to achieve this aim, this paper takes a diachronic corpus linguistic approach to the representation of videogames. Given the limitations of space, we will only focus on two time periods: first, in 2000 and second in 2020. The year 2000 was selected for two reasons: first, it was the year following the Columbine shooting noted by Hart (2020), meaning that 2000 would have been a time when the idea that videogames cause violence would have been a popular topic to report on. This event is what Partington (2010: 102) calls a 'seasonal' event, which is likely framed around its connections to social concerns. Secondly, the year 2000 was chosen because it represents the middle point of what has been called the 'sixth-generation' of computer and videogames, characterised by massive improvements to graphics and developments of more complex narratives in videogames (see Srinivasan & Venkatraman, 2018). During this time period, a number of popular consoles and videogame series were being released for the first time (for example, the PlayStation 2 was released in 2000). We selected 2020 as not only was it two decades after the first corpus, and the most recent years' worth of data available when the research began, but because of the COVID-19 pandemic and ensuing lockdowns. Both Balhara et al. (2020) and King et al. (2020) observe a significant increase in the time spent gaming (predominantly by young men) during this period. To that end, one might expect to see similar social concerns about gaming and the impact of gaming on young men's mental health raised in the British press. These years thus provide interesting points of comparison for different reasons, and evidence for whether there are regular portrayals of concerns about videogames.

This paper is structured as follows: first, we review the background literature on ludolinguistics, focusing particularly on work which has looked at videogame paratext. In addition, we discuss work which has taken a corpus linguistic approach to exploring the representation of social issues in the press from diachronic perspectives. We identify a number of gaps in the current literature with regards to the topic of the representation of videogames and formulate three research questions based on this gap. Following this, we discuss the corpora, how they were gathered and how they were analysed. We then discuss the key findings, focusing on violence and the industry/economy. This leads us to a brief discussion of our findings against the context of previous research. Finally, we provide some

concluding remarks, note the limitations of the research, and suggest directions for future studies.

2. Background literature

2.1. Ludolinguistics

Before we discuss how corpus methods have been applied to ludolinguistics and related fields, particularly from a corpus linguistic perspective, a distinction should be made between two sub-fields within ludolinguistics: studies which examine orthodiscourse and studies which examine paradiscourse (see Carter et al., 2012; Author #1, 2021a; 2022a). Orthodiscourse relates to the language used within videogames and how representations are constructed within videogames as a text (see Author #1 2021a, 2022a). On the other hand, paradiscourse relates to the language around videogames and how representations are constructed in these texts. This includes, for example, how communities and players use language to talk about representations (e.g., Ensslin, 2012; Gray, 2014; Richard & Gray, 2018; Author #1, 2022b). Importantly for the present study, paradiscourse extends to how different forms of media talk about videogames and the interdiscursive nature of multiple texts (e.g., Miller and Summers, 2007; Summers and Miller, 2014). Paradiscourse is not localised to texts produced by the players, but also includes texts from professionals associated with games, everyday observers, and the media to name a few. For example, Miller and Summers (2007) explored how gendered characters were constructed in videogame magazines, a text which was related to videogames but ultimately produced by people external to the game.

Previous research into videogame paradiscourse has utilised corpus linguistic methodologies to explore a range of text types. For example, Potts (2015) used a mixed-methods approach (including corpus methods) to explore how homosociality was constructed in an online community of people who watched 'let's play' videos on YouTube. Potts explored how the community appraised the YouTuber by examining the adjectives and verbs that collocated with the YouTuber's name. In addition to the use of collocational analysis, Potts also utilised explored multiple concordance lines for the collocates to investigate the different ways those words (and word combinations) were used. Similarly, Author #1 (2022b) used concordance line analysis to explore salient phraseology about transphobia across several comments in posts about World of Warcraft's first transgender character. Author #1 explored the language used on the World of Warcraft official fora in a thread about this transgender character and argued that multiple concordance lines could be compared to explore how people with different ideologies interacted with each other and how they framed their support for/against the inclusion of the character.

However, despite an extensive search of the literature, there appears to be a lack of research that explores paradiscourse about games in a general sense. In other words, previous research appears to have looked at how people react to social topics within videogames (such as about how people talk about gender in games), rather than videogames as a social topic. In addition to this, work on (corpus) linguistic approaches to videogame paradiscourse often takes data from a single temporal point. In other words, there are a lack of studies which approach paradiscourse from a diachronic perspective. As others (e.g., Baker, 2001; 2005; Partingon, 2010) have noted, media portrayals of issues which were once social concerns can

change drastically over a short period of time, which may also be the case with media representations of videogames.

2.2. Modern Diachronic Corpus Assisted Discourse Studies (MD-CADS)
Corpus linguistic methodologies have previously been used to explore changes diachronic changes in language and attitudes towards certain topics (e.g., Baker, 2011; Brookes & Wright, 2020; McEnery et al., 2019; Taylor, 2022a). One common data set used to conduct such analysis is newspaper articles (e.g., Baker, 2005; Taylor, 2022a; Wilkinson, 2019). In particular, a number of studies have explored how concepts and people are represented in British newspapers, as these texts are engaged with by millions of people and have the power to both normalise and legitimise ideologies towards such concepts or people (see Jeffries, 2007). In addition, digital archival databases have become so commonplace that these texts are some of the most well documented and accessible forms of media.

Ultimately, as Partington (2010) notes, observing newspaper language through the lens of corpus linguistic methods allows for multiple types of patterns to be examined simultaneously; namely, linguistic, cultural and societal. For example, a report from the Association for UK Interactive Entertainment (UKIE), the trade association for the UK's games and interactive entertainment industry suggests that videogames have evolved from being a relatively niche hobby into a massive industry that is still witnessing record-breaking growth (see Osborn, 2022). By looking at newspaper data through corpus linguistic methods, it may be possible to see if whether reporting on videogames has increased to match the growth of the videogame industry, and whether this reporting is accompanied with changes in the industry. The examination of such changes allows for an exploration of what linguistic, cultural, and societal shifts related to videogames have occurred.

MD-CADS (Modern Diachronic Corpus Assisted Discourse Studies) is a sub-discipline of CADS (Corpus Assisted Discourse Studies), which is particularly concerned with using corpora in the analysis of discourse from different moments of contemporary time. Although there are numerous goals and traditions associated with both CADS and MD-CADS, the tradition most relevant to the current study is hypothesis testing (see, e.g., Taylor, 2010; Duguid, 2010). This can be done in different ways – though this paper takes a more top-down analysis to the data (i.e., exploring pre-selected words, phrases, and constructions for how they are used). By taking a modern-diachronic approach, we may be able to see the development of meaning in how words interact with each other by observing newspaper reporting over time, which allows for different concepts to be drawn together through repeated word combinations (Baker, 2005, 2006, 2014). By using MD-CADS as a basis for our study, we can explore more rudimentary hypotheses of the nature of newspaper reporting of videogame violence to see whether this relationship exists, and if it changes over time.

Finally, it is worth situating the current study against previous MD-CADS papers with regards to the time-period of just 20 years. Taylor (2022b) explored the periods of time that different studies, which were presented at the Corpora and Discourse Studies International Conference 2022, used for their corpora. Taylor grouped the research presented into three categories: short-term change (e.g., around 20-50 years), medium-term change (e.g., around 50-10 years), and long-term change (e.g., 100+ years). Taylor argues that there are several

approaches to diachronic CADS, but that many also neglect long-term change. One reason for this might be due to the accessibility of corpora and corpus-building databases. While the current study would likely be considered to only analyse short-term change (in line with the notion of modern-diachronic analyses), given that videogames are a relatively recent phenomenon (they have only been available as home-consoles since the early 1990s), this study uses data which is somewhat long-term for the availability of the data.

The above literature points to several research gaps. To date, little work has been done to address 1) how videogames have been represented within the media through systematic analysis. And 2) whether the representation of videogames may have changed across time. To contribute towards bridging these gaps, we aim to answer the following research questions:

> 1) How are videogames portrayed in the British press?
>
> 2) Does the portrayal of videogames align with previous assertions that they are represented as violent?
>
> 3) Has there been any change in how videogames are represented across a 20 year period?

3. Data

As is consistent with previous literature which uses corpus linguistic methods to explore press representations (e.g., Wilkinson, 2019; Brookes & Wright, 2020;), the corpora were collected using the online newspaper archive 'Nexis'. The search terms 'videogame' and 'video game' were selected because they were significantly more frequent than their counter parts 'computergame' and 'computer game'. It is also worth noting that the term 'computer game' has become somewhat phased out in favour of the more general videogaming by the community, as videogaming is more inclusive of a wider range of consoles beyond computers, thus possibly not providing a good representation of actual discussion of videogames in the corpus. We also opted not to use the search term 'gaming' as it has an almost opposite effect of 'computer gaming', being more general. 'Gaming' could potentially include other forms of games, such as board games and tabletop role-playing games, which have similarly become mainstream forms of entertainment over the time period represented in the corpora. By using 'videogame' as our primary search term, we achieve the optimal balance between specificity of the topic whilst also being inclusive of all that topic may cover.

The files were initially downloaded as Word documents (owing to the limitations of Nexis) but were converted to .txt files using a module in Word which changed their file extensions from .doc. This module has been included in Appendix 1. After conversion, a number of processing stages were applied. Firstly, using regular expressions (regex), the metadata information at the top of each file were removed. The metadata removed included all the text that appears above the string 'body' (which marks the start of the article) and all the text following 'load-date' (which marks the end of the article). Removal of this additional text

meant that the corpus files only contained the article text, rather than the additional copyright information and metadata which accompanies collection from Nexis.[1]

Secondly, all the text within the corpora was changed to lowercase, to improve the recall of searches. In addition, this step enabled the files to be compared for any duplicates via string matching within Python. At this stage, a total of 742 files were identified as duplicates and removed from the corpora. Another processing stage included standardising different spellings of search terms. For consistency, singular and plural forms were changed from 'video game(s)' to 'videogame(s)' to create a standard form, which allows for easier collocational analysis . This change is not dissimilar from the change implemented in Author #1 (2021b), who changed the search term '[redacted for peer review]' to '[redacted for peer review]' for collocational (and collocational network) analysis. In addition, the related terms 'videogamer(s)/video gamer(s)' and 'videogaming/ video gaming' were also changed to a standard one lexical unit form. The final processing stage involved renaming the original Nexis files from the article titles to a regular format (e.g. '2000s_1.txt'). The new filenames were stored in a spreadsheet alongside the original Nexis output, again, to enable cross-checks with the original Nexis files.

The data were then loaded into #Lancbox (Brezina et al., 2015) to ascertain their sizes. The first corpus – containing data from newspapers published in English in the UK between 1st January 2000 and 31st December 2000 inclusive totalled 1,238,413 words and contained 1,667 articles. We refer to this corpus as the VG2000 corpus. The second corpus, constructed using the same parameters of VG2000 but using data published between 1st January 2020 and 31st December 2020, contains 3,410 newspaper articles consisting of a total of 2,480,279 words. We refer to this second corpus as VG2020.

4. Methods

In order to explore how videogames were written about in both corpora, we explored the use of the singular and plural forms (i.e., 'videogame' and 'videogames'). The term 'videogame' occurred 1942 times in VG2000 (normalised frequency: 1,568 per million words) and 3757 times in VG2020 (normalised frequency: 1,514 per million words). Further, the term 'videogames' occurred 799 times in VG2000 (normalised frequency: 645 per million words) and 1071 times in VG2020 (normalised frequency: 432 per million words). We investigated the top 30 collocates of each team in each corpus. For the purposes of this paper, we used a window of 5 words to the left and 5 words to the right of the search term and used the Log-likelihood statistic to calculate whether a collocated was statistically significant (see Brezina, 2018). The collocates are discussed in Tables 1-4. For data visualisation purposes, we do not include the statistics for each collocate, but all collocates presented in this paper were at least LL = 17.855 (meaning they were statistically significant at $p<0.0001$).

Once we generated the list of the top 30 collocates, we then contrasted each possible variation from both angles (diachronic and plurality) to explore similarities and differences. Importantly, we 'look in both directions' (Taylor, 2013) and acknowledge that both similarities and differences are important in how such phenomena are constructed in the

---

[1] All the original formatted Nexis files were retained in a separate location, in case of need of consultation.

media. We noted a number of semantic themes within the collocates, but given the limitations of space, we pay particular attention to the themes of violence and industry. To do this, we explored how the collocates were used in context, by examining the concordance lines for such collocation.

5. Analysis

Once we generated the lists of the top 30 collocates, we contrasted these across the corpora. Shared and unique collocates for both search terms are highlighted in Tables 1 and 2.

Table 1. Collocates of 'videogame' (singular) in VG2000 and VG2020

Corpora	Number of shared/unique collocates	Collocates (Content words)	Collocates (Function words)
VG2000 and VG2020	22	based, playing,	the, a, which, has, that, like, on, in, for, by, and, with, of, is, to, from, it, was, as, an
VG2000	8	market, new, playstation, sega, computer, nintendo	or, be
VG2020	8	popular, character, fortnite, fifa, world, series, developer, online	

Table 2. Collocates of 'videogames' (plural) in VG2000 and VG2020

Corpora	Number of shared/unique collocates	Collocates (Content words)	Collocates (Function words)
VG2000 and VG2020	20	children, violent, playing, industry,	the, a, has, that, on, in, for, with, and, of, are, is, to, have, as, more
VG2000	10	market, inner, sega, computer, nintendo	by, be, but, they, it
VG2020	10	people, played, time, said, play	or, who, about, from, at

      While these collocate lists demonstrate some differences in collocates across time for both singular and plural forms, we recognise that this kind of visualisation is more suited to demonstrating diachronic changes, as opposed to changes caused by the plurality of the noun.

As such, to more easily demonstrate the change that the use of singular/plural forms has on collocates, we demonstrate collocates unique to the singular and plural forms in Tables 3 and 4.

Table 3. Collocates of 'videogame' and 'videogames' in VG2000

Corpora	Number of shared/unique collocates	Collocates (Content words)	Collocates (Function words)
VG2000 - plural VG2000 – singular	21	market, sega, playing, computer, nintendo	the, a, has, that, on, in, for, by, be, with, and, of, is, to, it, as
VG2000 – singular	9	new, based, playstation,	which, like, or, from, was, an
VG2000 – plural	9	children, inner, violent, industry,	but, they, are, have, more

Table 4. Collocates of 'videogame' and 'videogames' in VG2020

Names	Total	Collocates (Content words)	Collocates (Function words)
VG2020 - plural VG2020 – singular	15	playing,	the, a, has, that, on, in, for, and, with, of, is, to, from, as
VG2020 – singular	15	popular, based, fortnite, character, fifa, world, series, developer, online	which, like, by, it, was, an
VG2020 – plural	15	children, violent, about, people, played, time, said, play, industry	or, who, are, have, at, more

The above collocates demonstrate that videogames are represented in complex ways, depending on whether they are referred to in the singular or plural form. In addition, the above tables of collocates demonstrate that there has been diachronic change in how videogames are represented. The data also suggests that there are some similarities and differences in the kind of semantic domains of collocates, such as how terms like 'market' occur in both 2000 and 2020, and how 'industry' occurs as a collocate of the plural form in 2020. In the next two subsections, we explore two of these groups of collocates more closely.

5.1. Videogames and violence

Tables 1-4 demonstrate that videogames are typically more likely to be discussed as 'violent' when referred to in the plural form (this collocation occurs 40 times in VG2000 and 17 times in VG2020). This could be an indication that the British media are less likely to name

specific violent videogames and that this has reduced in frequency over time. The reference to several potentially violent videogames might be one way in which a social concern is constructed: as there being several potential threats and/or harms that the reader has little control over, rather than a single issue that can be addressed (Victor, 1998). In addition to the above, terms denoting children tend to occur with references to videogames as a plural noun in both VG2000 and VG2020. Taken together, these collocates could indicate that one way the media constructs videogames is as a social concern for the safety of children. This would not be surprising, given that social concerns are usually centred around a 'save the children mentality' (see Markey & Ferguson, 2017). In order to test whether the media constructed violent videogames as a risk to children (i.e., whether these collocates were related in some way), a concordance line analysis was conducted for all instances of the collocation between 'violent' and 'children'. However, given that there is a 5L 5R window from 'videogame' as the node word, this could mean that 'violent' and 'children' could occur within a 10L;10R window of each other. As such, parameters were increased to account for this difference and all concordance lines were explored with this increased window-range. This collocation occurred 41 times in VG2000 (LL = 59.906) but was not a collocate in VG2020. The concordance line analysis demonstrated that these kind of constructions were indeed used in VG2000, such as:

(1) there is much controversy over the impact of violent videogames on children (VG2000)
(2) the US senate last week approved an inquiry into the impact of violent videogames on children (VG2000)

In the 41 instances of collocation between 'violent' and 'children' in VG2000 (in a 10L;10R window), 37 were specifically about violent videogames. Of these 37 occurrences, 35 related to social concerns about children being exposed to violence in videogames. This conflation between violent media and violent actions, especially as with regards to children's behaviour, draws on notions of cultivation (see Morgan et al., 2017). While cultivation was designed for television media, the evidence as to whether videogames have a cultivation effect is somewhat mixed (see Author #1, 2021a: 7-9). It appears as though media narratives on videogames in VG2000 purported ideas similar to cultivation to the general British public, without much critical engagement with whether such notions were correct.

One linguistic element of these instances we noted was the agency of children. In both examples provided above, children have a low degree of semantic agency – and are often constructed as having an influence 'on' them, rather than, for example, 'engaging' with violence (see Darics and Koller, 2019 for a discussion of semantic agency). There are different ways to approach agency and action – but for this paper we are particularly concerned with semantic agency (who does what and who has what done upon them). Indeed, our closer analysis of the concordance lines revealed that of the 35 instances where children were written about in relation to social concerns about violence in games, 32 positioned the children as having actions done upon them/being exposed to the violence in videogames with little to no semantic agency.

A closer analysis of the collocation between 'violent' and 'videogames' revealed that the British press typically reported on this concept as evidenced by professionals (e.g., academics) and people in positions of power (e.g., politicians), as also demonstrated in

example (2). This kind of representation occurred in 30 instances where 'violent' and 'videogame' collocated, and these are further shown in the following examples:

> (3) Craig Anderson, of Iowa State University, the study's leading author, said: "we found that students who reported playing more violent videogames in junior and high school engaged in more aggressive behaviour." (VG2000)

> (4) A judge has banned six violent videogames throughout Brazil, saying they inspire real violence. Judge Claudia Maria Resende Neves Guimaraes ordered shops to remove the videogames Doom, Postal, Mortal Kombat, Requiem, Blood and Duke Nukem from their shelves or face £6,250 fines (VG2000)

While these examples demonstrate a prominent representation of videogames, the only time when counter-discourses appeared within the reporting of violence in videogames was when comments and letters to the editor were published. There were 5 instances where letters to the editor countered these arguments, and these typically came from other experts explaining that more research was needed, or that the evidence for violent videogames creating violent 'real world' actions was unclear. This can be seen in the following example:

> (5) your article about violent videogames […] raises interesting questions about the effects of videogame violence on social behaviour. […] there has been little systematic research […] there are also problems defining "violent" or "aggressive", […] therefore, research into the effects of long-term exposure remains speculative. Dr Mark Griffiths, Nottingham Trent University

The inclusion of letters to the editors like this are important for representations of videogames: they demonstrate that press representations of videogames are not monolithic. However, in line with notions associated with MD-CADS, considering the context of these letters is also important. Letters to the editor typically come after sensationalised news stories, and often are glossed over because they are given less prominence than commissioned articles. This means that readers may see these counter-discourses at a lower frequency in comparison to discourses found in the main body of articles. Although the inclusion of such discourses in letters to the editor are important, because they represent a range of ideologies and editors can chose to in/exclude certain discourses. In terms of corpus linguistic methodologies, such an approach also demonstrates the need to retain meta-information in corpus construction, as this kind of contextual information (i.e., about the file being from the letters to the editor section of a paper) was removed at earlier stages and only confirmed by cross referencing the corpus to original files.

By contrast, the way videogames are typically written about is different in VG2020. There were 27 instances of collocation between 'videogames' and 'children'. The concordance lines for this collocation revealed that there was only one instance where these terms occurred with "violence". In this one instance, the author refuted the conflation of violent videogames making children violent:

> (6) playing violent videogames such as grand theft auto and call of duty won't make children more aggressive (VG2020)

Indeed, an exploration of the collocation between 'videogames' and 'violent' (of which there were 17 instance) demonstrated that the majority of concordance lines (11 instances) rejected

the hypothesis that violent videogames make people (especially children) violent. This is evidenced in the below examples:

> (7) videogames do not lead to violence or aggression, according to a reanalysis of data gathered from more than 21,000 young people around the world. (VG2020)

> (8) more recent research also revealed that playing violent videogames such as grand theft auto and call of duty won't make children more aggressive. (VG2020)

These kinds of representations demonstrate a shift in how videogames are represented: while they were previously positioned as the cause for social concerns (i.e., that violence in videogames would cause children to become violent), more recent reporting of academic research appears to be refuting this idea. This kind of analysis also demonstrates the need to look deeper into perceived similarities in collocates (see Taylor, 2013). Despite the term 'videogames' sharing similar collocates in both VG2000 and VG2020, the concordance lines for this collocation demonstrate different representations, which appear to have changed to a more nuanced and neutral stance in 2020 compared to 2000.

Finally, within our analysis of violence and videogames, we were concerned that data from VG2000 and VG2020 would differ because Columbine and COVID-19 were very different seasonal events. The former has been discussed particularly with regards to gun violence (see Hart, 2020), while such violence did not necessarily occur during the 2020 COVID-19 lockdowns. To explore whether the violent nature of Columbine as a seasonal event impacted the reporting of videogames as a social concern, we also built a third corpus of news reporting on videogames following a violent shooting in the UK. In August 2021, Jake Davidson shot and killed 6 people (including himself) and wounded two others. Davidson was later associated with online hate groups (particularly the incel community) (see Author#1 forthcoming). Some newspaper articles which reported on the event suggested that he was also a fan of first-person shooter videogames and had previously posted some gameplay footage of the videogame *Killzone* prior to his attack. For example, Metro note that Davidson's: 'Youtube cover image featured a still from a video game showing a character wielding an automatic weapon' (Seddon, 2021). In these reports, the relationship between violence and videogames appears to be implicit rather than made explicit through collocation. However, such cherry-picked examples may not be representative of broader reporting, and it could be the case that videogames are still represented as violent after a seasonal event involving gun violence from people who play such games. If the media were to have the same social concerns about videogame violence, it would be expected that reporting from the month following Davidson's attack would also contain such concerns. A third supplementary corpus was built using the same parameters as VG2000 and VG2020 but was only collected for the month following Davidson's shooting (12th August-12th September 2022 inclusive) and totalled 33,791 words.

Comparison to the supporting corpus showed that no terms denoting 'Jake', 'Davidson', 'Plymouth', 'incel', or 'attack' appeared in the corpus. There were two instances where 'shooting' occurred, both in the context of a function of game mechanics (e.g., '"hone your skills," which adds dribbling, shooting, dunks and alley-oops') rather than referencing gun-violence. There were only 9 statistically significant collocates for 'videogames' in the supplementary corpus, which are outlined below in Table 5. Given the much smaller size of the supplementary corpus, a minimum frequency of 1 was used for collocation.

Table 5. Collocates of "videogames" in the supplementary corpus

Collocate	Rank	Freq(Scaled)	Frequency	Likelihood
announced	1	170	4	26.591
superheroes	2	50	2	15.393
homes	3	60	2	14.657
clearly	3	60	2	14.657
education	5	80	2	13.502
wars	6	90	2	13.031
marvel	7	100	2	12.612
leave	7	100	2	12.612
star	9	110	2	12.233

As the data from Table 5 suggests, there were no collocates which explicitly discuss violence or the impact of violent videogames on children. The one collocate which might indicate this is *wars*, which is only ever used to talk about the franchise *Star Wars*. Combined with other collocates, this shows that when the media were talking about videogames in the month following the Jake Davidson shooting, they were typically focused on videogames around superheroes and fantasy (e.g., 'marvel' heroes). In turn, and in line with other work in MD-CADS, it is also important to consider what does *not* appear in collocational lists or within a given corpus (see, e.g., Jaworska, 2016). Similarly, a collocational analysis of 'videogame' revealed that no terms denoting violence occurred, but the second most frequent collocate was "industry" (LL = 35.659).

5.2. The videogame industry and brands

The previous section has problematised the hypothesis that the media constructs a social concern around videogames through linking violence in them with offline violence. The findings above demonstrated that while there are some references to violence in games in the media in 2000 and 2020, these are not the only representations. Another common way the British press write about videogames is in relation to the videogame industry/the economic benefit of the videogame industry. This is not surprising given how lucrative the videogame industry is – a report by Accenture in 2021 forecasted that the global value of the industry would be $300 billion in 2022 (Accenture, 2021). In Tables 1-5, several collocates of 'videogame*' indicate references to: specific videogame companies, specific videogame titles, specific videogame products (e.g., systems), or the industry more broadly. Taking the data from VG2000 and VG2020 presented in Tables 1-4, this section turns to explore how videogames are constructed in relation to the videogame industry more broadly. These collocates from Tables 1-4 have been reproduced and categorised by the groups noted above in Table 6.

Table 6. Collocates related the videogame industry

Category	Collocates unique to VG2000	Collocates unique to VG2020	Shared collocates
Specific videogame companies	Sega, Nintendo		
Specific videogame titles		Fortnite, Fifa	
Specific videogame products	PlayStation, computer		
Videogame industry broadly	Market	Developer	Industry

The categorisation of collocates in Table 6 demonstrates some shift in how the videogame industry is referenced in the British press. The data suggests that specific videogame companies were more likely to occur with the 'videogame(s)' in 2000 in comparison to 2020, as were specific videogame products. This is not to say that the names of these companies or products did not occur within the corpus, rather they are now less likely to be referred to in tandem with the 'videogame(s)'. A concordance line analysis was conducted on the collocates presented in Table 6, to explore the contexts in which they were used with/without videogames. For example, articles in VG2020 referred to:

> (9) 2. Fortnite (PlayStation4, Xbox One, Nintendo switch, Pc/Mac and IOS/Android […]) when life is as dark as it is right now, it's important to remember the importance of "fun" […] although a competitive shooting game at heart, Fortnite distances itself from similar games through its cartoon-like aesthetic (VG2020)

> (10) After a raft of dire adaptations, gamers weren't jumping for joy over the prospect of a movie based on Sega's coincollecting hedgehog. (VG2020)

Example (9) refers to the online videogame Fortnite, developed by Epic Games. Example (10) refers to a movie based on Sega's Sonic the Hedgehog series. In these examples, it is somewhat interesting to see how more the more recent game (Fortnite) does not have its developer referred to. This might suggest less emphasis is placed on who develops the game and more on the game itself. However, these examples could also demonstrate how popular and well-known various videogames have become. By comparison to data found in VG2000, there appears to be less of a need to note that Sonic the Hedgehog is a videogame in 2020, likely due to the continued success of the videogame since it was released in 1991.

By contrast, when the names of videogame companies collocated with the 'videogame(s)' in VG2000, the typical construction they appeared with was [possessive form] + 'videogame(s)'. For example:

> (11) The real excitement this Christmas will be over Pokemon, Nintendo's videogame cartoon character for which Hasbro makes spin-off figures. (VG2000)

Further, when specific products occurred with 'videogame(s)', this was used in a noun phrase, such as

> (12) because of high demand, Sony has had to delay the sale of PlayStation 2 videogame consoles in Europe until November 24 (VG2000).

However, such constructions do not appear in VG2020. While noun phrases like "videogame consoles" are used, they do not appear with the names of videogame companies. The shift from noting the videogame as being owned by a company or using noun phrases such as "videogame consoles" could potentially have been due to how novel they were in 2000. However, given the extensive growth of the videogame industry, and the fact that most videogame companies are now household names, editors in the British press may now view such references as redundant. From an MD-CADS perspective, understanding this shift is particularly important because it can highlight changes in social attitudes towards videogames as well as videogame companies, and might demonstrate how such companies are becoming more well-known and accepted.

Finally, it is worth noting how the British press refer to the videogame industry more broadly. In 2000, the British press typically wrote about videogames so as they would collocate with 'market' as opposed to 'industry'. However, both terms were used to primarily discuss how various videogame companies were doing within the market, for example:

(13) Sony computer entertainment, is to launch a portable version of its older 32-bit PlayStation console in a bid to remain ahead of the pack in the lucrative videogame market (VG2000)

(14) Electronic Arts' European PR head, David Willis, said: "we think we have got a tremendous story to tell: we have broken the billion dollar barrier for the first time in our industry"

By contrast, the term 'market' was less statistically likely to occur with 'videogame(s)' in 2020[2]. When it did occur, videogame markets were spoken about less in terms of how specific videogame companies were doing on the market, such as:

(15) there is a buoyant vintage consoles and games market, and all the paraphernalia that go with it (VG2020).

This could demonstrate a shift how the media view videogame companies: when videogames were more fringe, there may have been concerns around the survivability of companies. Such concerns may have dissipated by 2020, given the massive increase in the market for videogames and the clear economic benefit of the industry.

However, as noted previously, the papers in VG2020 were more likely to refer to the videogame 'industry'. Unlike the use of 'market', 'industry' was statistically like to collocated with both 'videogame' and 'videogame(s)' (although, the term appears below the arbitrary top 30 cut-off for the singular form in 2020). In the concordance lines for the collocation between "videogames" and "industry", the role of the UK in the international videogame industry was the most commonly referenced topic, such as:

(16) "the UK has been at the forefront of the videogames industry for decades," said Royal Mail's Philip Parker. (VG2020)

---

[2] Note, the collocation between 'videogame' and 'market' is still statistically significant, but is below the arbitrary top 30 cut-off point established for this paper. 'Market' is the 349th most statistically significant collocate of 'videogame' (LL = 62.2).

Furthermore, the videogame industry is also referenced as a collectivised social actor who puts pressure on the government for policy changes, such as:

> (18) Scotland's videogames industry has called for government action to maintain growth in the sector as new figures show it is worth more than £236 million to the UK economy. (VG2020)

This kind of representation suggests a change over time to a more focused and localised conceptualisation of the videogame industry. This way of viewing the videogame industry appears to have moved from one which conceptualises individual companies as competing to survive to one where multiple companies support the livelihoods of workers and have an influence on national politics.

The change in how the videogame industry is viewed might also account for the shift in how violence in videogames is constructed in the British press. In an earlier subsection, we demonstrated a move away from the construction of social concerns around the violence in games. Given that there also appears to have been a shift in the normalisation of the videogame industry and a greater focus on the benefits of gaming to the economy, if the press were to also report on the social concerns they had for children's safety, this could be constructed as contradictory news values (see Bednarek and Caple, 2017 for a full theorisation of news values).

6. Discussion, conclusions, and directions for future research

This research was conducted to address the presupposition that the media represents videogames as violent. To explore this, we have conducted an MD-CADS analysis of two representative samples of language from the British press: one from 2000 and one from 2020, drawing on collocational and concordance line analysis. We have demonstrated that the way the British press has represented videogames is much more complex than simply only representing videogames as the cause of violence, and the way they have been represented has changed between these years. The data suggests that in 2000 videogames were constructed as a social concern because of the potential risk they had of normalising violence to children, but this was not the only way they were represented. Additionally, claims of violence in videogames causing offline violence appear to be more contested in 2020. Compared to, for example, a close reading of two or three newspaper articles, the use of corpus linguistic methodologies has allowed us to explore social attitudes constructed in the press in a more generalisable sense. The findings presented here could only be gained through the use of corpus linguistic methods as our research questions were particularly concerned with exploring representative data from the media to explore whether general claims made about the press were accurate (or not).

The research has also demonstrated that there are differences in how videogames are represented depending on if they are referred to in the singular or plural, with singular forms typically relating to the videogame industry and plural relating to overarching trends in videogames, such as violence. Importantly, the analysis of singular and plural demonstrates that no one videogame, or a small number of videogames referred to in the singular across several instances, is represented as the root of social concerns or representations. In other words, there do not appear to be constructions such as 'the violent videogame "x" is causing

children to become violent'. Similarly, the use of corpus linguistic methodologies has allowed us to tentatively suggest a need to reframe interdisciplinary arguments around videogames: that there are differences in how individual videogames and specific videogames are represented, and this could lead to future research on the representation of specific games, accompanied by hypotheses to be tested through corpus linguistic methodologies.

The data also demonstrates that it is not just violence that the media draws attention to. The data suggests that several outlets also discussed the videogame industry, the money videogames can generate, and how companies are doing in terms of profits. This appears to have been a consistent theme in both the data from 2000 and 2020, although the focus of the economic impact of videogames has changed over time. While some previous work has focused on the representation of violence, it is also important to remember and acknowledge that the representation of videogames is not monolithic, and that they can (and have) also been represented in positive ways.

The paper here also has demonstrated the need to consider the seasonal nature of some news reports, which is particularly important for other research in CADS. Our implementation of a supplementary corpus was useful because it demonstrated that impressionistic findings were not due to the seasonality of the social events, but rather were accurate reflection of media reporting on videogames more broadly. Although this supplementary corpus was only used for cross-comparison, it increased the reliability of the overarching findings.

However, this paper is not without limitations. While we have explored the representation of videogames with relation to violence and industry, there are several other ways that videogames might be represented – for example, it might be interesting to explore how they are discussed in relation to education, or government policy more broadly. Some collocates and concordance lines suggested that these representations might be present, but given the limitations of space, we have focused on the two most prominent themes. In addition, we have only focused our comparison on two points in time. While a third period was used as a supplementary dataset to confirm our analysis, it would be interesting and useful to more fully explore how such representations have shifted on a finer-grained level through an analysis of each year 2000-2020 inclusive.

Finally, we have only implemented collocational and concordance line analysis. There are several other corpus linguistic and discursive frameworks which might reveal different and nuanced representations of videogames, such as conducting a usage fluctuation analysis (see McEnery et al., 2019), investigating the semantic categories of texts where videogames occur (see Rayson, 2008), or exploring the transitivity (including process types) of videogame companies (see Berber Sardinha and Barbara, 2009). Nevertheless, the methodologies employed in this study have allowed us to test a pervasive hypothesis about videogames and the media, and we are able to offer a tentative conclusion that videogames are discussed in relation to more concepts than just violence, and that the way they have been represented has changed between 2000 and 2020. This change is certainly worthy of more academic inquiry and consideration, and more methodologies could be used to investigate such a change.

Videogames are a popular form of media, and studies which explore the relationship between the media, violence, and videogames are only going to continue to be produced. However, it is also important to test underlying assumptions about media representations before conducting such analyses or making claims about how the media represents something. Our findings also suggest that other fields and studies looking at the relationship between the media and different phenomena may wish to undertake more critical analyses of how newspapers represent those topics and test their own foundational hypotheses.

References:

Author #1. 2021a.

Author #1. 2021b.

Author #1. 2022a.

Author #1. 2022b.

Accenture. 2021. A new wave of social-seeking gamers is driving industry growth to the next level. Retrieved April, 2023 from https://www.accenture.com/us-en/insights/software-platforms/gaming-the-next-super-platform?c=acn_glb_thenewgamingexpbusinesswire_12160747&n=mrl_0421

Agnew, R. 2013. Social Concern and crime: Moving beyond the assumption of simple self-interest. *Criminology*, *52*(1), 1–32.

Baker, P. 2001. Moral panic and alternative identity construction in Usenet. *Journal of Computer-Mediated Communication*, *7*(1).

Baker, P. 2005. *Public discourses of gay men*. Routledge.

Baker, P. 2006. *Using corpora in discourse analysis*. Bloomsbury

Baker, P. 2011. Times may change, but we will always have money: Diachronic variation in recent British English. *Journal of English Linguistics*, *39*(1), pp 65-88.

Baker, P. 2014. *Using corpora to analyze gender*. Bloomsbury

Balhara, Y. P. S., Kattula, D., Singh, S., Chukkali, S., and Bhargava, R. 2020. Impact of lockdown following COVID-19 on the gaming behavior of college students. *Indian Journal of Public Health, 64(*6), pp 172-176.

Bednarek, M., and Caple, H. 2017. *The discourse of news values: How news organizations create newsworthiness*. Oxford University Press.

Berber Sardinha, T., and Barbara, L. 2009. Corpus linguistics. In F. Bargiela-Chiappini (ed.) *The handbook of business discourse*, pp 105-118. Edinburgh University Press.

Brezina, V. 2018. *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.

Brezina, V., McEnery, T., and Wattam, S. 2015. Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics, 20*(2), 139-173.

Brookes, G. and Wright, D. 2020. From burden to threat: A diachronic study of language ideology and migrant representation in the British press. In P. Rautionaho, A. Nurmi , & J. Klemola (eds.), *Corpora and the Changing Society: Studies in the Evolution of English*, pp 113–140. John Benjamins.

Carter, M., Gibbs, M. and Harrop, M., 2012, May. Metagames, paragames and orthogames: A new vocabulary. In *Proceedings of the international conference on the foundations of digital games*, pp 11-17.

Darics, E., and Koller, V. 2019. Social actors "to go": An analytical toolkit to explore agency in business discourse and communication. *Business and Professional Communication Quarterly, 82*(2), pp 214-238.

Duguid, A. 2010. 'Investigating anti and some reflections on Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS)', *Corpora 5* (2), pp 191-220.

Ensslin, A. 2012. *The language of gaming*. Bloomsbury.

Ferguson, C. J. 2008. The school shooting/violent video game link: Causal relationship or moral panic?. *Journal of Investigative Psychology and Offender Profiling*, *5*(1-2), pp 25-37.

Goode, E., and Ben-Yehuda, N. 1994. Moral panics: Culture, politics, and Social Construction. *Annual Review of Sociology*, *20*(1), pp.149–171.

Gray, K.L. 2014. *Race, Gender, & Deviance in Xbox Live: Theoretical Perspectives from the Virtual Margins*. Routledge.

Hart, M. 2020. Social scientists as the architects of their own defeat in the study of video games. In C. Kelly, A. Lynes, and K. Hoffin (eds.) *Video games, crime and next-gen deviance,* pp 23-48. Emerald Points.

Hunt, A. 1997. "Moral Panic" and Moral Language in the Media. *The British Journal of Sociology*, *48*(4), 629–648. https://doi.org/10.2307/591600

Jaworska, S. 2016. Using a corpus-assisted discourse studies (CADS) approach to investigate constructions of identities in media reporting surrounding mega sport events: the case of the London Olympics 2012. In: I.R Lamond & L. Platt (eds.) *Critical Events Studies: Approaches to Research. Leisure Studies in a Global Era,* pp 149-174. Palgrave Macmillan

Jeffries, L. 2007. *Textual construction of the female body: A critical discourse approach*. Palgrave Macmillan.

Kelly, C., Lynes, A., and Hoffin, K. 2020. Introduction: Reorienting the debate. In C. Kelly, A. Lynes, and K. Hoffin (eds.) *Video games, crime and next-gen deviance,* pp 1-22. Emerald Points.

King, D. L., Delfabbro, P. H., Billieux, J., and Potenza, M. N. 2020. Problematic online gaming and the COVID-19 pandemic. *Journal of Behavioral Addictions*, *9*(2), pp 184-186.

Markey, P. M., & Ferguson, C. J. 2017. *Moral combat: Why the war on violent video games is wrong*. BenBella Books.

McEnery, T., Brezina, V., and Baker, H. 2019. Usage Fluctuation Analysis: A new way of analysing shifts in historical discourse. *International Journal of Corpus Linguistics*, *24*(4), pp 413-444.

Miller, M. K., and Summers, A. 2007. Gender differences in video game characters' roles, appearances, and attire as portrayed in video game magazines. *Sex roles*, *57*(9), pp 733-742.

Morgan, M., Shanahan, J., and Signorielli, N. 2017. Cultivation theory: Idea, topical fields, and methodology. In P. Rössler, C. Hoffner, & L. Zoonen (eds.), *The International encyclopedia of media effects*. Wiley.

Osborn, G. 2022. UK Consumer Games Market Valuation 2021 – how much money did people in the UK spend on video games? *UKIE*. Retrieved April, 2023: https://ukie.org.uk/consumer-games-market-valuation-2021

Partington, A. 2010. Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: An overview of the project. *Corpora 5* (2), pp 83–108

Potts, A. 2015. 'LOVE YOU GUYS (NO HOMO)' How gamers and fans play with sexuality, gender, and Minecraft on YouTube. *Critical Discourse Studies*, *12*(2), pp 163-186.

Rayson, P. 2008. From key words to key semantic domains. *International journal of corpus linguistics*, *13*(4), pp 519-549.

Richard, G.T. and Gray, K.L. 2018. Gendered play, racialized reality: Black cyberfeminism, inclusive communities of practice and the intersections of learning in gaming. *Frontiers: A Journal of Women's Studies, 39*(1), pp 112-148.

Seddon, S. 2021. Plymouth gunman was body-building virgin who said he was too ugly to have sex. *Metro*. Retrieved April, 2023: https://metro.co.uk/2021/08/13/plymouth-shooting-jake-davidsons-youtube-videos-detail-sick-views-15086125/

Srinivasan, A., and Venkatraman, V. N. 2018. Architectural convergence and platform evolution: Empirical test of complementor moves in videogames. *IEEE Transactions on Engineering Management*, *67*(2), pp 266-282.

Summers, A., and Miller, M. K. 2014. From damsels in distress to sexy superheroes: How the portrayal of sexism in video game magazines has changed in the last twenty years. *Feminist Media Studies*, *14*(6), pp. 1028-1040.

Taylor, C. 2010. Science in the news: a diachronic perspective. *Corpora*, *5*(2), pp 221-250.

Taylor, C. 2013. Searching for similarity using corpus-assisted discourse studies. *Corpora*, *8*(1), pp 81-113.

Taylor, C. 2022a. The affordances of metaphor for diachronic corpora & discourse analysis: water metaphors and migration. *International Journal of Corpus Linguistics 27*(4), pp 451-479.

Taylor, C. 2022b. The long-distance view: Analysing discourses over time. Paper presented at *6th Corpora & Discourse International Conference*, University of Bologna, Italy.

Thompson, K. 2013. *Moral panics*. Taylor and Francis.

Victor, J. S. 1998. Moral panics and the social construction of deviant behavior: A theory and application to the case of ritual child abuse. *Sociological Perspectives*, *41*(3), pp 541-565.

Wilkinson, M. 2019. 'Bisexual oysters': A diachronic corpus-based critical discourse analysis of bisexual representation in The Times between 1957 and 2017. *Discourse & Communication*, *13*(2), pp 249-267.

Appendix 1:

```
Sub ConvertDocuments()

Application.ScreenUpdating = False

Dim strFolder As String, strFile As String, strDocNm As String, wdDoc As Document

strDocNm = ActiveDocument.FullName

strFolder = GetFolder

If strFolder = "" Then Exit Sub

strFile = Dir(strFolder & "\*.doc")

While strFile <> ""

  If strFolder & "\" & strFile <> strDocNm Then

    Set wdDoc = Documents.Open(FileName:=strFolder & "\" & strFile,
AddToRecentFiles:=False, Visible:=False)

    With wdDoc

      .SaveAs FileName:=strFolder & "\" & Split(strFile, ".doc")(0) & ".txt",
FileFormat:=wdFormatText, AddToRecentFiles:=False

      .Close SaveChanges:=True

    End With

  End If

  strFile = Dir()

Wend

Set wdDoc = Nothing

Application.ScreenUpdating = True

End Sub
```

```
Function GetFolder() As String

Dim oFolder As Object

GetFolder = ""

Set oFolder = CreateObject("Shell.Application").BrowseForFolder(0, "Choose a folder", 0)

If (Not oFolder Is Nothing) Then GetFolder = oFolder.Items.Item.Path

Set oFolder = Nothing

End Function


Function GetFolder() As String

Dim oFolder As Object

GetFolder = ""

Set oFolder = CreateObject("Shell.Application").BrowseForFolder(0, "Choose a folder", 0)

If (Not oFolder Is Nothing) Then GetFolder = oFolder.Items.Item.Path

Set oFolder = Nothing

End Function
```