Department of Media and Communications

Course Code	
Your 5 digit Candidate Number (from LFY) NOT your 9 digit student number	
Word Count:	10, 884
Supervisor's Name:	Professor Shakuntala Banaji
MSc Programme:	Media and Communications(Data and Society)
Date of submission:	26/08/2020

DISSERTATIONSUBMISSION and PLAGIARISM DECLARATION 2019–20

Refer to the LSE Calendar and the MSc Handbook.

"All work for classes and seminars as well as scripts (which include, for example, essays, dissertations and any other work, including computer programmes) must be the student's own work. Quotations must be placed properly within quotation marks or indented and must be cited fully. All paraphrased material must be acknowledged. Infringing this requirement, whether deliberately or not, or passing off the work of others as the work of the student, whether deliberately or not, is plagiarism."

Submitting the same piece of work for assessment twice may be regarded as an offence of 'self–plagiarism'.

DECLARATION(without signature, to preserve anonymity): I hereby confirm, by completing this form, that the work submitted is my own (or part of a joint submission where appropriate). I confirm that I have read and understood both the LSE policy on assessment offences and the relevant parts of the MSc Handbook.

Start your dissertation and your word count here:

Afros, Guns, and Voodoo: the Black Descriptions that Sell Video Games

A Term Frequency-Inverse Document Frequency and Content Analysis of Race in Video Game Reviews

Supervisor: Professor Shakuntala Banaji

Dissertation (MC499) submitted to the Department of Media and Communications, London School of Economics, August (2020), in partial fulfilment of the requirements for the MSc in Media and Communications (Data and Society)

Acknowledgements

I would like to thank my mother for all of the support she has given me throughout my life. Without the love and sacrifices she has made, I truly would not be the person I am today, nor would I be where I am.

I would like to thank my father, whose eternal flame of a love for learning transferred to me when his life's light went out.

I would like to thank my brother for introducing me to the Sega Saturn when I was three years old, for cheering me on even when he knew the controller was unplugged, and for always believing in me.

I would like to thank Professor Shakuntala Banaji for supporting me through these pursuits. I will forever be grateful for her guidance and for her ability to make a safe space wherever she goes.

ABSTRACT

This research project investigates the importance of racialized terms in U.S. gaming reviews. As a form of media, video games provide a space of learning for consumers. As social productions, video games themselves have the power to both perpetuate and destroy racist hegemony, specifically through their support or condemnation of racial stereotypes. However, as a media, video games are significantly lacking in regard to diverse representation, and although there has been an increased demand by marginalized gamers for the industry to incorporate more diverse portrayals, video games featuring characters of color are still limited. In the gaming community, professional video game reviews occupy a space of authority, as they give consumers insight into whether they should invest their money and time into purchasing and playing a game. With this idea in mind, the primary research aim of this project is to explore how the racial identity of a protagonist affects language used by reviewers in video game reviews.

By applying a mix of term frequency-inverse document frequency and content analysis to a sample of 80 game reviews, this research project seeks to answer the question: does the race of a protagonist in a video game affect the language used in gaming reviews? Some of the key findings suggest that video game reviews with Black protagonists have higher average TF-IDF scores for terms related to racial identity and racial stereotypes. However, the largest average scores occurred for terms related to gameplay for both Black protagonists and white protagonists. Using a conceptual framework of critical race theory, feminist theory, and socio-linguistic theory, this project aims to investigate how racialized language permeates into the gaming community.

1. INTRODUCTION

To take a stance and believe that representation matters should not be a revolutionary statement—and yet, it is. This is a particularly rebellious statement in the gaming community. When faced with demands for more progressive, nuanced, and diverse forms of representation, the misogynistic and racist toxicity of the Western gaming community—a community constructed of mostly white males—boiled to the surface. For instance, the ex-partner of Zoe Quinn, a prominent voice in the movement and video game developer, alleged that they engaged in sexual acts with multiple game reviewers in exchange for favorable reviews for their video game (Gray & Leonard, 2018, p. 11). This malicious act of sexual harassment inspired an entire movement, known as GamerGate. Although GamerGate was a movement within the gaming

community meant to inspire positive change through the inclusion of diverse representation, one of its most defining features is the narrative of white victimhood that feeds the Far-Right (Lees, 2016). Rhetoric deeming progressive individuals in the gaming community as "snowflakes" and "cry bullies" has been co-opted by the Alt-Right for the use against progressive voices in general (Lees, 2016). I will explore GamerGate in more theoretical terms later on in this paper, but I have chosen to highlight this important moment in gaming history to emphasize the fact that gaming culture is not insular. Video games, and the culture fostered by them, do not dwell in a space isolated "from other media forms" and social entities (Shaw, 2014, p. 4). While there is a belief in gaming studies that video games are neutral media forms, it is important to remember that video game culture aided in the growth and perpetuation of Far-Right culture and tactics online.

GamerGate formed the catalyst for my critical interest in race and video games. The scholarship dedicated to studying race in video games is lacking (Shaw, 2014, p. 19). However, this could be a result of the limited representation in the gaming community. Being one of the more recent types of media, one would expect video games to incorporate more open and inclusive images promoting representation (Shaw, 2014, p. 6). From a historical standpoint, games "developed alongside media representation critiques, numerous civil rights movements, and increased visibility of marginalized groups across popular media," which should have some effect on how video games present storylines (Shaw, 2014, p. 7). However, from a technological standpoint, the gaming industry's ability to introduce more racially diverse characters is a recent technological accomplishment. According to Everett and Watkins (2008), "a limited number of pixels" and "technological constraints" restricted developers' ability to create detailed, racially diverse characters in the early era of faming (p. 143). However, in the contemporary era of video games, technological advancements "have opened the way for upgraded representational depictions, and a more diverse range of themes and characters" (Everett & Watkins, 2008, p. 143). Nonetheless, the gaming industry continues to disappoint individuals from marginalized communities by neglecting to incorporate more diverse racial representations.

Studying video games through a critical media lens is important as video games continue to become more popular. As I will explore later in this paper, video games are ideological vehicles that teach and reaffirm hegemonic beliefs. For instance, video game depictions of Black men often incorporate "dominant discourses (and fears) of Black and brown masculinity" to

portray Black men as hyperviolent thugs (Leonard, 2006, p. 85). Another popular stereotype is the repeated portrayal of Arab men as terrorists, which, I would argue, arises from Anti-Arab sentiment in U.S. society (Leonard, 2006, p. 84). While I have chosen to focus on Blackness in video game reviews for this project, I cannot stress enough how important it is to explore racial representation in video games. As video games grow in popularity, harmful stereotypes continue to become reified in society, and as critical media scholars, we have a responsibility to explore how both old and new forms of media affect society. This research project aims to investigate how racialized ideas in video games permeate into video game reviews.

2. THEORY

2.1 Video Games as Forms of Media

There is a belief in gaming studies that video games are disconnected from other forms of media, and thus, are also disconnected from notions of race, gender, sexuality, and class (Harrer, 2018; Shaw, 2014). However, video games, like other forms of media teach, reaffirm, and reify social ideologies. According to Davis and Gandy (1999), representations found in media play an important role in "informing the ways in which we understand social, cultural, ethnic, and racial difference" (p. 367). Media is a tool that facilitates learning, and as a form of media, video games play a significant role in supporting social learning (Leonard, 2006). However, racial representation in video games often depict people of color in a "binary form of representation" that segregates them from white people and portrays them in simultaneously positive and negative lights (Hall, 1997, p. 229). As interactive forms of media, video games act as racialized pedagogical zones that teach "entrenched ideologies of race and racism" (Everett & Watkins, 2008, p. 150).

2.1 The Myth of Internet Neutrality

When the Internet was first popularized in the early 90's, it was marketed as a "utopian space for identity play, community building, and gift economies" (Nakamura, 2008, p. 3). This version of the Internet is a myth, as the individuals who created it were humans with biases and prejudices informed by society. Benjamin (2019) reflects on this idea, stating that "tech designers encode judgements into technical systems but claim that the racist results of their designs are entirely exterior to the encoding process" (p. 6). The entirety of the Internet industry is marked by racial inequality, being "predominantly led by white men and a few white women"

(Daniels, 2012, p. 697). For this reason, ideas surrounding race, both empowering and disparaging, continue to exist upon and within the Internet (Daniels, 2012, p. 696). In fact, race is physically built into the Internet, as seen in the "white hand pointer" that "becomes 'attached' to depictions of white people" (Daniels, 2012, p. 696). In the present era, individuals can change the white hand; however, it, much like whiteness in general, acts as a technological default and reminder of societal norms. Although some hope that technological advancements, like the Internet, have the power to aid in the reformation of a more accepting society, these advancements "too often reinforce racism and other forms of inequity" (Benjamin, 2019, p. 1). Therefore, to believe that the Internet, along with its successors, does not reflect issues of the social world is irresponsible, especially when looking at how people of color, specifically Black people, and women are treated in this space.

2.1 Anti-Blackness and Technology

According to Bailey (2016), "misogynoir describes the co-constitutive, anti-Black, and misogynistic racism directed at Black women, particularly in visual and digital culture" (p. 1). One particular case in which the Internet weaponized misogynoir against Black women occurred in 2013 when Far-Right trolls from 4Chan launched a series of cyberattacks against Black feminists on Twitter (Donovan, 2019). The Far-Right trolls' dedication to studying and embodying the language of Black feminists for the purpose of posing as them is a type of identity play (Nakamura, 2008). While Butler (1988) explores the performative nature of identity as it relates to gender, I would argue that her assertions apply to any form of social identity. Race, in the terms of this paper, is not simply a biological factor; rather, it is performative because race is identity. Anti-Blackness and racial performativity intersect in the gaming world, particularly in the urban/street genre where players can engage in digital minstrelsy by playing characters based on harmful stereotypes (Everett & Watkins, 2008, p. 149). While the 2013 example of cyber-misogynoir did not occur in the gaming world, it is important to recognize that in both instances, anti-Blackness is facilitated through the internet by "whites' fear of and fascination with black bodies" (Everett & Watkins, 2008, p. 149).

2.3 Toxic Gaming Culture and the Fight Against Representation

Gaming culture as a whole is built upon the foundations of homophobia, misogyny, and racism, which are all linked to what bell hooks (1997) defines as the "white supremacist capitalist patriarchy" (p. 7). The white supremacist capitalist patriarchy, in essence, provides a

wide consideration of intersectional oppression, which not only exists as a hegemonic pillar in US society; it informs the principles for gaming culture and the gaming industry. This white supremacist capitalist patriarchal structure exposed itself during the most important event to occur in the gaming world: GamerGate.

The white supremacist capitalist patriarchy is directly supported and perpetuated by the gaming industry. The video game industry is a space controlled by white men, which means that white, heterosexual men have the authority when it comes to the portrayal of narratives in video games (Gray & Leonard, 2018, p. 6). Additionally, within the wider culture of gaming, it is widely accepted "that white male lives are the only ones that matter" (Gray & Leonard, 2018, p. 8). I would add that it is not just white male lives that are thought to matter, but white, cis-gendered, heterosexual white male lives. The backlash of GamerGate was triggered by a perceived loss of hegemonic control by white men. According to Gray and Leonard (2018), white male gamers flocked to the reckoning that GamerGate brought, as it "empowered this narrative of white male victimhood" (p. 9). White men did not want to see more diverse representations because they believed that it would ruin the white male haven that gaming has provided them for decades. GamerGate was a war for white men, and in this space, they felt that they had to maintain control by any means necessary.

The catalyst for GamerGate occurred in 2014 when female gamers dared to question the lack of diverse representation within the gaming community (NJ Coalition Against Sexual Violence, 2014). The wave of backlash women faced during GamerGate included sexual harassment that reflected gamer culture and the assertion of white male dominance. This backlash included horrific examples of misogyny, misogynoir, harassment, and threats of rape, which, according to the New Jersey Coalition Against Sexual Violence (2014), are often used in in-game chats to declare a sense of dominance. Gray and Leonard (2018) assert this idea, as well, stating that in-game chats are filled with, "toxic practices of antisocial behavior, racism, heterosexism, and misogynistic language" (p. 5). The casual use of language associated with sexual and racial violence in these spaces are examples of "invisibly normalised colonial values" that further reify white male domination in the gaming community (Harrer, 2018, p. 15). Gamergate exposed the toxicity of the gaming community and the racist, misogynistic sentiment inspired by the white supremacist capitalist patriarchy.

2.4 Colorblindness

Games, as well as the culture of the gaming world, are deeply tied to race and gender norms, as expressed by Shaw (2014), who writes that the gaming world is a space for the "U.S. heterosexual, white, cisgendered-male, adolescent gamer stereotype" (p. 42). One of the most prominent racial norms supported by the video game industry is colorblindness, which prioritizes the idea that "race no longer 'matters' in American society" (Doane, 2014, p. 15). In theory, this notion shifts the United States into a post-racial space; however, as a country and society built both physically and ideologically upon a white supremacist patriarchy, I would argue that it is incredibly difficult to remove race and oppression from the heart of U.S. society. A prominent belief perpetuated by colorblindness is that oppression and struggles manifested through racial inequality are not products of a racist system; rather, racial inequality exists as a consequence of people of color's actions (Doane, 2014; Venegas, 2013). Through this lens, racism and racial inequality become "individualized," meaning that the responsibility of fighting for racial equality falls upon people of color rather than the social systems that continue to privilege whiteness (Venegas, 2013, p. 78).

Colorblind discourses contribute to the lack of representation in the gaming industry. In the gaming industry, companies prioritize white representation in order to cater to a majority white male audience. Exploring the market philosophy in the media industry as a whole, Shaw (2014) argues, "the market-logic argument seems to presume that dominant audiences do not wish to see people unlike them in their media" (p. 152). Thus, the assumption in the gaming industry that white gamers do not want to play games focused on characters of color shifts the responsibility of fighting for diversity onto marginalized gamers (Shaw, 2014, p. 153).

2.5 The White Default

In the gaming industry, whiteness is accepted as the default. As the dominant audience, white gamers do not have to do the extra work for diverse representation; rather, the games created and developed are made for a more readily available consumption. Exploring the social power of whiteness, Benjamin (2019) writes:

> The presumed blandness of White American culture is a crucial part of our national narrative. Scholars describe the power of this plainness as the invisible 'center' against which everything else is compared and as the 'norm' against which everyone else is measured (p. 2).

From a postcolonial perspective, invisible whiteness situates "the colonizer's point of view as the default," which is a tactic that privileges whiteness and disenfranchises marginalized perspectives (Harrer, 2018, p. 3). In the gaming world, there are clear postcolonial implications, specifically in regard to "systematic and ethnic oppression," as well as notions of white supremacy and sexual violence (Harrer, 2018, p. 1). Postcolonial implications are built into the nature of video games due to the emphasis on control (Brock, 2011, p. 432). The belief that games can be separated from historical, cultural, and social contexts is, within itself, a colonizer's perspective because once an object is disconnected from "implications of race, gender, and class," it becomes a "fun object" (Harrer, 2018, p. 3). This is only mostly true for white consumers, who do not have to confront the "detrimental consequences of seeing one's people repeatedly vilified or victimized" (Harrer, 2018, p. 3). Games normalize colonial values by reaffirming notions of "naturalized representations of events and situations relating to race which have racist premises and propositions inscribed in them" (Hall, 1995, p. 20). These notions are reaffirmed through the use of racial stereotypes in video games.

2.5 Racial Representation in Video Games

In recent years, there has been a call to action by gamers who want video game companies to create games featuring individuals with more diverse backgrounds (Russworm, 2017, p. 109). While the industry recognizes that it does, in fact, have issues with standing by the idea that representation matters, it continues to prioritizes white, heterosexual men, who, as assumed by video game companies, do not want to play games with diverse characters and representation (Russworm, 2017, p. 109). Looking at the racial demographics of video game characters, it is clear that the video game industry is resistant to abandon the hegemonic values of the white supremacist patriarchy upon which gaming culture has been built. According to Brock (2010), "racialized representations within videogames build upon hegemonic discourses of non-White difference, deviance, and alienation" that perpetuate "Whiteness as the default, 'normal' identity" (p. 444). The racial makeup of video game characters—particularly protagonists—speaks to the reification of whiteness as the default, as white characters represent the majority of protagonists and heroes. White male protagonists dominate the digital gaming realm, accounting for 52% of video game protagonists while Black male protagonists appear at a rate of roughly 40% (Everett & Watkins, 2008, p. 145). Regarding female protagonists, 80% of female protagonists are white while roughly 10% of female protagonists are Black (Leonard, 2006, p.

84). Additionally, looking at heroic representation, white characters have the highest likelihood of being the hero while Black characters have the least likelihood of being represented as the hero of a video game (Leonard, 2006, p.84). Thus, white characters, male and female, not only dominate the digital realm, but they do so in a positive light.

2.7 Racial Stereotyping and the Commodification of Black Digitized Bodies in Video Games

Thinking broadly about stereotypes, and their intimate relationship with racial representation, it is important to recognize that stereotypes are, in their own way, a form of racial representation (Hall, 1997). According to Hall (1997), "stereotyping reduces people to a few, simple, essential characteristics, which are represented as fixed by Nature" (p. 257). Racial stereotypes in video games are dangerous, as "repeated exposure" to harmful depictions of racial and marginalized groups reinforces the idea that a harmful interpretation is "a relevant schema for processing members of that particular group" (Brown et al., 2011, p. 292). While Shaw (2014) argues that "stereotypes are neither inherently good nor bad" due to their socially constructed and malleable presence in society, I would argue that stereotypes are, in fact, negative in nature. Racial stereotypes reinforce notions of 'otherness,' solidify notions of power within the social majority, legitimize false and harmful beliefs about groups in the social imagination, and encourage exoticization (Hall, 1997). Shaw (2014) asserts that stereotypes "must be constantly repeated and reaffirmed through media texts and social discourse to maintain their hold on the social imaginary" (Shaw, 2014, p. 20). Video games reaffirm racial stereotypes through the repeated negative depictions of Black people.

In video games, Black characters are often portrayed at the forefront of hyperviolence and street life (Brown et al., 2011; Everett & Watkins, 2008). In gaming spaces, Black characters are "often restricted to athletic, violent, and victim roles, or rendered entirely invisible" (Everett & Watkins, 2008, p. 43). The placement of Black characters in violent and victim roles are inspired by a Western history of enslavement and colonial violence (Hall, 1997, p. 239). However, the oversaturation of Black characters in sports games emphasizes stereotypes of athletic exceptionalism and sexual fetishism (Hall, 1997, pp. 230-231). Black characters also incorporate "some sort of voodoo or 'undead' quality" that has evolved from the archetype of the Magical Negro (Christensen et al., 2008, p. 25). Notions of misogynoir relegate Black women as being "more likely than any other group to be victims of violence" in video games (Everett &

Watkins, 2008, p. 148). They are also regarded as sexual objects that solidify "lessons about race and sexuality, especially the sexual mores" that accompany hypersexual stereotypes attached to Black woman (Everett & Watkins, 2018, p. 148). I would argue that the simultaneous hypersexualization and victimization of Black women in video games arises from the dehumanization, fetishization, and fascination surrounding Saartje Baartman (Hall, 1997, p. 264). The violence, victimization, hypersexuality, and criminality used to depict Blackness is directly related to anti-Blackness in mediated spaces.

2.7 The Significance of the Game Review

Most research on game reviews concerns the usability of the game review in relation to identifying points of improvement for game designs and "consumer advice" (Johnson et al., 2009; Strååt & Verhagen, 2017). Video game reviews act as a pillar in the gaming community, being "one of the primary forms of video game journalism" and "one of the prevalent forms of discourse about games" (Johnson et al., 2009, p. 215). Game reviews also act as a form of "consumer advice" that either encourages or dissuades players from buying games (Johnson et al., 2009, p. 215). Video game reviews as a source for consumer advice is particularly important. Video games are expensive, and many of them provide endless hours of gameplay; therefore, the game review helps consumers determine whether the monetary and temporal investments are worth it. Over the years, there has been criticism in gaming studies regarding game reviews, claiming that they lack a sense of consistency, professionalism, polish, and critical perspectives (Johnson et al., 2009; Nieborg & Sihvonen, 2009). According to Johnson et al. (2009), the anatomy of the professional game review routinely features "description, personal experience, reader advice, design suggestions, media context, game context, technology, design hypotheses, and industry" (p. 216). While these are the results of one study, I find it particularly interesting that the gaming industry, which claims it wants to incorporate more diverse representation, does not include character identity as a stable factor in its professional gaming reviews. For this reason, I am interested in exploring race in game reviews.

2.8 Research Question and Conceptual Framework

This project undertakes a critical race theory approach to explore racial representation and notions of race in game reviews. In particular, Stuart Hall's (1997) theory surrounding the spectacle of the 'other' plays a large role in the understanding of how historical contexts associated with colonialism and imperialism influence the ways in which Black bodies are

portrayed in the media. Ideas related to Butler (1988)'s theories surrounding the construction of identity have also informed the scope of this project. Joining theory from Hall (1997) and Butler (1988), I have framed this project to explore the construction of racial identity and the reification of racial stereotypes as racial identity. The concept of colorblindness and the systematic denial of racial difference will also play a large role in the framing of this project due to their prominence in the gaming industry as a whole. Colorblindness connects to the gaming industry's refusal to abandon its marketing philosophy that games with marginalized characters at the forefront do not sell. Colorblindness directs this marketing philosophy by placing the responsibility of the fight for representation on gamers from marginalized communities. In connection with colorblindness, van Dijk (1992) expresses that "contemporary racism is its denial," which not only solidifies colorblindness as a racist ideology but connects the denial of the existence of racism to language (p. 87). I am aware that video game reviewers may not use openly racialized language to discuss racial identity or portray racial stereotypes; therefore, I will employ the aid of van Dijk (1992), who argues that people who use disparaging language to discuss marginalized communities know that "they may be understood as breaking the social norm of tolerance or acceptance" (p. 89). In a colorblind society, it is not acceptable to openly use explicitly disparaging language, which means that in order to reaffirm racial ideology and stereotypes, the language used must be implicit. According to van Dijk (1992), the institution of media propagates the widespread "reproduction of racism" in order to affirm the superiority of whiteness in society (p. 88). This lens has informed the way that video game reviewers are perceived in this project. As the communication stewards of the gaming industry, video game reviewers and journalists act as authorities in the video game community. The language used by reviewers to describe video games, especially the characters, has the power to reinforce norms surrounding representation and race in the gaming community.

With this conceptual framework in mind, the objectives of this research aim to investigate how race is discussed in video game reviews. In particular, the research will focus on video game reviews for video games with Black protagonists and white protagonists. The main research objective of this project is to answer the question: *Does the race of a protagonist in a video game affect the language used in gaming reviews?*

One of the intentions of this project is to confront the lack of research surrounding race and video game reviews. As authorities in the gaming community, game reviewers have the power to

perpetuate ideas surrounding race, and I hope that this research will contribute to a wider conversation about the power that reviewers have in relation to reproducing hegemonic beliefs. There is a shortage of research on race in gaming studies; however, with the events of the Black Lives Matter Movement and the focus on bringing race to the forefront of social conversations, I am hopeful that more research will be produced. Finally, this research intends to introduce machine learning algorithms as a valid method within media and gaming studies for the purpose of increasing efficiency in text analysis. Based on the literature, I hypothesize that the race of the protagonist will influence the racialized language used in the reviews.

3. METHODOLOGY

In this chapter, I will highlight the methodologies used to investigate the research question: *Does the race of a protagonist in a video game affect the language used in gaming reviews?* Here, I will outline my research strategy, justify the methods used in this project, explain the parameters of my sample, and illustrate my research tools. I will also take some time to discuss ethics and reflexivity in regard to the research project.

3.1 Research Strategy

In order to investigate the question of how a video game protagonist's race affects language used in gaming reviews, I employed a mix of methods—term frequency-inverse document frequency (TF-IDF) and content analysis—to examine a sample of video game reviews on the popular review website IGN. I decided to use TF-IDF as a primary method with content analysis as an aid because I believe that TF-IDF can be used as a tool for automated and computer-aided content analysis. According to Conlon et al. (2001), research addressing information extraction often focuses on the evolution of software rather than "the effective application of such technology" (p. 43). They add that automated content analysis can aid researchers in "their decision making better than ever before" (Conlon et al., 2001, p. 43). The research conducted for this project questions the language used in video game reviews and if the language used treats race as relevant. I decided to pair content analysis with TF-IDF because it is a method that focuses on "counting the frequency," which directly relates to TF-IDF (Hansen, 1998, p. 96). Furthermore, conducting a content analysis simultaneously with the TF-IDF aided in increasing the understanding the implicit meaning that may accompany a writer's "lexical choice" (Hansen, 1998, p. 113). Thus, the content analysis acted as a training tool for interpreting

the nuances of the TF-IDF statistical scores of term importance that the algorithm may not understand.

As I will detail in this section, I decided to use IGN reviews because the company is an authority in the gaming community. In fact, according to the Alexa Web Services ranking system, IGN is the #1 gaming review website on the Internet. Regarding race, I identified video games with Black protagonists and white protagonists before searching for their respective IGN reviews. After identifying these video game reviews, I operationalized TF-IDF through Python and simultaneously conducted a content analysis to compare the TF-IDF machine learning algorithm to a human reading.

3.2 Term Frequency Inverse Document Frequency

I decided to incorporate term frequency-inverse document frequency because this quantitative method of text analysis provides a statistic for how important and relevant a term is in a document within a collection of documents. As a method, term frequency-inverse document frequency breaks down into two parts: term frequency and inverse document frequency. Term frequency incorporates the "frequency at which a term is found in a particular document" while inverse document frequency embodies "how frequent a term is found in all documents examined" (Chen, 2017, p. 113). Working together, term frequency and inverse document frequency calculate "a composite weight for each term in each document" within a corpus of documents (Manning et al., 2008, p. 118) The formula for term frequency-inverse document frequency is as follows: $tfidf\ (t, d, D) = tf\ (t, d) * idf\ (t, D)$ (Nyberg, 2018, p. 2). *Tfidf* is a function that takes a current term *t* in a current document *d* in a collection or corpus of documents *D* (Nyberg, 2018). While *tf* and *idf* represent separate functions, the TF-IDF formulation takes both *tf* and *idf* statistics into account to produce a statistical score of term relevance. Regarding the term frequency in the formula, *tf* is a function that takes term *t* in the current document *d* (Christenson et al., 2020; Nyberg, 2018). Term frequency calculates a score for the frequency of a term used in a document. It focuses on the "number of occurrences" for each term within a given document (Manning et al., 2008, p. 117). The inverse document frequency function (*idf*) takes the log of the total number of documents *D* divided by the number of documents where term *t* appears (Christenson et al., 2020; Nyberg, 2018). A significant factor of inverse document frequency in the TF-IDF formulation is that it applies a log weighting scale to term frequency that creates "a composite weight for each term in each document" within a

corpus of documents (Manning et al., 2008, p. 118). Additionally, applying a log weight scale to term frequency ensures that when a term has a high rate of appearances within the sample of documents, the ratio of the number of total documents in the collection where that term appears gets closer to 1. Therefore, the log becomes 0. The log weight penalizes words found in a high proportion of documents, and it stabilizes the weights given to rare terms in a collection to prevent skewed results. Log weight also minimizes document length bias, which ensures that longer documents are not favored (Buckley & Salton, 1987, p. 8). This is particularly important because the video game reviews included in the sample vary in length. Thus, the log weight scale supports "the enhancement of retrieval effectiveness" (Buckley & Salton, 1987, p. 6). The log weight is a filter, ensuring that the terms retrieved are relevant to each individual document in a sample while simultaneously discarding common terms that may create noise (Buckley & Salton, 1987, p. 6).

After running a TF-IDF, a score is given for each term within a document, ranging from the most relevant term to the least relevant term. According to Manning et al. (2008), "The score of a document *d* is the sum, over all query terms, of the number of times each of the query terms occurs in *d*" (p. 119). Essentially, the score communicates the statistical representation of how relevant a term is within a given document within a collection.

3.5 Sampling of Video Game Reviews

As a Black gamer, I am interested in investigating how the race of a protagonist affects the terms used in video game reviews. From an empirical standpoint, white and Black protagonists are the most prominently featured demographics in the gaming world with white protagonists as the most represented racial group. To investigate the research question, I created a sample by collecting 80 of the most representative video game titles with Black and White protagonists.

I decided to collect a sample of 80 video game reviews total from IGN featuring Black and white characters as the main characters, meaning that 40 of the video game reviews are for video games with Black protagonists while 40 are for white protagonists. For the purpose of this project, I define protagonists as not just main characters, but characters who have a central role in the narrative of the game. Due to the limited amount of games featuring Black protagonists, I was forced to make this distinction. Within the sample of video games with Black protagonists, there are reviews for games that feature multiple main characters that a player can play. I chose

to include these games because, I would argue, the representation of Black protagonists in games like *Left 4 Dead 2* and *Afterparty* matter, and I believe that it would be unfair to exclude them from the sample. I would also argue that my inclusion of these video games exhibits the lack of video games featuring solely Black protagonists—an issue that plagues the gaming industry. Although there is a diverse collection of genres included in the sample, I only included games where character customization does not exist as a gaming mechanism. Some of the games incorporate character personalization, but this mechanic only allows changes to physical features such as hair and clothing—not race.

For the content analysis, 16 of the 80 video games reviews were randomly selected using an online random number generator. I chose 16 out of 80 because it represents 20% of the total sample. Breaking down the sample even further, eight video games each were randomly chosen from my datasets of Black protagonist video game reviews and white protagonist video game reviews. Without time constraints, I would have conducted a content analysis on all 80 games to accompany the TF-IDF for more statistically significant results.

3.6 Data Collection and Processing for TF-IDF

I began by creating two separate text documents for the URLs with Black protagonist video game reviews and for white protagonist video game reviews. Once all 80 of the URLs were organized into their respective text documents, I used the Python package BeautifulSoup to scrape the text from the online reviews. I specified the data for my program to scrape by identifying the particular class in the source code. After looking at the source code for a few of the articles, I discovered that the class identifier used to discern the review text from the rest of the source code was the same for every review in the sample. This enabled me to use (class_ = “article-content”) in the code for web-scraping to return only the text included in the video game reviews. To ensure the data collected was a faithful representation of what readers would see on the IGN website, I coded the scraper to appear as a browser when scraping. The final piece of the web-scraping code saved all of the reviews in a corpus with columns identifying each URL, game name, and the review.

In the field of data mining, preprocessing refers to the act of cleaning data before feeding it into a model (Christenson et al., 2020, p. 11). This act of cleaning in data preprocessing allows the researcher the opportunity to give the data a “desired form” (Nyberg, 2008, p. 2). For TF-IDF and other Natural Language Processing algorithms, data preprocessing incorporates the

removal of words, known as stop-words, as well as punctuation. According to Nyberg (2008), "Stop words are words that usually contain little or no information by itself" (p. 2). Stop words are also words that are considered to be common, like articles and common verbs (Christenson et al., 2020; Nyberg, 2018). Because these words, along with punctuation, offer no valuable information, I removed them from the text reviews. The packages used for preprocessing include Python's regular expression library and the Natural Language Toolkit (Christenson et al., 2020, p. 11).

3.7 Design of Codebook for Content Analysis

When designing a codebook, it is imperative that the topics and themes of the categories include "dimensions or characteristics of texts" that will supply important information for analysis (Hansen, 1998, p. 106). Additionally, the categories should be created with "aims, objectives and foci of the research, on the theoretical framework and questions stated as part of the formulation of the research problem" (Hansen, 1998, p. 106). Kassarjian (1977) reinforces this idea, arguing that categories represent "the conceptual scheme of the research design" (p. 12). With these recommendations in mind, the codebook for this project was designed with concepts from past research on video game reviews and my research aims in mind. Research on both professional and player video game reviews suggest that the most common themes of video game reviews include game mechanics, usability, game issues, and suggestions (Harrer, 2018; Johnson et al., 2009; Ståål & Verhagen, 2017). More specifically, according to Johnson et al. (2009), the most common themes found in video game reviews include "description, personal experience, reader advice, design suggestions, media context, game context, technology, design hypotheses, and industry" (p. 216). I created categories and themes with these topics in mind, specifically gameplay, environment, references to character personalization, references to controls, and references to the "feel" of the game. As exhibited in Appendix B, the original codebook incorporated a code for game design; however, for the final codebook, I removed this theme as it created confusion during the coding process. I also eliminated the physical counts for Variable 6, opting to measure prevalence instead, as physical counts have the tendency to create variability. Because my project focuses on discussions of race in video games, I ensured that I included categories that record such topics. For instance, I included racial identity, racial stereotyping, presence of racialized language, and themes pertaining to praising diversity and lamenting a lack of diversity. While racial identity, racial stereotyping, and presence of racialized

language present some overlap, I would argue that they may also exist independently of each other. For instance, racial identity refers to the writer's explicit detailing of a protagonist's racial identity while racial stereotyping concerns the writer's use of racial stereotyping to describe or identify a protagonist. The presence of racialized language indicates a writer's use of racialized language, like using Black English in the video game review ironically or referring to a Black protagonist in a particularly disparaging way that reflects prejudice. I included variables to record the expressed sentiment of a reviewer towards a protagonist to investigate whether race has an effect on a reviewer's view of a protagonist. Variable 6 is particularly important, as it records the prominence of topics mentioned in the reviews, which directly connects with TF-IDF.

For the comparison of TF-IDF and human-coded content analysis, I decided that it would be appropriate to use the categories of Variable 6 to code the top terms words with the highest importance from the TF-IDF. The only modification made to this particular codebook and code schedule is that a variable to record gameplay was incorporated.

3.8 Intercoder Reliability for Content Analysis

As explored earlier, the content analysis for this project doubles as a companion study for a methodological comparison between the TF-IDF algorithm and a human-coder. To establish a sense of intercoder reliability, I chose a second coder to act as my co-coder. I also chose to incorporate a co-coder to aid in my own interpretation of the TF-IDF scores. Because the TF-IDF uses one of the variables from the companion content analysis, I wanted to ensure that my biases would not be skewed; therefore, I employed a co-coder for intercoder reliability to mitigate my own biases in the later interpretation of the TF-IDF scores. The individual chosen was a Lebanese American woman who, in the past, has dedicated much of her academic time to exploring notions of imperialism and postcolonialism. I decided to choose this individual because she does not have an extensive knowledge of video games, which, I would argue, minimizes bias. Together, my co-coder and I coded 16 randomly selected video game reviews, which is 20% of the overall sample used for the TF-IDF.

According to Krippendorff (2004), the acceptable level of agreement occurs when alpha is greater than or equal to .800 or 80% while "tentative conclusions" may be drawn when alpha is above .667 or 66.7% (p. 429). To calculate intercoder reliability and produce statistical analyses, I used R, an open-source coding and statistical software. Specifically, I used a package

known as ICR, which was created for the purpose of calculating intercoder reliability. The package contains a function that calculates Krippendorff's alpha coefficient, which formed the basis of my statistical analysis for intercoder reliability. I decided to consult Krippendorff's alpha coefficient because it is a widely used statistical measure used by researchers to compute agreement between coders (Gwet, 2015; Krippendorff, 2004). Regarding the calculated Krippendorff's alpha coefficients, I calculated two separate groups based on whether they were nominal or ordinal. Variables 3-5 were calculated nominally, as they are all variables that fall into distinct categories. Variable 6 was calculated on an ordinal scale due to the fact that the codes are ranked in that particular category. The alpha coefficient for Variables 3-5 is .7, which, according to Krippendorff (2004) is, in fact, an acceptable alpha coefficient for tentative conclusions. Additionally, the most variability between the intercoder and myself exists in the identification of themes of Variable 4—specifically the themes of violence, presence of racialized language, and notions of hypersexuality. I believe that these spaces of disagreement arise from my own experience as a gamer who is familiar with game reviews. I would also argue that my identity as a Black woman and experiences with Blackness aided in my ability to discern racialized language directed at Black protagonists. The alpha coefficient for Variable 6 is .889, which is above the threshold for acceptable results.

3.9 Ethics and Reflexivity

As explored earlier, video game reviews represent a form of video game journalism. IGN video game reviews are produced by professional gaming journalists and reviews for consumption in the public sphere. Due to the public nature of the reviews included in this study, I have determined that there is nothing harmful about this research. Additionally, I would like to stress that although this research focuses on how race is discussed in video game reviews, I, as the researcher, am not making any claims about any of the reviewers' social standings in regard to race.

Regarding reflexivity, I am a Black woman who is familiar with the U.S. hegemonic discourses and history surrounding race. I am also a gamer with a considerable amount of knowledge as it pertains to the gaming industry. I have attempted to mitigate bias by including a co-coder for the companion content analysis who does not have an extensive knowledge of the gaming industry, video games, or video game reviews, who, prior to this study, has an extensive knowledge of the gaming industry

4. RESULTS

In this section, the results for both the content analysis and the TF-IDF will be presented. For the content analysis, the results will include tables detailing the findings for the chi-square test of independence, as well as the degrees of freedom and the p-value. The TF-IDF results include a table, as well; however, the results are the averages of the TF-IDF scores for each variable for the video game reviews featuring Black and white protagonists. Once again, the results attempt to answer the research question: *Does the race of a protagonist in a video game affect the language used in gaming reviews?*

4.1 Results of the Content Analysis

While the quantitative variable of interest in this research project is Variable 6, I believe it is important to briefly discuss Variables 3-5. These variables include *video game genre, protagonist's racial identity, protagonist's gender identity, top 5 themes,* and *sentiment toward protagonist.* The most common genre included in this sample was the action genre, followed by the shooter genre. There was one game included that is a part of the urban/street genre, known as *Mafia 3.* Regarding *protagonist's racial identity,* discerning racial identity was not a difficult feat. In fact, only 3 out of the 16 games included in the sample were identified as having protagonists with unknown races. For variable *protagonist's gender identity*, the results only include protagonists with the gender identities: male and female. There were not any protagonists identified as transgender, non-binary, or genderqueer. Regarding Variable 4, the top five themes, as they appear in the game review, mostly include gameplay, game environment, violence, and racial identity. However, it appears that racial identity appears more frequently for games with protagonists who are coded as Black. Regarding character sentiment, the results do not suggest that race plays a role in a reviewer's positive or negative opinion of them.

The quantitative variable of interest in this research project is Variable 6, which measures the prevalence of particular topics present in the video game reviews. The selection of these topics was informed by past research on video game reviews and the research aims of this project. The topics in Variable 6 are as follows: references to character physical appearance, references to racial identity, references to racial stereotyping, references to diverse representation, references to character personalization, references to controls, and references to the feel of the game. To explore the relationship between a protagonist's race and the topics included in Variable 6, I ran a chi-square test. To specify, Variable 6 records the prevalence of

references in a game review in five categories: not mentioned, mentioned once, somewhat prevalent, prevalent, very prevalent. The chi-square statistical test is particularly appropriate for Variable 6, as it is comparing multiple categorical variables. According to Miller (2017), the chi-square test of independence is used to investigate the statistical significance between "multiple categorical variables from the same population" (p. 443). The null hypothesis for statistical tests using the chi-square test of independence typically state that there is no relationship between variables and that they are independent of each other.

Table 1: Chi–Square Test between Race and Topics Referenced in Video Game Reviews

	Character Appearance	Racial Identity	Racial Stereotyping	Diverse Representation	Character Personalization	Controls	Feel of the Game
Pearson Chi–Square	3.4061	3.6606	3.6923	3.6923	1.7692	4.941	6.2444
df	6	4	4	2	4	6	8
P–value	0.7564	0.4539	0.4492	0.1578	0.7781	0.5514	0.6199

For this chi-square test, a standard p-value of less than 0.05 was observed as the threshold to reject the null hypothesis in favor of the alternative hypothesis introduced above (Bauer et al., 1997). As Table 1 suggests, the null hypothesis cannot be rejected, meaning that a protagonist's race is, in fact, independent from the variables of character appearance, racial identity, racial stereotyping, references to diverse representation, references to character personalization, controls, and mentions of the feel of the game. This sample includes incredibly high p-values, which means that the results are statistically insignificant. The high p-values and statistical insignificance could be attributed to the small sample size used for the content analysis, which, for future research, I would recommend using a larger sample for greater statistical significance. Nonetheless, the results in Table 1 are still relevant, as Chi-Square coefficients tell the narrative of potential variable associations.

The chi-square test in Table 1 suggests that there could be some dependence between race and the topics included in Variable 6, but because the p-values are high, the results are not statistically powerful enough to reject the null hypothesis. Therefore, further studies will need to be conducted for more statistically significant results. The results for the Chi-square tests in Table 1 suggest that there could be a relationship between race and variables *character appearance, racial identity, racial stereotyping*, and *diverse representation*. This is clear due to

the fact that lower measures of chi-square statistics indicate the possibility of a relationship existing between variables. The variables listed above all incorporate measures that are within a 3.4 to 3.6 range. A particularly notable result is the fact that racial identity, racial stereotyping, and diverse representation share nearly identical chi-square measures at 3.6. Also, the chi-square statistical measures for racial stereotyping and diverse representation as they relate to race are the exact same statistic. The smallest chi-square measure in Table 1 is between race and the variable *character personalization*, which suggests that there could be an association between protagonist racial identity and reviewers' mentions of character personalization. Regarding popular topics in video game reviews, like the mechanical aspects of game controls and the expression of how it feels to play a particular game, the results suggest that there is a weaker association between said variables and the race of a video game protagonist. According to these results, more research is required, but I can tentatively state that there could be a relationship particularly between the race of a video game protagonist and the prevalence of references to character appearance, racial identity, racial stereotyping.

4.2 Results from the TF-IDF Study

The results of this chi-square test of independence helped to inform my decision to continue towards a TF-IDF. As explored earlier, TF-IDF scores communicate a statistical measure for the importance of terms within a document in a corpus (Christenson et al., 2020; Manning et al., 2008; Nyberg, 2018). To achieve the results in Table 2, I looked at the top ten most important terms calculated by the TF-IDF for all 80 games included in the sample and coded the sums of scores pertaining to the topics of Variable 6. From the content analysis, I decided to add an additional variable *gameplay* since that was a prevalent theme throughout the reviews, and I felt that it was an important theme to understand game reviews further. I chose to look at the top ten terms because the scores computed beyond ten lacked significance and, in many cases, appeared to be more common than the terms found in the top ten scores. Table 2 contains the average scores for each topic in Variable 6 for video game reviews featuring games with Black protagonists and games with white protagonists.

The results show some clear differences in the average scores of the topic between the two groups of Black protagonists and white protagonists. The variable *racial identity* differs greatly between reviews with Black protagonists and white protagonists. The variable *racial identity* shows a higher importance in Black protagonist games than white protagonist games

with the average score being double. In this sample, terms relating to racial identity are, on average, considered twice as important in video game reviews featuring Black protagonists than white protagonists. The average TF-IDF scores for the variable *racial stereotyping* shows a higher importance within Black protagonist games than white protagonist games with the average score being approximately seven times higher. Regarding the variables for terms pertaining to character appearance, controls, and the feel of the game, the results are comparable, suggesting that the race of a protagonist does not affect the importance of terms used to discuss these topics. For both Black and white protagonists, variable *diverse representation* calculated a score of 0. For variable *character personalization,* the average score is higher for video game reviews featuring Black protagonists than it is for reviews featuring white protagonists. The highest scores, on average, relate to gameplay for both groups within the sample. However, terms for variable *gameplay* are twice as important for video game reviews with white protagonists than they are for video game reviews with Black protagonists.

Table 2: Average TF–IDF Scores for Video Game Reviews with Black and White Protagonists

	Character Appearance	Racial Identity	Racial Stereotyping	Diverse Representation	Character Personalization	Controls	Feel of the Game	Gameplay
Black Protagonist	2.3876	11.0453	7.4663	0	1.0623	3.3063	4.3075	33.4728
White Protagonist	2.8933	6.7003	1.1078	0	0	3.303	5.5893	67.7739

5. DISCUSSION

The objectives of this research aimed to answer the question of how the race of a protagonist affects the language used in gaming reviews. Reflecting on the results, it is clear that the race of a protagonist does, in fact, affect the language used in gaming reviews. More specifically, the protagonist's race affects the importance of terms associated with racial identity and stereotyping. As explored in the results section, the findings for the content analysis suggest that there could be a relationship between a protagonist's race and the prevalence of language pertaining to racial identity and racial stereotyping. The TF-IDF scores corroborate this idea, as well, suggesting that words related to racial identity are more important in reviews featuring Black protagonists than they are for reviews featuring white protagonists. This can be attributed

to a multitude of factors, one of which being that "racialized representation" communicates "non-White difference" (Brock, 2010, p. 444). This essentially separates Black protagonists from white protagonists through racial identification. From a socio-linguistic perspective, the language of difference is significant because difference "construct[s] meaning through a dialogue with the 'Other'" (Hall, 1997, p. 235). Thus, language used to express racial identity achieves its goal through implications of constructed racial difference. Another potential contributing factor is that whiteness is not discussed in video game reviews due to the invisibility of whiteness in society and in the gaming community. As Benjamin (2019) argues, whiteness constructs "the invisible 'center'" and the "norm" (p. 2). With the status of the racial and societal default, whiteness does not require an introduction; rather, it is assumed. Therefore, it comes as no surprise that the results suggest that the most important terms for reviews with white protagonists do not include terms related to racial identity and racial stereotyping. Reflecting on the TF-IDF scores for reviews featuring white protagonists where the top ten most important terms include words associated with racial identity, the terms differentiate the white protagonist from identities considered to be 'other'. For instance, in *Red Dead Redemption,* terms associated with Mexico have high TF-IDF scores, meaning that they are important. While the reviewer does not discuss the protagonist's whiteness, the importance exhibited in the scores related to Mexico and Mexican people both highlights the protagonist's invisible whiteness and makes Mexican people hypervisible in the context of the game. A similar case presents itself with the *Wolfenstein* games. In these games, racial identity is exhibited through the emphasis and importance of terms related to Nazi Germany—and ideals associated with the racist ideology of Aryanism. Although whiteness is not explicitly mentioned, the semantics associated with Nazi Germany and the white idealism of Aryanism makes it clear that the racial identity of whiteness is important in the *Wolfenstein* games included in the sample. The highest TF-IDF score for racial identity and racial stereotyping in the sample of reviews with white protagonists is *Super Mario Odyssey,* where Mario's Italian identity is considered highly important. I do not believe that this is a coincidence, especially when considering the U.S. context where Italian people were constructed as a "lesser" group in the white racial hierarchy during the 20th century (Gerber, 1999, p. 440).

Regarding racial stereotyping in reviews for games with Black protagonists, some particularly interesting results arose that connect to the research aims of this project. In both the content analysis and the TF-IDF, racial identity and racial stereotyping overlapped, which

suggests that racial stereotypes act as tools of racial identification. This idea directly relates to notions of social and racial recognition as they connect to racial stereotypes. Venegas (2013) argues, "games take a racial meaning through the use of racial stereotypes, semiotics, and language" (p. 77). The overlap between racial identity and racial stereotype separates Blackness from whiteness by employing a separation of difference and the default (Hall, 1997, p. 258). Regarding difference and the default, whiteness represents "civilization" while Blackness represents "savagery," which is a binary that openly presents itself in the use of stereotypes in video game reviews. (Hall, 1997, p. 243) For instance, while protagonist names were included in the top ten most important TF-IDF scores, I chose to code names in racial identity and racial stereotype because "names are racially coded" (Benjamin, 2019, p.1). Names like Leroi in *Shadowman* and Akuji in *Akuji the Heartless* are not names associated with the invisibility and "blandness of White American culture" (Benjamin, 2019, p. 2). Rather, they act as racial markers separating so-called ghetto and exotic names from socially respectable white names. Further, important terms that communicated racial stereotypes included words, like urban, arrest, afro, and voodoo-affiliated expressions. These terms are directly linked to the stereotypical roles that Black protagonists and characters are limited to, which include urban archetypes, criminals, and voodoo priests (Christensen et al., 2008; Everett & Watkins, 2008). In particular, the afro hairstyle acts as a point of social tension due to the politicization of Black hairstyles in society. Also, voodoo, as a Black practice, represents the antithesis of white, Western Christianity. Additionally, these stereotypes perpetuate the "binary form of representation" by placing importance upon stereotypical terms when discussing Black protagonists, who are supposed to be the heroes of their own stories (Hall, 1997, p. 29). The importance of stereotypical terms used in the game reviews both produces and reproduces an image for "how 'difference' and 'otherness'" is portrayed in the gaming community (Hall, 1997, p. 232). While the results of the content analysis suggest that there could be a relationship between a protagonist's race and the prevalence of racial stereotypes used in video game reviews, the TF-IDF offers statistical evidence proposing that language pertaining to racial stereotyping is important when discussing video game reviews for games with Black protagonists.

Notions of game mechanics and gameplay had higher statistical importance in reviews featuring games with white protagonists. For the content analysis, the chi-square measure suggests that there may not be an association between race and references to the feel of a game

and controls. The TF-IDF statistical scores offer statistical evidence that terms associated with the feel of a game and gameplay are considered to be more important in game reviews for games with white protagonists than for reviews featuring Black protagonists. While the results for the feel of the game are only slightly higher for white protagonists, terms associated with feel are considered more important in reviews with white protagonists. The results for gameplay represent the highest average score for both racial demographics; however, for games featuring white protagonists, terms connected to gameplay are considered more important with an average that is double that of games featuring Black protagonists. This suggests that gameplay is the most important topic when discussing video games, which relates to claims made by Johnson et al. (2009) that game reviews break down into nine parts, including themes related to gameplay. Additionally, according to Nieborg and Sihvonen (2009), video game journalism and reviews have shifted to prioritize gameplay. Furthermore, the results connect to postcolonial notions of fun and the focus on "a set of mechanics rather than representation" in video game reviews (Harrer, 2018, p. 15).

While the results for the content analysis and TF-IDF largely agreed, a fascinating point of contention is in the variable *diverse representation*. The results of the content analysis suggest that there could be a relationship between the race of a protagonist and references to diverse representation. However, in the TF-IDF results suggest that in the list of the top ten most important words in the game reviews, terms associated with diverse representation are absent. I accredit this discrepancy to the lack of statistical significance of the sample for the content analysis. Thus, a suggestion I would make for a future study would be to include a larger sample for the content analysis to support a better comparison. Nonetheless, looking at the results for the TF-IDF, it is particularly fascinating that terms connected to diversity are not included in the most important terms, especially since the gaming industry has openly expressed that it wants to challenge itself by including more diverse representations (Russworm, 2017). As communication stewards of the gaming industry, game reviewers and journalists have a responsibility to rise to the occasion. However, the lack of importance placed upon diverse representation suggests that colorblind ideology not only exists in the process of creation for video games, but in the journalism associated with it, as well. I would also argue, in connection to colorblindness, the choice to treat diverse representation as an unimportant factor in the gaming community can be interpreted as the reproduction of racist hegemony. As van Dijk (1992) argues, "contemporary

racism is its denial" (p. 87). While the denial van Dijk (1992) discusses is verbal denial, I would argue that the choice to stay silent is also a form of denial. To withhold one's voice is a verbal choice; therefore, by treating diverse representation as a trivial topic in game reviews, reviews perpetuate racist hegemonic beliefs that prioritize white pleasure and punish marginalized gamers. Johnson et al. (2009) assert that video game reviews occupy the role of consumer advice in the gaming community. As consumer advice, gaming reviews help gamers to make decisions when purchasing video games. By neglecting to express the importance of diverse representation in their video games, reviewers and game journalists place the labor of celebrating multivalent forms of representation, lamenting the lack of representation, and fighting for more diverse representations onto marginalized individuals (Shaw, 2014, p. 152). I would encourage game reviewers to use their influential voices to place an emphasis and importance on the need for diverse representation.

For further research, I would suggest increasing the statistical power of the content analysis by introducing a larger sample. I believe that increasing the sample would provide clearer results with greater statistical significance. Increasing the diversity of the sample by including video game reviews from different professional review companies would provide interesting results, as well. A comparison of TF-IDF scores for reviews by different gaming companies could yield fascinating results regarding what different companies value in their reviews. Another interesting topic to explore would be a study of how the race of a video game reviewer affects the language used in their reviews when discussing characters of color. While gathering data from IGN, I noticed that they included the Black Lives Matter symbol in their logo for a few months during the widespread protests around the world. Future research comparing game reviews for video games with Black and white protagonists before and after the global movement would be particularly helpful for investigating how social movements affect game journalism. Finally, expanding on the TF-IDF portion of this study, a future research project could program the TF-IDF to create and organize the most important terms into different topic groups.

6. CONCLUSION

The main research objective of this project was to answer the question: *does the race of a protagonist in a video game affect the language used in gaming reviews?* After conducting the research and looking at the results, I conclude that the race of a video game protagonist does, in

fact, affect language used in video game reviews. Using TF-IDF scores to weigh the language used in video game reviews indicated that there is a larger importance for words pertaining to racial identity and racial stereotypes for Black protagonists than for white protagonists. Connecting the results to literature in critical race theory, feminist theory, and linguistic theory, it is clear that the importance of words pertaining to racial identity and racial stereotypes are used to indicate a sense of racial difference for Black protagonists. Reflecting on the results, the most disappointing is the 0 score for variable *diverse representation* for both groups. I was expecting for video games reviews featuring Black protagonists to have TF-IDF scores that rated diversity as an important term; however, I do believe that the score of 0 is particularly helpful for understanding the colorblindness that exists in the gaming industry. As the literature about gaming reviews suggested, gameplay is incredibly important when it comes to the content of video game reviews, and for both groups, variable *gameplay* had the highest average scores. I found it particularly interesting that video game reviews featuring white protagonists had a TF-IDF score average that was about double the score for reviews with Black protagonists. While I do not have any explanations for the difference in average scores for variable *gameplay*, I do think that further research could provide some insight.

A part of this research project was to showcase the benefits of TF-IDF. As explained earlier, I incorporated a content analysis to essentially act as a companion for my own interpretation of the TF-IDF terms and scores. Although the sample used for the content analysis was too small to be statistically significant, I still believe that the content analysis itself greatly aided in this project. The creation of the content analysis codebook, which was informed by research on game reviews and my own research aims, helped in the interpretation and organization of terms associated with the TF-IDF scores. I believe that TF-IDF could aid in future media text analyses by showing researchers what the most important terms are in a media text. I think that using TF-IDF could help to decrease monetary and time costs that accompany media text analyses.

For further research, I would continue to study race in video game reviews because the scholarship is incredibly limited. I would also push the boundaries even further with the TF-IDF by incorporating a more sophisticated techniques to further understand the importance of different topics in video game reviews.

References

Bailey, M. (2016). Misogynoir in Medical Media: On Caster Semenya and R. Kelly. *Catalyst: Feminism, Theory, Technoscience,* 2(2).

Bauer, P., Hung, J. H. M., Kohne, K., & O'Niell, R. (1997). The Behavior of the P-Value When the Alternative Hypothesis is True. *Biometrics,* 53(1), 11-22. DOI: 10.2307/2533093

Benjamin, R. (2019). Introduction: The New Jim Code. *Race After Technology: Abolitionist Tools for the New Jim Code*, (pp. 1-25). Jon Wiley & Sons.

Brock, A. (2011). '"When Keeping it Real Goes Wrong"': Resident Evil 5, Racial Representation, and Gamers. *Games and Culture*, 6(5), 429–452. https://doi.org/10.1177/1555412011402676

Brown, B. P., Burgess, M. C. R., Burgess, S. R, Dill, K. E., & Stermer,. S. P. (2011). Playing with Prejudice: The Prevalence and Consequences of Racial Stereotypes in Video Games. *Media Psychology*, 14(3), 289-311, DOI: 10.1080/15213269.2011.596467

Buckley, C., & Salton, G. (1987). Term Weighting Approaches in Automatic Text Retrieval. Cornell University Press.

Butler, J. (1988). Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory. *Theatre Journal*, 40(4), 519–531. JSTOR. https://doi.org/10.2307/3207893

Chen, C. H. (2017). Improved TFIDF in big news retrieval: An empirical study. *Pattern Recognition Letters*, 93, 113–122, DOI: https://doi.org/10.1016/j.patrec.2016.11.004

Christensen, J., Dickerman, C., Kerl-McClain, S. B. (2008). Big Breasts and Bad Guys: Depictions of Gender and Race in Video Games. *Journal of Creativity in Mental Health,* 3(1), 20-29, DOI: 10.1080/15401380801995076

Christenson, L., Dong, Z., Fulton, L., & Meng, L. (2020). Mining Public Opinion on Twitter about Natural Disaster Response Using Machine Learning Techniques. *ArXiv*: abs/2005.07019

Conlon, S., Mukhopadhyay, S., Simmons, L. L., & Yang, J. (2011). A Computer Aided Content Analysis of Online Reviews, *Journal of Computer Information Systems*, 52(1), 43-55, DOI: 10.1080/08874417.2011.11645521

Daniels, J. (2012). Race and racism in Internet Studies: A review and critique: *New Media & Society*. https://doi.org/10.1177/1461444812462849

Davis, J. L., & Gandy, O. H. (1999). Racial Identity and Media Orientation: Exploring the Nature of Constraint. *Journal of Black Studies*, *29*(3), 367–397.

Dijk, T. A. van. (2016). Discourse and the Denial of Racism: *Discourse & Society*. https://doi.org/10.1177/0957926592003001005

Doane, A. (2014). Shades of Colorblindness: Rethinking Racial Ideology in the United States. In S. Nilsen & S. E. Turner (Eds.), *The Colorblind Screen* (pp. 15–38). NYU Press; JSTOR. https://www.jstor.org/stable/j.ctt9qg55f.4

Donovan, J. (2019, August 15). FirstThey Came for the Black Feminists. *New York Times.* https://www.nytimes.com/interactive/2019/08/15/opinion/gamergate-twitter.html

Everett, A., & Watkins, S. (2008). *The Power of Play: The Portrayal and Performance of Race in Video Games*.

Gerber, D. A. (1999). Caucasians Are Made and Not Born: How European Immigrants Became White People. *Reviews in American History*, *27*(3), 437–443; JSTOR.

Gray, K., & Leonard, D. (Ed.). (2018). *Woke Gaming: Digital Challenges to Oppression and Social Justice.* Seattle: University of Washington Press; JSTOR. http://www.jstor.org/stable/j.ctvd7w7f6.

Gwet, K.L. (2011). On The Krippendorff's Alpha Coefficient.

Hall, S. (1995). The Whites of Their Eyes: Racist Ideologies and the Media. In: Dines, G., Humez J. M. (Eds.), *Gender, Race and Class in Media* (pp. 18-22). London: Sage.

Hall, S. (Ed.). (1997). *Representation: Cultural Representations and Signifying Practices*. Sage Publications, Inc; Open University Press.

Hansen, A. (1998). Content Analysis. *Mass Communication Research Methods*. NYU Press.

Harrer, S. (2018). Casual Empire: Video Games as Neocolonial Praxis. *Open Library of Humanities*, *4*(1), 5. https://doi.org/10.16995/olh.210

Hooks, B. (1997). Bell Hooks—Cultural Criticism & Transformation (S. Jhally, Interviewer). Retrieved from https://www.mediaed.org/transcripts/Bell-Hooks-Transcript.pdf

Johnson, T., Ladd, A., & Zagal, J. (2009). Characterizing and Understanding Game Reviews. 215-222. DOI: 10.1145/1536513.1536553

Kassarjian, H. H. (1977). Content Analysis in Consumer Research. *Journal of Consumer Research*, 4(1), 8-18. Oxford: Oxford University Press.

Krippendorff, K. (2004). Reliability in Content Analysis: Some Common Misconceptions and Recommendations. *Human Communication Research,* 30(3), 411-433.

Lees, M. (2016, December 1). What Gamergate Should Have Taught Us About the 'Alt-Right'. *The Guardian.* https://www.theguardian.com/technology/2016/dec/01/gamergate-alt-right-hate-trump

Leonard, D. J. (2006). Not a Hater, Just Keepin' It Real: The Importance of Race- and Gender-Based Game Studies. *Games and Culture*, *1*(1), 83–88. https://doi.org/10.1177/1555412005281910

Manning, C., Raghavan, P., & Schütze, H. (2008). Scoring, Term Weighting, and the Vector Space Model. *Introduction to Information Retrieval* (pp. 100-123). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511809071.007

Miller, S. J. (2017). The Chi-square Distribution. *The Probability Lifesaver: All the Tools You Need to Understand Chance* (pp. 427–446). Princeton University Press; JSTOR. https://doi.org/10.2307/j.ctvc7767n.20

Nakamura, L. (2008). Introduction: Digital Racial Formations and Networks Images of the Body. *Digitizing Race: Visual Cultures of the Internet* (pp. 1-36). University of Minnesota Press; JSTOR. http://www.jstor.org/stable/10.5749/j.ctttswwb.4

New Jersey Coalition Against Sexual Violence. (2014, November 14). *Gaming Culture & Rape Culture: How #GamerGate's Misogyny Prevents a Safer Space*. https://njcasa.org/news/gaming-culture-rape-culture-gamergates-misogyny-prevents-safer-space/

Nieborg, D., & Sihvonen, T. (2009). The New Gatekeepers: the Occupational Ideology of Game Journalism.

Nyberg, A. (2018). Classifying movie genres by analyzing text reviews. *ArXiv:1802.05322 [Cs]*. http://arxiv.org/abs/1802.05322

Russworm, T. (2017). Dystopian Blackness and the Limits of Racial Empathy in the Walking Dead and the Last of Us. In T. Russworm & J. Malkowski. (Eds), *Gaming Representation: Race, Gender, and Sexuality in Video Games* (pp. 109-128). Bloomington, Indiana: Indiana University Press. Retrieved from http://www.jstor.org/stable/j.ctt2005rgq.11

Shaw, A. (2014). *Gaming at the Edge: Sexuality and Gender at the Margins of Gamer Culture*. University of Minnesota Press; JSTOR. http://www.jstor.org/stable/10.5749/j.ctt1287nqh.

Strååt, B., & Verhagen, H. (2017). Using User Created Game Reviews for Sentiment Analysis: A Method for Researching User Attitudes*,GHITALY,* 2019.

Upton, G. J. G. (2000). Conditional Independence, the Mantel-Haenszel Test, and the Yates Correction. *The American Statistician*, *54*(2), 112–115. JSTOR. https://doi.org/10.2307/2686027

Venegas, M. (2013). The Racialized Experiences of Video Games. *McNair Scholars Research Journal,* 9(1).

Appendix A: TF-IDF Scores

enter-the-matrix-review	
matrix	61.117
enter	23.205
hype	18.805
film	16.575
bullettime	14.104
comparison	13.794
payne	12.888
movie	11.517
max	9.945
urban-chaos	
eidos	30.28
dreamcast	30.071
urban	25.219
chaos	20.976
rate	20.976
frame	20.975
arrest	18.805
city	18.391
nt	18.279
resident-evil-5-review-2	
chris	42.312
sheva	37.61
resident	24.14
evil	17.529
capcom	15.14
kijuju	14.104
umbrella	12.888
nt	12.186
wo	11.318
left-4-dead-2-review	
zombie	23.972
dead	23.085
left	19.92
survivor	19.183
jockey	17.183
versus	16.033
spitter	14.104
four	12.25
team	10.245
remember-me-review	
nilin	23.507
memory	22.043
combo	18.551
remember	16.118
combat	12.116
lab	12.024
section	11.022
really	10.518
ledge	9.592
the-walking-dead-the-game-review	
walking	40.778
dead	37.775
lee	28.057
telltale	24.049
episode	18.551
adventure	17.939
choice	14.436
mobile	13.794
ipad	12.888
the-walking-dead-the-final-season-review	
aj	103.429
clem	94.508
walker	46.835
episode	34.011
season	33.15
final	25.612
cycle	19.89
child	19.183
freeroaming	18.805
sanity-aikens-artifact-2	
talent	94.625
sanity	90.213
totem	85.917
psionic	47.254
monolith	47.013
cain	37.61
power	34.618
deathmatch	26.495
different	25.213
grand-theft-auto-san-andreas-special-edition	
andreas	21.479
san	18.925
software	14.411
auto	13.26
theft	13.26
grand	11.986
still	10.81
question	10.016
additionally	8.016
grand-theft-auto-v-review	
gta	49.725
andreas	47.254
theft	43.095
auto	43.095
san	41.635
trevor	38.663
franklin	37.61
grand	35.959
santos	28.208
starhawk-review	
starhawk	84.623
multiplayer	24.178
flag	21.479
graf	21.479
emmett	18.805
hawk	14.104
shooter	12.52
nt	12.186
vehicle	9.592
mafia-3-review	
mafia	65.818
bordeaux	18.805
loading	17.15

pc	16.449
lincoln	14.104
definitely	11.986
fast	11.986
setting	11.388
filled	11.318
watch-dogs-2-review	
dog	23.276
marcus	21.479
hackable	18.805
watch	18.799
bay	15.14
francisco	14.104
detonate	12.888
gang	12.789
hacker	12.024
prototype-2-review	
heller	56.416
prototype	55.846
mercer	14.104
absorbing	12.024
york	11.355
nt	10.663
soldier	9.945
blackwatch	9.403
blackboxes	9.403
the-suffering-2	
suffering	28.057
torque	14.104
opening	11.638
horror	9.945
hell	9.276
prison	8.016
crawling	8.016
thrill	7.57
meat	7.57
blade-2	
blade	46.41
quake	17.183
analog	13.26
nt	10.663
control	9.527
actually	9.49
activision	9.403
purported	9.403
hammerhead	9.403
assassins-creed-iii-liberation-review	
liberation	64.13
aveline	47.013
assassin	43.286
creed	43.232
persona	30.071
iii	26.495
slave	18.805
franchise	15.459
vita	14.104
crackdown-3-campaign-review	
crackdown	89.325
trigger	26.969
shooting	24.178
orb	24.049
nova	23.507
terra	23.507
agility	20.041
vehicle	19.183
efficient	17.183
bad-day-la-review	
humor	32.424
disaster	25.775
la	17.243
bad	11.763
day	11.295
style	10.682
gameplay	10.238
art	9.8
angeles	8.592
afro-samurai-review	
afro	84.623
samurai	56.416
ninja	32.065
slice	20.041
nt	18.279
focus	14.7
slicing	14.104
lop	14.104
card	12.367
shadow-man-review	
shadow	25.041
man	19.933
mike	18.805
legion	16.033
deadside	14.104
voodoo	11.355
using	9.802
leroi	9.403
acclaim	9.403
crysis-2-review	
crysis	77.325
nanosuit	28.208
suit	12.789
armor	10.808
ambush	9.403
crynet	9.403
weapon	8.654
console	8.266
subset	8.016
dead-island-riptide-review	
island	46.378
riptide	32.909
zombie	23.972
dead	23.085
coop	16.575
original	11.763
level	9.799
techland	9.403
gory	8.016
marc-eckos-getting-up-contents-under-pressure-2	
trane	28.208
graffiti	24.049

artist	20.691
graf	17.183
book	15.367
dream	14.411
hiphop	14.104
legend	11.355
urban	10.808
true-crime-new-york-city-2	
crime	24.735
marcus	12.888
york	11.355
TRUE	11.082
father	8.729
generally	8.266
duty	8.016
street	7.684
nt	7.616
dishonored-death-of-the-outsider-review	
dishonored	60.142
outsider	47.013
displace	37.61
foresight	23.507
billie	23.507
ability	21.538
mission	18.49
energy	16.575
death	14.69
half-life-alyx-review	
alyx	65.818
halflife	54.04
vr	38.663
magazine	28.208
combine	23.205
gravity	22.71
reloading	21.616
round	19.89
valve	18.925
akuji-the-heartless	
akuji	42.312
heartless	18.805
underworld	12.888
crystal	8.592
voodoo	7.57
raider	7.205
tomb	6.897
magic	6.184
brother	5.993
assassins-creed-origins-review	
origin	38.663
creed	25.219
assassin	24.735
bayek	18.805
quest	16.791
ancient	16.033
egypt	14.104
climbing	12.789
mission	10.085
afterparty-review	
afterparty	56.416
milo	47.013
lola	47.013
hell	40.194
school	36.027
night	25.578
choice	20.622
earth	20.041
oxenfree	18.805
killer-7	
killer	46.378
smith	24.049
capcom	18.925
assassin	15.459
seven	14.411
adventure	11.96
mask	11.355
guide	11.355
ask	9.592
virginia-review	
virginia	51.714
tarver	47.013
halperin	32.909
fbi	18.805
scene	18.244
emotionally	14.411
town	12.789
kingdom	12.024
haunting	12.024
sunset-review	
angela	21.479
sunset	20.691
ortega	18.805
tale	14.1
chore	10.346
meditative	9.403
torn	8.592
cared	8.592
sense	7.244
25-to-life	
lean	17.243
nt	10.663
freeze	9.945
cover	9.8
life	8.654
leaning	8.592
pretty	7.842
backup	7.57
officer	7.57
unreal-ii-the-awakening-review	
unreal	38.663
ii	21.616
approach	11.082
planet	10.808
bit	9.799
pt	9.403
sophisticated	8.592
review	8.249
level	8.166
saw-review	
jigsaw	23.507
tapp	18.805

victim	17.243
film	16.575
detective	15.14
trap	13.794
horror	13.26
asylum	12.024
saw	10.487
the-walking-dead-michonne-miniseries-review	
michonne	61.117
episode	34.011
walking	28.785
dead	25.183
miniseries	14.104
reviewverdict	14.104
season	13.26
book	10.245
telltale	8.016
falcon-age-review	
falcon	47.013
vr	42.958
age	35.959
bird	31.037
outerloop	18.805
pet	12.888
dualshock	12.888
care	11.994
explore	11.994
broken-age-act-i-review	
age	41.952
broken	37.824
adventure	25.912
shay	17.183
vella	17.183
puzzle	17.091
act	16.449
schafer	14.104
double	13.794
broken-age-act-2-review	
act	35.249
vella	21.479
puzzle	19.227
shay	17.183
age	14.983
adventure	11.96
broken	11.638
loved	9.276
timing	8.016
the-last-of-us-part-2-review	
ellie	122.234
part	58.603
dog	55.281
joel	47.013
naughty	42.958
last	39.874
seattle	25.775
original	21.565
combat	15.578
final-fantasy-7-remake-review	
midgar	61.117
remake	44.09
cloud	41.635
materia	37.61
atb	32.909
fantasy	29.835
party	29.835
filler	28.208
weapon	25.963
death-stranding-review	
stranding	192.753
death	83.944
sam	75.656
kojima	42.958
bts	28.208
cargo	26.495
bridge	22.71
consistently	19.183
hill	18.013
wolfenstein-the-new-order-review	
nazi	55.846
wolfenstein	51.55
order	32.434
blazcowicz	23.507
alternate	17.457
history	16.532
especially	14.69
xbox	14.392
machinegames	12.888
red-dead-redemption-au-review	
redemption	36.073
marston	25.775
red	23.597
dead	23.085
mexican	16.033
rockstar	14.411
gamers	9.592
compass	9.403
ethic	9.403
mass-effect-review-2	
mass	100.875
effect	56.853
nt	36.558
quest	33.582
power	24.232
biotic	23.507
shepard	23.507
bioware	23.507
knight	23.507
assassins-creed-4-black-flag-review	
flag	55.846
black	33.296
creed	28.822
assassin	27.827
edward	18.805
ac	14.104
lighter	14.104
amazing	11.638
ubisoft	11.355
metal-gear-solid-5-the-phantom-pain-review	

phantom	108.13
pain	76.245
gear	46.553
metal	43.643
base	31.124
online	18.551
resource	18.013
prisoner	17.183
outpost	14.411

rise-of-the-tomb-raider-review	
tomb	100.008
raider	75.656
lara	72.147
rise	59.67
croft	17.183
broad	14.104
card	12.367
hunt	11.986
storyline	11.022

the-witcher-3-the-wild-hunt-review	
witcher	98.727
geralt	32.909
ciri	14.104
main	13.723
quest	9.595
rpg	9.592
wider	9.403
open	9.34
wild	9.276

someday-youll-return-review	
someday	65.818
return	42.982
daniel	28.208
section	13.777
potion	12.888
order	10.136
bulk	9.945
campsite	9.403
benefitted	9.403

horizon-zero-dawn-review	
horizon	94.026
aloy	42.312
machine	35.959
thanks	15.515
mystery	14.548
concentration	14.104
dawn	14.104
beast	13.794
fight	13.286

uncharted-the-lost-legacy-review	
legacy	29.835
uncharted	25.219
lost	22.05
chloe	18.805
nadine	14.104
naughty	12.888
empty	11.355
adventure	9.966
india	9.403

god-of-war	
kratos	38.663
god	32.237
are	23.507
war	23.051
zeus	14.104
man	9.966
olympian	9.403
throne	9.403
olympus	9.403

control-review-2	
jesse	61.117
bureau	51.714
control	28.58
remedy	28.057
oldest	23.507
house	15.367
chapter	14.548
sector	14.104
telekinesis	14.104

super-mario-odyssey-review	
mario	90.213
odyssey	65.818
super	21.643
cappy	18.805
nintendo	16.033
plumber	14.104
hat	12.024
moon	12.024
possession	12.024

silent-hill-2-2	
hill	108.081
silent	99.45
james	54.04
konami	42.958
camera	31.85
fog	30.071
cg	30.071
player	28.021
survivalhorror	26.495

deus-ex	
deus	73.029
ex	73.029
rifle	26.969
color	24.14
augmentation	24.049
mac	23.507
guard	20.033
grenade	17.979
weapon	17.309

resident-evil-review	
resident	31.037
evil	25.041
resource	10.808
tense	9.276
zombie	8.99
remake	8.016
update	8.016
persistent	8.016
pressure	7.57

bioshock-infinite-xbox-360ps3-review	
infinite	115.988
bioshock	87.055
columbia	51.714
elizabeth	32.909
booker	28.208
vigor	28.208
daddy	17.183
skyline	17.183
float	12.888
half-life-2-review	
valve	30.28
halflife	21.616
shooter	15.024
review	12.373
released	10.346
chance	10.016
registration	9.403
firstperson	8.729
system	8.051
wolfenstein-youngblood-review	
youngblood	65.818
wolfenstein	42.958
nazi	34.367
twin	18.805
cheat	18.013
coop	16.575
blazkowicz	14.104
level	13.066
bj	12.888
devil-may-cry-5-review	
devil	109.765
cry	69.615
nero	47.013
may	44.071
dante	37.61
breaker	36.073
technique	24.14
orb	24.049
dlc	18.805
hitman-2-review	
hitman	80.163
target	21.149
mission	16.809
disguise	16.033
neck	10.808
brand	10.808
point	10.518
sandbox	10.346
coin	10.346
vampyr-review	
vampyr	56.416
vampire	47.254
citizen	23.972
london	21.479
blood	19.806
reid	18.805
supernatural	18.013
dr	17.183
quest	16.791
dishonored-review	
dishonored	34.367
corvo	30.071
power	17.309
player	16.813
guard	15.024
mission	13.447
grabbing	12.888
oil	12.024
code	11.355
dead-rising-4-review	
rising	34.485
shelter	23.507
dead	20.986
willamette	18.805
holiday	12.888
zombie	11.986
collectible	9.945
memorial	9.403
feelgood	9.403
quantum-break-review	
quantum	42.312
joyce	23.507
break	22.537
serene	14.104
half	11.022
max	9.945
bullet	9.592
eyepopping	9.403
hero	8.729
homefront-the-revolution-review	
revolution	40.082
homefront	32.909
district	18.013
resistance	17.183
philadelphia	14.104
stealth	13.554
shooter	12.52
weapon	12.116
mission	11.766
sherlock-holmes-the-devils-daughter-review	
sherlock	42.312
devil	26.495
daughter	19.183
loading	17.15
holmes	14.104
detective	11.355
room	10.878
happening	10.808
interpret	9.403
the-longest-journey	
april	30.071
adventure	27.906
longest	25.775
tlj	18.805
ryan	17.183
puzzle	17.091
journey	16.449

average	16.033
future	15.459
red-dead-redemption-2-review	
arthur	107.396
horse	60.142
redemption	60.122
dead	41.972
red	36.706
rockstar	32.424
dutch	28.208
marston	25.775
gta	23.205
marvels-spider-man-ps4-review	
spiderman	147.615
peter	65.818
insomniac	56.416
york	34.065
web	30.071
spidey	23.507
gadget	19.89
city	18.391
villain	17.457
far-cry-review	
ram	15.14
rig	14.104
foliage	12.024
graphical	11.355
cry	9.945
gigabyte	9.403
distance	8.059
fade	7.57
ambient	7.57
infamous-review	
cole	37.61
city	22.069
empire	20.041
infamous	14.104
rooftop	12.789
lightning	12.024
run	10.233
sucker	9.403
grab	9.276
splinter-cell-blacklist-review	
blacklist	75.221
fisher	51.55
cell	36.073
splinter	34.367
spy	34.367
mercs	23.507
conviction	21.479
mode	21.216
chaos	20.976
spelunky-review	
spelunky	47.013
cave	18.925
spike	12.888
treasure	10.808
damsel	9.403
overcome	8.592
tunnel	8.592
hour	8.404
pit	8.016
doom-review-2	
demon	47.958
doom	47.254
chamber	25.775
kill	20.23
health	18.799
shotgun	15.367
doomguy	14.104
bfg	14.104
weapon	13.847
braid-review-3	
braid	75.221
tim	44.09
doppelganger	23.507
rewind	21.479
puzzle	19.227
fairy	18.805
platform	16.575
lever	14.104
princess	12.888
thief-review	
thief	52.106
guard	20.033
garrett	18.805
scenario	14.983
stealth	13.554
spotted	13.26
undetected	12.888
arrow	12.789
limited	11.986

Appendix B: Content Analysis Codebooks

Master's Thesis
Final/Revised Content Analysis Code Book

Variable 1: Video Game Title

Variable 2a: Review URL

Variable 2b: Author name

Variable 2c: Original Date of Review

Variable 2d: If the review has been revised, what is the new date?

Variable 3a: Video Game Genre
For this variable, use genre given in game review. If there are multiple, choose the one that resonates the most with you while reading the review.
1 = Action
2 = Shooter
3 = Urban/Street
4 = Survival
5 = Platformer
6 = Horror
7 = Roleplaying. (RPG)
8 = Life simulation
9 = Strategy
10 = Adventure
11 = Thriller
12 = Narrative

Variable 3b: Protagonist's Racial Identity
1 = Black
2 = White
3 = Other
4 = Racial Identity not mentioned/Racial Identity unclear

Variable 3c: Protagonist's Gender Identity
1 = Male
2 = Female
3 = Transgender
4 = Non-binary/genderqueer
5 = Gender Identity not mentioned/Gender Identity unclear

Variable 4: Themes Present in Article
Identify 5 maximum, in order of mention.
Count 0 if NONE

1 = Gameplay
2 = Environment (Gameworld Location)
3 = Racial Identity
4 = Racial Stereotypes
5 = Violence
6 = Hypermasculinity
7 = Hypersexuality
8 = Urban/Street Life
9 = Praising Diversity
10 = Presence of Racialized Language
11 = Lamenting Lack of Diversity

Variable 5: Sentiment toward Protagonist

1= Overall positive toward the protagonist
2 = Neutral toward the protagonist
3 = Overall negative toward the protagonist

Variable 6: Topics
Count the number of words/phrases per category
[0 = Not Mentioned, 1 = Mentioned Once, 2 = Somewhat Prevalent, 3 = Prevalent, 4 = Very Prevalent]
6a: References to Character's Physical Appearance
6b: References to Racial Identity
6c: References to Racial Stereotyping
6d: References to Diverse Representation
6e: References to Character Personalization
6f: References to Controls
6g: References to the "feel" of the game

Variable 7: Mentioning or presence of racial identity

If yes, describe how race is identified by the reviewer.

Variable 8: Coder Inferred race

If yes, describe how race was inferred during coding.

Variable 9: Mentioning or presence of racial stereotyping

If yes, describe in what sense the reviewer uses racial stereotypes to describe protagonists.

Variable 10: Other racial identities
If yes, describe how race was inferred during coding.

Variable 11: Quotes from Reviews
Include name of reviewer, quote that stood out, and how it relates to race or racial stereotyping.

Master's Thesis
Original Content Analysis Code Book

Variable 1: Video Game Title

Variable 2a: Review URL

Variable 2b: Author name

Variable 2c: Original Date of Review

Variable 2d: If the review has been revised, what is the new date?

Variable 3a: Video Game Genre
For this variable, use genre given in game review. If there are multiple, choose the one that resonates the most with you while reading the review.
1 = Action
2 = Shooter
3 = Urban/Street
4 = Survival
5 = Platformer
6 = Horror
7 = Roleplaying. (RPG)
8 = Life simulation
9 = Strategy
10 = Adventure
11 = Thriller
12 = Narrative

Variable 3b: Protagonist's Racial Identity

1 = Black
2 = White
3 = Other
4 = Racial Identity not mentioned/Racial Identity unclear

Variable 3c: Protagonist's Gender Identity
1 = Male
2 = Female
3 = Transgender
4 = Non-binary/genderqueer
5 = Gender Identity not mentioned/Gender Identity unclear

Variable 4: Themes Present in Article
Identify 5 maximum, in order of mention.
Count 0 if NONE
1 = Video Game Design
2 = Gameplay
3 = Environment (Gameworld Location)
4 = Racial Identity
5 = Racial Stereotypes
6 = Violence
7 = Hypermasculinity
8 = Hypersexuality
9 = Urban/Street Life
10 = Praising Diversity
11 = Presence of Racialized Language
12 = Lamenting Lack of Diversity

Variable 5: Sentiment toward Protagonist

1= Overall positive toward the protagonist
2 = Neutral toward the protagonist
3 = Overall negative toward the protagonist

Variable 6: Topics
Count the number of words/phrases per category
Count 0 if NONE
6a: References to Character's Physical Appearance
6b: References to Racial Identity
6c: References to Racial Stereotyping
6d: References to Diverse Representation
6e: References to Character Personalization
6f: References to Controls
6g: References to the "feel" of the game

Variable 7: Mentioning or presence of racial identity

If yes, describe how race is identified by the reviewer.

Variable 8: Coder Inferred race

If yes, describe how race was inferred during coding.

Variable 9: Mentioning or presence of racial stereotyping

If yes, describe in what sense the reviewer uses racial stereotypes to describe protagonists.

Variable 10: Other racial identities
If yes, describe how race was inferred during coding.

Variable 11: Quotes from Reviews
Include name of reviewer, quote that stood out, and how it relates to race or racial stereotyping.

Appendix C: Sample of Texts Used

Regarding the five samples, I have included links to five randomly generated game reviews used for this study, as well as the opening paragraphs. I have chosen to include brief screenshots of the game reviews with the URL to save space, as some of the game reviews included in the sample of this study are quite long.

1) Sunset (https://www.ign.com/articles/2015/05/21/sunset-review)

The act of housekeeping sounds like a curiously mundane fit for a video game, but in its latest game Sunset, developer Tale of Tales has turned it into a welcome constant in a world of chaos. Sunset is a meditative and unnerving experience, and one I won't forget quickly.

As Angela, a university graduate from Baltimore, Sunset tasks you with completing dull chores in a single penthouse over a series of days. I found these repetitive - the intricacies of the chores are boiled down to the click of a button - but as I quickly learned, that's the point. The less I cared about folding clothes, the more I cared about the world around me.

Sunset is set in a fictional South American city at the beginning of the 1970s. There's a loud revolution going on outside its floor-to-ceiling windows, yet Angela is cocooned in her wealthy trappings, and sirens and gunfire are only ever heard from a distance. Being so removed from the action encourages a sense of isolation and helplessness, echoed in Angela's own inner monologues.

2) Sanity: Aiken's Artifact(https://www.ign.com/articles/2000/09/26/sanity-aikens-artifact)

Ice T's played many roles during his life: new jack, family man, heavy metal superstar, actor, dreamer, and original gangsta. Lately, though, his work tends to get pigeonholed into B-grade classification. But contrary to what the playa haters have to say, the guy's got talent. Like any average putz could've done a fraction of the stuff this artist's managed to find time for in the last decade. Stop frontin' and give the Iceman his props. If Bruce Campbell can get away with working the CD-ROM circuit, than the OG's certainly earned himself a shot at video game immortality right here and now.

And boy what a debut Mr. T (not to be confused with "Whatchoo talkin' bout foo" guy) has made. After disgruntled consumers hounded industry types for the better part of a decade to create a title based on psionic powers, Monolith's finally stepped up to the plate. Using the versatile LithTech engine as a backbone, Seattle's finest has created a one of a kind action/adventure that dares to defy description. Equal parts DreamWeb, Chaos Engine, MageSlayer, and Magic the Gathering, Sanity's a seething cauldron of creativity that's ready to boil over. All right, maybe it borrows a leeeeeetle heavily from cheesy sci-fi action flicks, but you can't refute the product's amazing approach to the world of espers.

Let's not get ahead of ourselves, however. Before we begin, let me introduce you to Cain, a rather large African-American male known for his psionic talents and piss poor attitude. This lucky chum's an agent who's been trained from birth to combat superhuman criminals. Because Cain's got a problem keeping civilian casualties to a minimum, an aggression control chip has been implanted into homeboy's brain that'll shut him down in an instant if mom, pop or little Billy is accidentally whacked. So, like uh, don't kill anyone unintentionally, or at least wait until they've given you a nasty look as a go ahead. And can you guess that this caveat's gonna make for a Rosie O'Donnell sized pain in the butt when the gray matter starts flyin'. Some people just don't know how to get out of harm's way it seems.

3) Metal Gear Solid 5: The Phantom Pain(https://www.ign.com/articles/2015/08/24/metal-gear-solid-5-the-phantom-pain-review)

Editor's Note: Review updated on 10/15/2015 to include impressions of Metal Gear Online.

I had planned it all very carefully. There were way too many guards still looking for me, and with sunrise coming shortly, I had almost no chance of making it out to the nearest safe landing zone with an injured prisoner on my shoulders. But I wouldn't have to. During the night, I planted some C4 on this outpost's radio communication equipment, the anti-aircraft battery, and most importantly, their AA radar. So I took a deep breath, detonated all three at once, called in a chopper, and watched it all unfold. After a short while, my ride swooped in, blasting an APC to bits with a ferocious rocket barrage, and cutting infantry down with heavy machine gun fire as I scrambled from my hiding place to the main courtyard, prisoner in tow. I hopped in with my precious cargo, and then jumped on the side-mounted minigun to keep the newly arrived reinforcements at bay as my chopper smoked and sputtered its way out of the hotzone.

4) Rise of the Tomb Raider(https://www.ign.com/articles/2015/11/09/rise-of-the-tomb-raider-review)

Rise of the Tomb Raider is the most fun I've had with a Lara Croft game since 1996. Its story is full of the right kind of danger and intrigue, its tombs are dastardly, and I was as struck by its huge, romantic environments as I was as a kid playing the original. Although I could have done with a few more puzzles and fewer firefights overall, I enjoyed every rollicking, big-hearted second of it.

After All, You're a Croft

It's broad, Indiana Jones stuff that gallops along at a great clip.

"

Like its predecessor, Rise of the Tomb Raider revels in an ever-so-slightly-sci-fi and ultimately very fun high-concept involving a hunt for an artifact that grants eternal life. It's broad, Indiana Jones stuff that gallops along at a great clip through gloriously over-the-top sequences grounded with a strong emotional throughline. Rise of the Tomb Raider is, at its core, about Lara and her late father, and actress Camilla Luddington's thoughtful performance as Lara sells us on the complicated relationship she has with the ghosts he left behind.

5) Vampyr (https://www.ign.com/articles/2018/06/04/vampyr-review)

With a fresh and genuine take on the familiar supernatural mythology, Vampyr's bold RPG ambition is to tempt you into eating your own quest givers. Beguiling the citizens of London and suffering the consequences of quenching your terrible thirst sets up some big choices that generally pay off, though its combat doesn't quite have the bite needed to force you out of your comfort zone and into the darker, morally gray areas it so clearly wants you to live in.

Where Vampyr sets itself apart is in its excellent recreation of London during the first World War and in the throes of the Spanish Flu epidemic. It's a gloomy, somber city, explorable through snaking alleyways, cobblestone courtyards, dingy sewers, and expansive buildings that combine with the moody string-heavy soundtrack to create a dense, sad atmosphere of a city on the edge.