How personality can predict video game preferences: a data-driven approach

Alexandre Vessereau

MASTER Economics & Psychology

Supervisors: Thomas Renault* and Todd Lubart†

June 13, 2021

*Université Paris-I Panthéon-Sorbonne, 12 Place du Panthéon, 75231 Paris
†Université de Paris, Institut de Psychologie, LaPEA, 71 Avenue Édouard Vaillant, 92100 Boulogne-Billancourt

Table of Contents

Acknowledgements

I would like to express my deepest thanks to my supervisors Mr. Renault and Mr. Lubart, who continuously helped me throughout the writing of this Master's thesis. I would also like to thank some anonymous individuals who helped me to clarify and make the experiment as understandable as possible. Finally, I would like to thank all my friends and family who supported me throughout this year.

Abstract

This study aims to investigate the link between personality traits and video game genre preferences. To do so, we first built a new nomenclature of video game genres, using topic modeling tools and a corpus of millions of video game reviews written by millions of unique players. In total, 7 relevant topics were extracted, each capturing a singular dimension of video games. We then conducted an online experiment to determine how personality traits could relate to individuals' preferences towards these topics. The results showed that individuals with high scores in Openness To Experience were more likely to favor games containing strategy and simulation elements, whilst high Conscientious scores positively predicted to prefer games with action-oriented features. Findings also showed that individuals with high Neuroticism scores were more likely to avoid games with puzzle components, but would favor games with role-playing and multiplayer aspects. We also showed that highly Extroverted individuals were more likely to be attracted to games with multiplayer and violent/shooting elements. With respect to our findings, this study helps to demonstrate that segmenting consumers based on their personality traits is a viable strategy, as each personality trait relates to specific preferences towards particular video game elements. Such an approach for firms would allow them to be more competitive and maximize their sales as well as to please their personality-segmented consumers, by advertising adequate and relevant video games.

Keywords: Market Segmentation, Psychographic consumer segmentation, Five-Factor Model, Personality, Video games, Games genre classification, Consumer preference

1 Introduction

One of the constant challenges for firms is to understand what factors drive the preferences of their consumers. If such a link is found, then firms can segment their market according to this factor, in order to optimize their resources on the one hand, and on the other hand, to advertise products likely to be preferred by consumers (McCarthy, 1960). Along with demographic or cultural factors, more latent factors such as personality can be used to determine whether it influences products preferences (Wedel & Kamakura, 2000). Indeed, studies have shown that personality traits can influence many situations in life (Nyhus and Pons, 2005; Weber, 2015) and that they play a role in shaping consumers' preferences on goods, especially cultural goods, such as fine arts (Pitt et al., 2020), cinema (Karumur et al., 2016) or video games (Braun et al., 2016).

In the recent years, this lucrative industry has become one of the most important cultural sectors for the youth, with a market value exceeding that of the sports or film industry (Monahan, 2021). Given those figures, it is logical and relevant to examine whether personality can shape video game preferences, in order to possibly segment consumers according to it. Several studies (Braun et al., 2016; Peever et al., 2012; Potard et al., 2019; Worth and Book, 2015) have already tried to examine whether such a link could exist, by linking personality traits to specific game genres (action, strategy, simulation...). However, all these studies have chosen, to the best of our knowledge, to use genres based on game genre classification manually constructed (Apperley, 2006; Wolf, 2001). In other words, it implies that those genres used were imposed and chosen manually, by one or a few authors, according to specific video games characteristics (e.g.: visual, interactivity). This approach is somewhat problematic, as these imposed genres may not correspond to those imagined by the players themselves, thus creating a discrepancy between what the authors and the players deem relevant in video games. Generally speaking, it is better for academics, and for firms, to use a vocabulary closer to that formulated by video gamers (Clarke et al., 2015). Failure to do so could lead to misunderstandings on the part of the players and thus possibly to biases in the results found (Clarke et al., 2015).

This paper aims at tackling this issue. It is using millions of different reviews, written by millions of unique individuals, and a topic modeling approach to build a classification of game genres closer to those of players, because it is indirectly produced by them. This classification is then used to examine how personality traits can shape preferences towards those new game genres. Therefore, this study contributes both to building a new classification and to offer substantial insights regarding the link between personality traits and video games. In addition to contributing to the burgeoning literature regarding video games, our findings could interest video game firms that want to segment their market according to personality traits, in order to better offer them relevant advertisements and products.

This paper is structured as follows. First, we review how implementing market segmentation is relevant and crucial for firms that want to maximize their profits, and explain how and why personality could be used to segment consumers. We also highlight the relevance of studying the video game market and emphasize the importance of building a proper video game nomenclature. Second, we describe how we extracted and built a new taxonomy of the video game genre, using data-driven and topic-modelling methods. We outline subsequently how we used those new game topics in an online experiment to investigate whether personality traits would influence preference towards those topics. To conclude, we discuss how the results obtained can contribute to giving new insights on a theoretical and practical level, while also acknowledging the limitations and suggesting avenues for future research.

2 Literature review

2.1 Market Segmentation

Market segmentation is one of the tools that companies can rely on to divide individuals into groups that share the same consumption patterns (C.-F. Lin, 2002). First introduced by Smith (1956), *market segmentation* can be defined as "viewing a heterogeneous market as a number of smaller homogeneous markets, in response to differing preferences... attributable to the desires of consumers for more precise satisfaction of their varying wants" (p. 6). As preferences can dramatically vary from one consumer to another, it is paramount for firms to implement proper market segmentation. It allows these companies to attribute their limited resources most efficiently through the understanding of which variables influence consumers' behavior (Liu et al., 2019). Accordingly, market segmentation helps companies to maximize their sales and be more competitive (Hunt & Arnett, 2004), by developing marketing strategies tailored to their consumers' profiles (McCarthy, 1960).

According to the parameter the company wants to focus on, this segmentation of the market can be enforced differently. As such, this segmentation can be based on observable variables such as demographics (Blattberg et al., 1976), usage frequency (Twedt, 1964; Young et al., 1978), brand loyalty (Grover & Srinivasan, 1989) or latent[1] factors (i.e., that needs to be inferred) such as personality (Brody & Cunningham, 1968), values (Kamakura & Mazzon, 1991) or lifestyles (Yankelovich, 1964). Marketing academics often group these last set of variables under the umbrella of *Psychographics* (a contraction between psychology and demographics) variables, and is one of the main bases on which consumers can be segmented. This approach is driven by the idea that similar social groups (demographics, cultural) can expose vastly different psychographic profiles (C.-F. Lin, 2002).

2.2 Defining personality

The interest of companies in psychographic data, and more specifically in personality, is not without reason. Since Allport (1938), personality has been considered as an aggregation of traits (e.g., shyness, anxiousness, honesty...) that helps to define the uniqueness of one's identity (Larsen & Buss, 2017), while influencing the person's behavior (Almlund et al., 2011) and shaping preferences (Pitt et al., 2020). As demonstrated by numerous psychological experiments, personality has been found to influence leadership dispositions (Judge et al., 2002), student performance (Weber, 2015) and political orientation (Furnham & Fenton-O'Creevy, 2018). Economics scholars also sought to in-

[1] This classification originates from Wedel and Kamakura (2000), who classify market segmentation in 2 (General vs. Product-Specific) x 2 (Observable vs. Unobservable) categories. For an extensive review on market segmentation, see Wedel and Kamakura (2000).

vestigate the possible relationship between personality and various economic outcomes. Such studies have shown that personality can influence areas such as income increase (Boyce & Wood, 2011), behavior in ultimatum games (Swope et al., 2008) or wage determination (Nyhus & Pons, 2005).

2.2.1 The Five-Factor model

One of the main models developed to represent human personality is the Five-Factor[2] model (henceforth, FFM), more commonly known as the "Big Five Factor"[3] model. As the consequence of several decades of study and work (Allport and Oldbert, 1938; Cattell, 1945; Goldberg, 1990; Norman, 1963), the FFM suggests that personality can be decomposed into five factors (Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to experience), each of which contains fine-tuned personality traits (also called "facets").

Each factor measures a characteristic adaptation and a specific motivational system. For instance, Extroverted individuals tend to favor stimulation (Aluja et al., 2003), social interaction, rewards and attention (Lucas et al., 2000). In contrast, Neurotic individuals are more affected by uncertainty and are more susceptible to exhibit pessimism and low self-esteem, and favor safe and secure behaviors (Carver et al., 2000). Conscientious individuals can be characterized by their appeal to self-discipline, strong organization, success, achievement, and analytical thinking (Roberts et al., 2005), whilst Agreeable individuals will be interested in altruism, cooperation and in their connection to others (Graziano & Eisenberg, 1997). Openness to experience, as its name suggests, implies individuals who are sensitive to novelty, creativity, intellectual stimulation and are more open-minded[4] (McCrae & Costa, 1997).

To assess personality through the Five-Factor model, numerous scales have been developed, with the most popular being the NEO Personality Inventory (Costa & McCrae,

[2] While other models which evaluate personality do exist (Myers–Briggs Type Indicator (Myers, 1962), HEXACO Personality Inventory (K. Lee & Ashton, 2004)), the FFM remains one of the most dominant model of personality (Giluk & Postlethwaite, 2015). For an extensive review of existing personality models, see John et al. (2008).

[3] A term coined by Goldberg (1981).

[4] See Costa and McCrae (2008) for a review on the theoretical basis of those factors.

2008), the International Personality Item Pool (Goldberg et al., 2006) or the Big-Five Inventory (John et al., 2008). These scales mostly rely on items composed of self-describing adjectives or short phrases (e.g., *I see myself as someone who is talkative*), for which subjects have to report how much they relate to them. This approach, however, has several issues. First, self-reported questionnaires are often subject to social desirability bias, in which the subject might answer differently in order to present him or herself in a favorable light (Mcdonald, 2008). Second, and perhaps the most critical point, lies in the length of those scales: self-report questionnaires are often time-consuming and tedious to complete for individuals (Liu et al., 2019). This issue often results in a low response rate (Pitt et al., 2020) and therefore complicates data collection, for scholars and firms alike. Despite those limitations, the FFM has proven useful in explaining a wide variety of behaviors, such as predicting subjective well-being (DeNeve & Cooper, 1998), academic dishonesty (Giluk & Postlethwaite, 2015), job performance (Shaffer & Postlethwaite, 2012) or even risk aversion (Rustichini et al., 2016).

The FFM has also been found to be relevant in consumer behavior studies. For instance, it has been shown that advertisements tailored to the personality traits of the consumers increase their effectiveness (Hirsh et al., 2012). Besides, in addition to its use in predicting consumer preferences in cultural goods such as music (Fernàndez-Tobías et al., 2016), movies (Karumur et al., 2016) and fine arts (Pitt et al., 2020), the FFM has been particularly used to investigate how personality traits can shape video games preferences. Despite being a perceived cultural good aimed mainly at teenagers (Braun et al., 2016), video games are, in fact, a strong and very lucrative sector, with estimated revenues of $180 billion (Monahan, 2021), and a consumer base of over 2.7 billion players in 2020 (Wijman, 2020). Given these figures, it is only natural and relevant to focus on the link between personality traits and video games, both from a psychological and an economic standpoint.

2.3 Personality traits and games genre preferences

While the first "video game" *Tennis for Two* was merely an experiment developed by Willy Higinbotham playable on an oscilloscope (Wolf, 2007), this medium has since certainly evolved and has, because of its important presence in the current cultural horizon, been increasingly studied by researchers. Specifically, academics have investigated the possible link between personality traits[5] of the FFM and games genre preferences.

As such, findings show that high scores in Extraversion led people to favor action-oriented games (Braun et al., 2016), casual and easy games (Potard et al., 2019), massive multiplayer online games, but to avoid games with strategy (Peever et al., 2012) and role-playing (Potard et al., 2019) features. Other studies found that Openness to Experience was positively associated with role-playing (Braun et al., 2016; Potard et al., 2019) and violent games (Peever et al., 2012), whereas Conscientious individuals preferred simulation/creative (Worth & Book, 2015) and fighting games (Peever et al., 2012). It was also showed that highly Neurotic individuals exhibited preferences towards role-playing (Braun et al., 2016), multiplayer (Potard et al., 2019) and violent games (Chory & Goodboy, 2011). Regarding Agreeableness, it was found that this trait was positively associated with helping behaviors in online games (Worth & Book, 2015), but negatively related to violent games (Chory & Goodboy, 2011).

2.3.1 The complex task of labeling games

Whereas some of the studies cited above based their work on some components of video games (e.g., *violent*, *creative*, *helping behaviors*), others preferred to rely on pre-existing classifications[6] of video game genres to conduct their research. For instance, Braun et al. (2016) used Apperley's (2006) classification, which divided video games into 4 categories (*Action*, *Simulation*, *Role-Play* and *Strategy*) according to their ludological and esthetic characteristics. Another well-known classification is Wolf's (2001)

[5] Besides personality, some studies have focused on the negative influence of video games such as aggressive behaviors (Greitemeyer & Mügge, 2014) and addiction (Anderson et al., 2010), whereas other research has shown that playing video games could positively affect cognitive abilities (Green & Bavelier, 2012), alleviate anxiety (Russoniello et al., 2009) and promote pro-social behaviors (Gentile et al., 2009).

[6] For a review of existing game genres classifications, see Clarke et al. (2015).

exhaustive taxonomy, which discriminated games into 43 categories, based on the degree of players' interactivity with the game. Proposing a proper nomenclature for game genres is meaningful, as it allows consumers to identify the product quickly and help them find similar products.

However, if the intent to propose a suitable classification for video games is laudable, those nomenclatures might fail to capture all the complexity inherent to this specific medium (Clarke et al., 2015). Moreover, and perhaps most problematically, is that most of those existing classifications emerge more from scholars, than from the players themselves (Faisal & Peltoniemi, 2015). Given this fact, several limitations can be outlined. On the one hand, genres are quite unstable: new labels gradually emerge as the gaming community itself begins to differentiate games according to certain characteristics (Clarke et al., 2015). On the other hand, games are almost always a mixture of genres, and are rarely uni-dimensional. For instance, both *Call of Duty*® and *Mario*® could be categorized as action games. However, the first one is more of an action/shooter game that involves killing people, whereas the second consists more of an action/adventure game, in which the goal is to avoid obstacles and enemies. Therefore, because classifications are "fixed" when established by their authors, they fail to be flexible and to take into account those player-created and multi-dimensional genres (Clarke et al., 2015). Consequently, those mismatches between players' and external actors' perspectives of game genres create a potential disconnect, which could be critical if firms label their games in a way that is not or poorly recognized by their consumers.

Accordingly, an alternative strategy to those "top-down" nomenclatures by external actors would be to use a "bottom-up" approach to extract what the main actors, namely video game players, would say about their games. Such bottom-up, data-driven approach, has already yielded interesting results: using a co-occurrence analysis of game tags generated by game players, Li (2020) extracted 4 main genres: *Simulation & Strategy*, *Puzzle & Arcade*, *Action* and *Role-Play*. Similarly, Faisal and Peltoniemi (2015) used a Latent Dirichlet Allocation (LDA) topic-modeling approach based on textual descriptions from Mobygames.com, a wiki-based website on which users can collaborate and edit games descriptions freely. Their database contained games published between 1979 and 2010, from which they extracted 31 different topics/genres. One of the great strengths of data-driven approaches is that they allow the genres to be contextualized in the players' own words, therefore not being imposed by external actors (Faisal & Peltoniemi, 2015).

As a matter of fact, proposing a suitable taxonomy is valuable, especially in a cultural sector like video games. As demonstrated by J. Lee et al. (2015), 74% of gamers use the genre of a game to buy it or to discover a new one, thus stressing the importance of correctly labeling the game. Given this fact, one of the strategies often put in place by game companies is to assign any label that can be more or less related to their game (Clarke et al., 2015). Such an approach seems justified at first sight: more labels would definitely reach more consumers. However, and as pointed out by Clarke et al. (2015), the overuse of labels obscures the game's own identity and further damages the marketing campaign by confusing the players themselves on the true nature of the game being advertised. While such a strategy could pay off in a large-scale marketing campaign, it could fail and be detrimental if companies decide to advertise games in a "language" (i.e., labels) not recognized by the target segments (Bell & David, 2008).

3 Aim of the present paper

The goal of this study is twofold. First, we aim to create a new nomenclature of video game genres via topic modeling, in the same vein as Faisal and Peltoniemi (2015). The key difference between Faisal and Peltoniemi's (2015) study and ours lies in the material used to create topics. Whereas the authors used textual descriptions from MobyGames.com to construct their topics, we will use the words contained in reviews written by players. This approach allows us to create topics whose words used by players can be much more diverse and tinged with emotions and feelings than those of a factual description from a wiki. In this respect, we expect to get a classification closer to what an average player might imagine and formulate.

Second, we plan to use this new nomenclature to link personality traits with preferences for certain types of games. In comparison to studies that have attempted to do such a thing, we would not ask if users prefer a particular game genre per se, but rather a list of words that conceptualize an idea of a game. Furthermore, it would allow game genres to be multi-dimensional and thus, to be more nuanced. Although our study is exploratory by design, we can nonetheless draw on the previous research, as well as the five traits characteristics, to formulate our hypotheses:

H1: As highly Extraverted individuals are stimulation-seekers (Aluja et al., 2003), they will prefer games with strong action components (fighting, shooting games).

H2: Given that Neurotic individuals are often characterized by their low self-esteem and negative feelings (Carver et al., 2000), they might prefer games in which they can "escape" those negative emotions by playing a different character in a virtual world. Thus, we predict that highly Neurotic subjects will favor games in which such a thing is possible, such as games with role-play components.

H3: As previously stated, individuals with a high Openness To Experience trait are attracted to novelty and creativity (McCrae & Costa, 1997). These two characteristics define games with role-playing elements because, as previously stated, they allow players to be immersed in an imaginary world, letting them be creative on how they want to shape their virtual character (e.g., behavior, mindset, physical appearance). Therefore, we hypothesize that individuals with high scores in Openness To Experience will prefer games with role-play elements.

H4: We hypothesize that highly Conscientious individuals will prefer games with puzzle elements, as these games inherently require a high level of thinking, which is a core element of the Conscientious personality trait (Roberts et al., 2005).

H5: Finally, because Agreeable individuals are characterized by their pro-social behavior and their willingness to help others (Graziano & Eisenberg, 1997), we predict that they will prefer multiplayer online games letting them express this behavior.

Ultimately, the main contribution of this paper is to shed more light on how personality traits can be related to video game genres, derived from a new taxonomy originating from players themselves. This whole approach could be of great interest for firms that want to segment their market according to personality variables, using a vocabulary in their advertisements more likely to please those segments.

4 Experimental design

The experimental design was conceived in two stages. First, we extracted the necessary data in the form of reviews, which were then pre-processed and filtered according to specific criteria. After establishing the adequate number of topics, we used our dataset to build the main genres/topics (these terms are used interchangeably). Once these topics were uncovered, they were used as material for an online experiment. A workflow summarizing the entire process can be found in Figure 4, in the Appendix.

4.1 Data preparation

4.1.1 Extracting Steam Reviews

The first step was to extract the reviews from the Steam website, which is one of the largest distribution platforms of digital video games globally with 120 million monthly active users alone in 2021 (Clement, 2021b). Steam serves multiple purposes: besides selling games, it acts as a digital game library and as a social platform, on which users can share messages and invite each other to play together. One of the most interesting features of Steam is that it allows users to review games they own in the form of a "Recommended" or "Not recommended", which needs to be accompanied by a textual content written by the user.

To collect the Steam reviews, we used a set of Steam-Crawler scripts[7] (Esuli, 2018), that we modified through the implementation of some structural changes[8]. We fetched virtually every English review that could be found for each game available on Steam using the Steam API. The reviews used were published over an 11-year period, from 16 October 2010 to 5 January 2021. This process was conducted within the conditions and

[7] A crawler is a small program that navigates through a website and collects the data required.
[8] The scripts used to extract the reviews can be found on our GitHub repository: https://github.com/Arkhemis/steam-crawler.

terms of the Steam API[9] and the data collected was limited to information made public by Steam. In total, we recorded 12,758,976 reviews[10], written by 5,868,265 unique users and distributed over 32,788 different games. Each review collected was supplemented with other relevant metadata associated with information regarding its author and the game reviewed, as can be seen on the sample review[11] in Table 1 below.

Table 1: Two review examples from the reviews dataset

game_Name	game_n_review	user_ID	user_n_review	user_playtime_hours	user_review_text	user_recommended
Zengeon	97	765611980	57	0.2	I feel furious because this game looked amazing, anime+roguelike elements and when I played it... it was so boring	False
Terraria	291903	765611981	10	121.8	If you want to pass the time, Terraria is the best game in the world to do that! I highly recommend you get this game!	True

4.1.2 Pre-processing and filtering reviews

The second step aimed to pre-process the data, that is, to "clean" the reviews and then filter some of them out of our dataset according to several criteria. First, we tokenized all reviews, which is the process of breaking down a piece of text into small units called tokens. This step is needed if one wants to operate machine learning methods or, in our case, Natural Language Processing methods (NLP), on text content. Then, we proceeded to turn all reviews in lowercase and to remove non-alphabetic characters (punctuation marks, numbers, non-Latin characters), as well as "stop-words", which are very common words (e.g.: "*the*", "*at*", "*a*"...) that can be deleted without sacrificing the meaning of a sentence. In addition, we ran a lemmatization process using the Spacy Python library (Honnibal & Montani, 2017), which is a technique used to revert a word to its base form by deleting its suffix (e.g.: the word "looked" is lemmatized to "look"). An illustration of these pre-processing methods can be found on the Figure 1 below, by using the first user's text of Table 14 as an example. We also used the LangDetect library (Nakatani, 2010) to identify and discard non-English reviews that could have been misidentified and fetched by our crawler above.

[9] https://steamcommunity.com/dev/apiterms

[10] The corresponding dataset can be found here: https://archive.org/details/steamreview_dataset.

[11] Only the most relevant data was included on Table 1. To see all the metadata collected, please refer to Table 14 in Appendix.

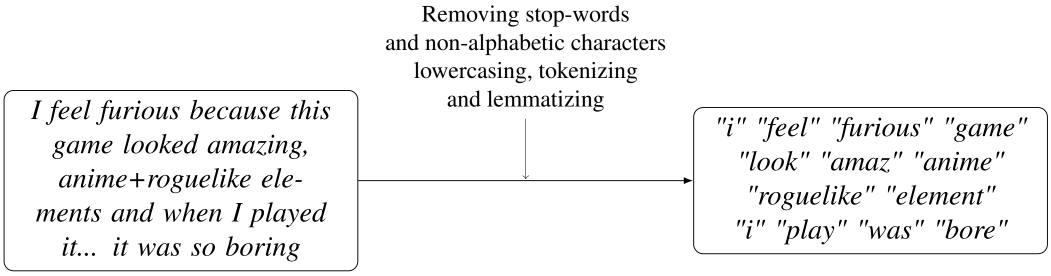


Figure 1: An illustration of pre-processing treatment

We then removed reviews for a given game for which users had a very low game playtime (less than 2 hours) when they published their review. We set this minimum requirement as some online gaming platforms (e.g., Nintendo) require this amount of gaming time before allowing users to post reviews (D. Lin et al., 2019). Very short reviews (less than 5 tokens/words) were also removed, as the tool we will introduce next does not perform well with small pieces of texts (Yan et al., 2013). Filtering according to these criteria reduced our review dataset to 4,892,472 reviews (2,793,097 unique users and 1095 different games), accounting for 38% of our original dataset.

4.1.3 Topic modeling

To group words used in the reviews into topics, we used Latent Dirichlet Allocation (henceforth, LDA) modeling (Blei et al., 2003). LDA modeling is an unsupervised learning method that identifies hidden ("latent") relationships between a set of documents (or a set of reviews, in our case) through the pattern of words they contain, using a bag-of-words approach (i.e., the order of words does not matter). First, LDA modeling assumes that documents are a combination of one or more topics. Second, it assumes that within one topic, certain words will be weighted differently, as they would be used much more frequently than others and be most likely to co-occur along with other words (Blei et al., 2003). Thus, using a probabilistic framework, a LDA model can find as many topics embedded in a set of documents requested by the user.

However, as the number of topics specified by the user increases, the coherence score of the LDA model —which represents how the model performs to create human-interpretable topics— will inevitably decrease. As a matter of fact, the algorithm will always find links between words to satisfy the number of topics required, even if those links seem incoherent from a human perspective. Consequently, one needs to compute the coherence score of the model for the optimal number of topics to use, which corresponds to the least number of topics before the coherence score starts to decrease.

Preparing the data for the number of topics to be used, we first constructed bi-grams and tri-grams, which are sequences of two and three adjacent tokens respectively. As was previously said, LDA modeling looks for the co-occurrence of words to construct topics, and bi-grams and tri-grams can be used alongside single tokens for this purpose. We also filtered frequent words (appearing more than 20% of the time). Using the LDA MALLET implementation (McCallum, 2002) which outperforms traditional LDA modeling, we found that 14 (with a coherence score of 48%) was the optimal number of topics considering our review dataset, as shown in Figure 2.

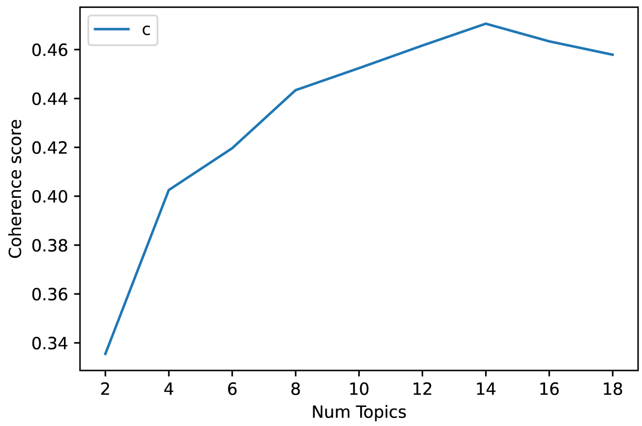


Figure 2: Coherence score as a function of the number of topics

After identifying the number of topic needed, we trained the model with 14 topics and using the 4,892,472 reviews of our dataset. Figure 3[12] yields a visual representation of topics and their relationships (semantically related topics are closer to each other), whereas Table 2 shows the 10 most important words per topic.

[12] The interactive version of the graph can be found at this address: https://bit.ly/3ul3TCN

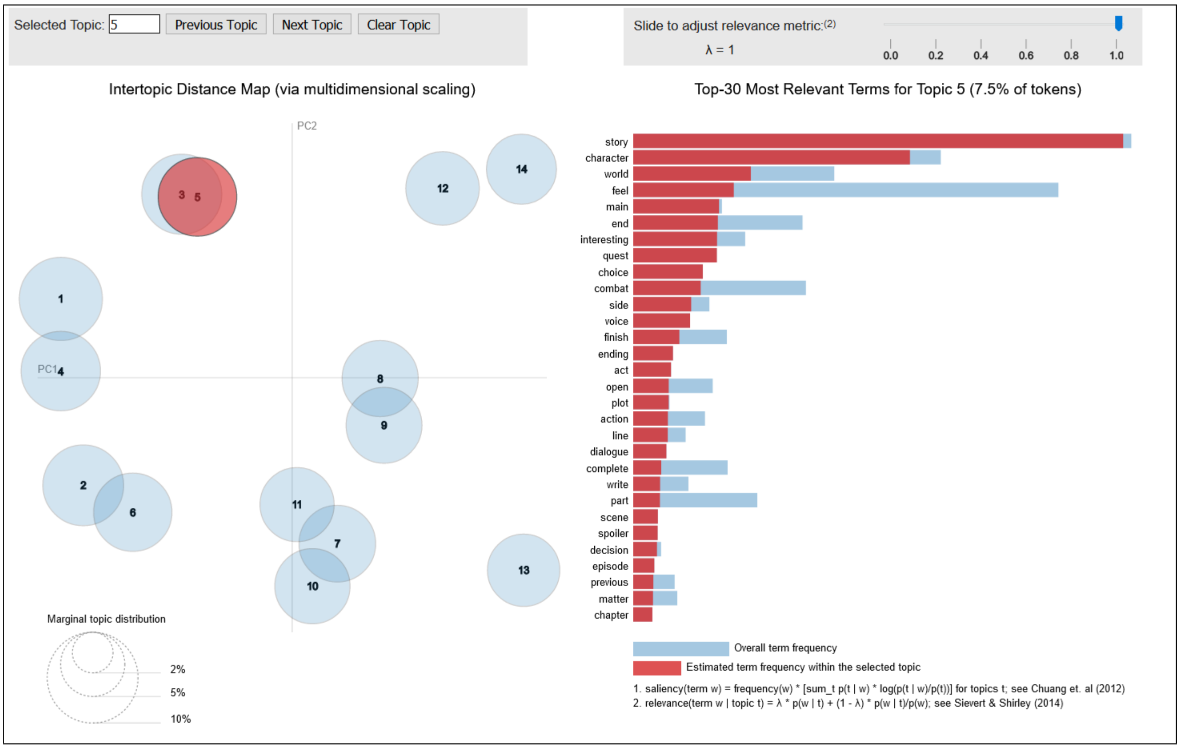


Figure 3: Intertopic distance and most relevant words associated with each topic, here Topic n°5

As depicted in Table 2 below, we looked for topics whose words could be linked to video game elements. On the one hand, some topics, such as Topic n°7, 8, 9, 13, and 14 contained words too related to video game consumption or technical issues, rather than to video games themselves and thus were discarded. On the other hand, Topics n°10 and 12 were more ambiguous. We decided to discard Topic n°10 as its most salient words (*"hell"*, *"doom"*) could directly remind a video game franchise (namely Doom®). As for Topic n°12, it contained too many opposed words (e.g: *"short"* vs. *"long"*, *"hard"* vs. *"easy"*) to really considering it to be related to video games. Consequently, we only kept topics containing salient words related to games, namely Topic n°1 (∼Challenging[13] games); 2 (∼Shooter games); 3 (∼Puzzle games); 4 (∼Strategy games); 5 (∼Role-playing games); 6 (∼Simulation games) and 11 (∼Multiplayer games).

[13] Those topic names were chosen to correspond as closely as possible, from a semantic perspective, to the most salient words of a topic.

Table 2: Top-10 salient words per topics and their relevance

Topic n°	Top-10 most salient words	Can be related to a video game?
1	"level", "enemy", "fight", "combat", "boss", "weapon", "skill", "system", "item", "ability"	Yes
2	"mission", "gun", "kill", "shoot", "enemy", "weapon", "car", "hit", "jump", "run"	Yes
3	"experience", "puzzle", "design", "feel", "sound", "atmosphere", "horror", "visual", "find", "environment", "mechanic"	Yes
4	"war", "battle", "space", "strategy", "system", "artificial intelligence", "ship", "turn", "build", "base"	Yes
5	"story", "character", "world", "feel", "main", "end", "interesting", "quest", "choice", "combat"	Yes
6	"find", "build", "start", "survival", "thing", "world", "explore", "day", "craft", "building"	Yes
7	"bug", "run", "issue", "fix", "work", "problem", "ca", "crash", "save", "setting"	No
8	"thing", "lot", "feel", "pretty", "bit", "bad", "give", "cool", "start", "stuff"	No
9	"update", "review", "add", "content", "early", "release", "developer", "hope", "access", "work"	No
10	"life", "back", "guy", "hell", "man", "doom", "die", "cry", "watch", "kill"	Maybe
11	"player", "friend", "people", "mode", "map" , "server", "team", "community", "single", "coop"	Yes
12	"hard", "easy", "long", "learn", "challenge", "hour", "music", "short", "beautiful", "enjoy"	Maybe
13	"buy", "hour", "worth", "money", "price", "wait", "sale", "spend", "year", "stream"	No
14	"love", "recommend", "amazing", "graphic", "awesome", "enjoy", "fan", "highly", "original", "series"	No

4.2 Main experiment

Having collected the material required, we designed an experiment[14] using oTree (Chen et al., 2016), a Python framework used for online experiments. The study was divided into four parts, had a duration of approximately 15 minutes, and was designed in both English and French.

4.2.1 Consent and introduction

The first page of the experiment consisted of a consent form. Subjects could consent to have their responses collected and analyzed anonymously (and thus continue the study) or refuse (and leave the study). If subjects gave their consent, they arrived on the

[14] The entire protocol can be found in the Appendix C, whilst the experiment is available at this address: https://otree-games-study.herokuapp.com/room/Experiment_EN/.

introduction page, explaining the experiment's purpose (preferences in terms of video games) and its duration. They were also told that if they completed the survey, they would receive a unique code to enter a prize draw to win either an Amazon or a Steam gift-card, both worth €60.

4.2.2 Part I: Measuring Willingness-To-Play

The first part of the experiment consisted of asking subjects if they would like to play a game, described by a list of 10 words. Each list contained the 10-top salient words from the 7 topics (Challenging, Shooter, Puzzle, Strategy, Role-playing, Survival and Multiplayer games) extracted from Table 2, and were each displayed on 7 different pages, whose order of presentation we randomized to avoid any order bias. For each list by page, subjects were asked to imagine that a game could be defined according to the words presented and were invited to give their Willingness-To-Play (WTPlay) this game, on a 7-point Likert-Scale ("*Not at all*", "*A little bit*", "*A bit*", "*Maybe*", "*Likely*", "*Very likely*", "*Definitely*"). This approach allowed us to gain an initial insight into participants' video game preferences, although subjects could set the same WTPlay to multiple lists, thus allowing for preference indifference.

4.2.3 Part II: Assessing subjects' personality

In the second part of our experiment, we sought to assess the personality of the participants. To this end, we used the Big-Five-Inventory[15] (BFI; John et al., 2008)[16], which is a self-report scale composed of 44 short sentences describing precise behaviors, each linked to a precise personality factor (Extraversion, Openness to Experience, Agreeableness, Neuroticism, Consciousness). For each behavior, subjects were asked if they identified with it on a 5-point Likert-Scale ("*Disagree strongly*", "*Disagree a little*", "*Neither agree nor disagree*", "*Agree a little*", "*Agree Strongly*").

[15] Because we were concerned by drop-outs issues in our online experiment, we did consider shorter scales, like the short BFI-10 items version of Rammstedt and John (2007). Yet, according to John et al. (2008), the trade-offs in the lost of reliability is too important if the main focus is the measure of the personality, which is why we ultimately chose the longer version of the BFI here.

[16] We used Plaisant et al.'s (2010) translation for the French BFI counterpart.

4.2.4 Part III: Measuring Willingness-To-Spend-Time

After evaluating personality, subjects arrived on a new page that displayed the 7 lists of 10-words used in Part-I. We asked subjects to imagine that they had to devote some time between 0 and 120 minutes, as they saw fit, to each game characterized by the words of each list. However, it was stressed that they had to assign different values for each of the 7 lists: subjects could not proceed further (the Next button was hidden) if this condition was not met. This step allowed us to measure the Willingness-To-Spend-Time (WTST) of subjects and offered an alternative to WTPlay measures by providing a distinctive ranking in terms of video game genre preferences.

After the allocation phase was over, participants were invited to justify why they spent their minutes in this manner. Subjects could either respond that they preferred to spend time on known games genres or because they wanted to explore games they know little or nothing about. Moreover, subjects could indicate that they found it difficult to allocate their minutes on the lists or could justify in their own words their allocation ("Other" option).

4.2.5 Exit survey

Once subjects completed the previous task, they arrived on a new page, which comprised questions regarding their video game behavior. More specifically, subjects were asked if they were video game players. If so, subjects were asked to indicate the gaming platform on which they played the most ("*Computer*", "*PlayStation*", "*Nintendo*", "*Xbox*", "*Handheld consoles*", "*Old consoles*", "*Other*"). Participants could also (although it was optional), give their second and third most-played platform. Besides, subjects were given the opportunity (also optional) to specify their favorite video game and to justify it in a couple of lines. If subjects were not video game users, they could simply leave all the fields blank and proceed to the next page.

On the penultimate page of our experiment, subjects had to answer an exit questionnaire. We asked participants to indicate their age, sex, location, field of study, socio-professional category and educational attainment. We also asked subjects whether they wanted to participate in the prize draw (to win either an Amazon or Steam gift card worth €60). If so, the following page included the code required to enter the draw. On this last page, each subject was thanked for participating in the experiment and invited to exit the survey.

5 Results

The results section is divided into two parts. First, we exposed descriptive information concerning the sample collected, followed by correlations between the different personalities and other variables of interest. Then, we investigated whether personality traits had a specific way of ranking games and used logistic models to find whether personality could predict the preference for a given game topic. We used Python and R software to conduct our statistical analysis[17].

5.1 Descriptive analysis

One of the main challenges of online experiments is keeping participant's presence throughout it: results obtained here show that this is a difficult task. Out of the 434 individuals who arrived on the Consent page, 150 finished the experiment, or 34.5%. Participants were recruited through multiple social websites such as Reddit or LinkedIn. The sample was composed of 42 females and 108 males, with an average age of 28.3 years old (Median = 25, S.D = 9.42). Additional demographics information concerning the field of study/work, the socio-professional category, the language used in the experiment by participants (English or French) as well as information about being a video game player or not is given in Table 3 below.

[17] All the scripts used can be found in the corresponding repository: https://github.com/Arkhemis/M2_EcoPsycho_Thesis

Table 3: Demographic information of the sample ($N = 150$)

Demographics		N	Frequency
Gender	Female	42	28%
	Male	108	72%
Age	18 - 24	61	41%
	25 - 34	63	42%
	35 - 44	16	10.5%
	45 - 54	3	2%
	55 & over	7	4.5%
Field of study/work:	I.T	35	23.3%
	Sciences	24	16%
	Social Sciences	23	15.3%
	Economics	12	8%
	Mathematics	12	8%
	Sales Agent	8	5.3%
	Law	6	4%
	Foreign languages	6	4%
	Others	24	16%
Socio-professional category:	Student	58	38.7%
	Executive	33	22%
	Freelance	10	6.7%
	Teacher	9	6%
	Blue collar	7	4.7%
	Others	33	21.9%
Video game player:	Yes	123	82%
	No	27	18%
Language used	French	96	64%
	English	54	36%

5.1.1 Personality traits distribution

We first calculated the internal reliability for each of the five traits and the proportion of subjects based on their predominant personality. To find the predominant personality, we followed Rose et al.'s (2002) approach by adding up the item loadings scores for a given trait and then divided the sum into quartiles. A personality trait was said to be predominant if the sum of its items loadings lied in the third (upper) quartile (so that 75% of the data lied below this score). As personality is multi-dimensional, such a process allows for individuals to have possibly two or more main personality traits. Regarding reliability, results on Table 13 showed that the BFI scale yielded high internal reliability across all five traits with Cronbach's α ranging from 0.72 to 0.88, thus being in line with previous findings (John et al., 2008).

In addition, we computed the Pearson correlation between personality trait scores and Age, gender and video game use. As shown in Table 13, only the Age ($r = -0.26$, $p < 0.001$) and being a Female subject ($r = -0.44$, $p < 0.001$) had a significant negative correlation with being a video game player. Additionally, being a Female was significantly related to Age ($r = 0.24$, $p < 0.001$), and positively correlated with each personality trait, with the exception of the Extraversion trait. Furthermore, almost all personality variables were correlated with each other. However, even though several of our variables of interest did relate to each other, the degree of correlation remained low to moderate.

Table 4: Internal reliability for personality trait measures and Pearson Correlation between Age, gender, video game player and the five personality traits scores

	M	SD	1	2	3	4	5	6	7	8
1. Age	28.3	9.42	-							
2. Gamer	0.82	0.38	-0.26***	-						
3. Gender	0.28	0.45	0.24***	-0.44***	-					
4. Openness ($N = 50$)	4	0.59	0.06	-0.03	0.18***	**0.72**				
5. Agreeableness ($N = 41$)	4.5	0.65	0.14	0	0.18***	0.26***	**0.76**			
6. Conscientiousness ($N = 44$)	4.1	0.66	0.16	-0.15	0.16***	0.18**	0.31***	**0.80**		
7. Neuroticism ($N = 41$)	3.7	0.88	0	0.04	0.15**	-0.07	-0.25***	-0.24***	**0.87**	
8. Extraversion ($N = 43$)	3.6	0.89	-0.01	-0.05	0.08	0.25***	0.23***	0.17**	-0.22***	**0.88**

Notes: Being a gamer was coded 1 if True, 0 else. Similarly, Gender was a binary variable coded 1 if subject was a female, 0 else. Regarding personality trait distribution, it was said to be predominant for a given subject if the sum of its item loadings scores lied into the upper quartile. Given this fact, it allowed individuals to possess one or more predominant personality traits.
Cronbach's alpha for each personality trait internal reliability are reported in bold.
*** $p < 0.01$, ** $p < 0.05$
$N = 150$, M = Mean, SD= Standard Deviation

5.2 Video games preferences by personality

To evaluate which personality traits predicted the preference for a game topic, we relied on the Willingness-To-Play (WTPlay) and Willingness-To-Spend-Time (WTST). The two measures had a highly convergent validity, as demonstrated by the strong significant positive Pearson's correlation for each game topic ($p < 0.05$)[18] between the two. As such, the next section is divided into two parts. First, we examined for WTPlay whether the ranking system was significantly different for each personality trait, and then investigated which personality traits could predict a preference for a game topic.

[18] Because Pearson's correlation can be influenced by non-normality and because it is less suitable for ordinal data, we also computed Spearman's correlations between the two measures. Findings indicated that the two measures were highly correlated ($p < 0.05$) for each game topic ranking. Results showing Pearson's and Spearman's correlations can be found in Table 10 of the Appendix.

This same process was conducted thereafter, this time for the WTST measures.

5.2.1 Willingness-To-Play (WTPlay) games ranking analysis

In the first step of the experiment, subjects were asked to rank the 7 topics on how much they would be willing to play them (from 1 to 7), thus allowing for preference indifference. Since ranks were ordinal and not normally distributed, we relied on Friedman's Chi-Square, which is the non-parametric version of repeated measure ANOVA, to examine whether topics were ranked significantly differently. If significant at the 0.05 level, a Nemenyi post-hoc test (equivalent of the Tukey-test for ANOVA) for all pairwise combinations was performed.

Table 5: Means WTPlay ranks (and Standard Deviations) for the 7 different game genres across the 5 FFM traits

	Challenging	Shooter	Puzzle	Strategy	Role-Play	Simulation	Multiplayer	χ^2 (p-value)
OPE	3.14 (1.71)	$4.22^{\mathbf{a,b,c}}$ (1.98)	3.82 (1.97)	$3.08^{\mathbf{a}}$ (1.64)	$2.72^{\mathbf{b}}$ (1.58)	$3.04^{\mathbf{c}}$ (1.59)	3.76 (2.03)	29.28 (<0.001)
AGR	$3.22^{\mathbf{a,b}}$ (1.96)	$4.51^{\mathbf{a,c,d}}$ (1.72)	$4.56^{\mathbf{b,e}}$ (1.94)	$3.80^{\mathbf{f}}$ (1.65)	$2.22^{\mathbf{e,f,g,h}}$ (1.37)	$3.24^{\mathbf{c,g}}$ (1.45)	$3.29^{\mathbf{h}}$ (1.38)	60.60 (<0.001)
CON	3.25 (1.99)	$4.02^{\mathbf{a}}$ (2.15)	$4.11^{\mathbf{b}}$ (1.89)	3.57 (1.87)	$2.66^{\mathbf{a,b}}$ (1.51)	3.23 (1.43)	3.23 (1.71)	30.27 (<0.001)
NEU	3.44 (1.64)	4.27 (1.94)	$4.02^{\mathbf{a}}$ (2.02)	3.93 (1.72)	$2.68^{\mathbf{a}}$ (1.46)	3.00 (1.70)	3.88 (1.99)	38.35 (<0.001)
EXT	3.33 (1.86)	$3.33^{\mathbf{a,b}}$ (1.96)	$3.91^{\mathbf{c,d}}$ (1.86)	$3.21^{\mathbf{e}}$ (1.74)	$2.70^{\mathbf{a,c,e}}$ (1.47)	$3.63^{\mathbf{b,d}}$ (1.70)	3.07 (2.91)	20.05 (<0.001)

Notes: the χ^2 Friedman test corresponds to reject the null hypothesis that repeated ranks of the same individuals have the same distribution. Personality traits rejecting this null hypothesis at the 0.05 level are reported in bold. Topics that share the same subscripts row-wise indicates a significant difference at the < 0.05 level through a Nemenyi post-hoc test.
WTPlay = Willingness-To-Play , OPE = Openness To Experience, AGR = Agreeableness, CON = Conscientiousness, NEU = Neuroticism, EXT = Extraversion

As demonstrated on Table 5, we found that each of the five traits had a specific manner to rank game topics. For instance, individuals with an Openness To Experience predominant trait ranked game topics differently $\left(\chi^2(6) = 29.28, p < 0.001\right)$, with a Nemenyi post-hoc test indicating that the Shooter game topic was ranked significantly lower when compared to Strategy, Role-Play and Simulation type of games. Focusing on Agreeableness, a Friedman Chi-Square test showed that Agreeable individuals ranked game topics differently $\left(\chi^2(6) = 60.60, p < 0.001\right)$. A Nemenyi post-hoc test indicated that Agreeable individuals ranked Role-Play game topic significantly higher than Shooter, Strategy, Simulation and Multiplayer game topics, and that they ranked Shooter game topic significantly lower than Challenging and Puzzle game topic. In addition, post-hoc tests also revealed that Agreeable individuals significantly ranked the Challenging game topic higher than games with Puzzle elements. Regarding Conscientious individuals, a Friedman test showed that game topics were ranked differently

$\left(\chi^2(6) = 30.27,\, p < 0.001\right)$, with Role-Play game topics ranked significantly higher when compared to games featuring Shooter and Puzzle elements. Neurotic individuals also ranked games differently $\left(\chi^2(6) = 38.35,\, p < 0.001\right)$, and a Nemenyi post-hoc test revealed that these individuals ranked Role-Play games higher when compared with Puzzle game topic. Finally, the Friedman test also revealed that Extraverted individuals ranked topics differently $\left(\chi^2(6) = 20.05,\, p < 0.001\right)$, and post-hoc tests results showed that Role-Play game topic was ranked higher when compared to Shooter, Puzzle and Strategy game topics. Additionally, Extraverted individuals ranked Simulation significantly higher than Puzzle games, but lower when compared to Shooter games.

As the previous step only revealed if game topics were ranked differently according to predominant personality traits, it remains to be found if the score of each personality trait could lead to favor a particular game topic. Given this fact, a logistic regression was required to examine whether personality could predict a particular preference for a given video game genre. We constructed 7 logit models, one for each game topic yielded by our LDA model. The dependent variable was a dichotomous variable coded 1 if the game was ranked first by the individual, 0 otherwise. Regression models included the personality traits scores as predictors, controlling for age, gender and game use. Although correlations did exist between our variables of interest and were rather low, we checked for potential multicollinearity by computing variance inflation factors (VIF). Results[19] revealed that VIFs ranged from 1.121 to 1.430, thus indicating that multicollinearity was not an issue for our models.

Results[20] are displayed in Table 6. We reported the β coefficients and their respective standard errors, Odds-Ratios (OR) and the p-value given by a two-tailed z-test to check if each coefficient was different from 1. As results showed, Age[21] was a significant negative predictor of ranking first Challenging ($\beta = -0.145$, OR$= 0.865$, $p < 0.05$), Shooter ($\beta = -0.141$, OR$= 0.868$, $p < 0.1$) and Role-Play ($\beta = -0.057$, OR$= 0.944$,

[19] See Table 9 in Appendix for detailed results.

[20] Although we were mainly interested in the influence of traits on game topics being ranked first (i.e.: the one that was preferred), we also constructed 7 ordered logit models, to see if our variables of interest increased the likelihood of ranking a game higher (i.e.: which variable influenced the ranking). Data in Table 11 in the Appendix show results that are relatively consistent with those obtained through the logit models.

[21] Each variable effect is given *ceteris paribus*.

$p < 0.1$) game topics. Similarly, female subjects were negatively associated with ranking first games with Shooter ($\beta = -1.902$, OR$= 0.149$, $p < 0.05$) and Strategy game ($\beta = -1.967$, OR$= 0.140$, $p < 0.1$) components, whereas being a gamer positively predicted ranking first Shooter game topic ($\beta = 0.127$, OR$= 1.136$, $p < 0.05$).

Concerning personality traits, results indicated that a one-unit increase on the Openness to Experience sub-scale increased the odds of ranking Strategy game topic first by 7% ($\beta = 0.087$, OR$= 1.092$, $p < 0.05$) and Simulation game topic by 11% ($\beta = 0.104$, OR$= 1.110$, $p < 0.05$). Likewise, an increase of one-unit on the Conscientiousness sub-scale showed a 16% increase in the odds or ranking the Shooter first ($\beta = 0.153$, OR$= 1.166$, $p < 0.001$), whereas a one-unit increment on the Agreeableness sub-scale led to a 8% increase in the odds of ranking first games with Role-Play elements ($\beta = 0.082$, OR$= 1.086$, $p < 0.1$). High Extraversion scores positively predicted ranking first Shooter ($\beta = 0.100$, OR$= 1.106$, $p < 0.001$) and Multiplayer game topics ($\beta = 0.122$, OR$= 1.130$, $p < 0.001$), with a 10% and 13% positive change in the odds respectively for each additional unit. As for the Neuroticism trait, it was positively associated with Multiplayer game topic ($\beta = 0.082$, OR$= 1.086$, $p < 0.1$), yielding an increase of 8% in the odds for ranking it first for each increment on the sub-scale.

5.2.2 Willingness-To-Spend-Time (WTST) on games ranking analysis

This second approach differs from the Willingness-To-Play in that it required subjects to provide a precise ranking of their preferences through their willingness-to-spend-time on a type of game. As values were minutes given on a 0 to 120 scale, we converted them into ranks, with the highest value ranked first and the others in descending order. Similar to the WTPlay, ranks were not normally distributed, which is why we also relied on the Friedman Chi-Square to check whether topics were ranked differently. Likewise, if the Friedman was significant at the 0.05 level, a Nemenyi post-hoc for all pairwise combinations was conducted. Corresponding results can be found on Table 7 below.

Table 6: Results of logistic regressions on personality traits predicting preference for video game topics through Willingness-To-Play, controlling for Age, Gender and video game use ($N = 150$)

	WTPlay_Challenging		WTPlay_Shooter		WTPlay_Puzzle		WTPlay_Strategy		WTPlay_RolePlay		WTPlay_Simulation		WTPlay_Multiplayer	
	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)
Constant	-6.354 (3.318)	-	-1.019 (2.762)	-	-3.165 (2.608)	-	-3.424 (2.773)	-	-7.096 (2.451)	-	-7.048 (3.083)	-	-26.743 (1.870×10^3)	-
Age	-0.145** (0.066)	0.865	-0.141* (0.058)	0.868	-0.040 (0.036)	0.960	0.005 (0.038)	1.005	-0.057* (0.033)	0.944	-0.059 (0.044)	0.943	0.024 (0.039)	1.025
OPE	0.057 (0.051)	1.059	-0.070 (0.048)	0.932	0.068 (0.047)	1.070	0.087** (0.043)	1.092	0.044 (0.039)	1.045	0.104** (0.051)	1.110	0.041 (0.048)	1.042
AGR	0.082 (0.059)	1.086	-0.058 (0.053)	0.944	-0.054 (0.049)	0.947	-0.051 (0.051)	0.950	0.082* (0.041)	1.086	0.042 (0.049)	1.043	-0.084 (0.058)	0.919
CON	0.068 (0.056)	1.071	0.153*** (0.056)	1.166	0.056 (0.049)	1.058	-0.002 (0.047)	0.998	0.056 (0.040)	1.058	-0.018 (0.048)	0.982	0.082 (0.054)	1.086
NEU	0.033 (0.046)	1.034	0.026 (0.041)	1.027	-0.001 (0.039)	0.999	-0.031 (0.038)	0.969	0.049 (0.033)	1.051	0.051 (0.042)	1.053	0.082* (0.046)	1.086
EXT	0.057 (0.044)	1.059	0.100** (0.041)	1.106	-0.006 (0.037)	0.994	-0.001 (0.037)	0.999	-0.045 (0.033)	0.956	-0.021 (0.041)	0.979	0.122*** (0.046)	1.130
Gender (0=Man, 1=Woman)	-20.559 (5.629×10^3)	0.000	-1.902** (0.940)	0.149	0.006 (0.0673	1.007	-1.967* (1.098)	0.140	-0.084 (0.563)	0.919	-0.009 (0.0683)	0.991	-0.284 (0.710)	0.753
Gamer (0=No, 1=Yes)	-0.277 (1.218)	0.758	0.127** (0.922)	1.136	0.127 (0.780)	1.135	1.140 (1.153)	3.128	1.380 (0.879)	3.975	1.217 (1.165)	3.377	18.349 (1.870×10^3)	-
Pseudo-R² (McFadden)	.274		.230		.049		.123		.121		.113		.203	
Log-Likelihood	-42.765		-49.493		-52.365		-53.293		-68.373		-48.810		-45.403	

Notes: For each type of game, the dependent variable was a dummy variable coded 1 if the game was ranked first, 0 else.
*** $p < 0.01$ two-tailed, ** $p < 0.05$ two-tailed, * $p < 0.1$ two-tailed
WTPlay = Willingness-To-Play, SE = Standard Error, OPE = Openness To Experience, AGR = Agreeableness, CON = Conscientiousness, NEU = Neuroticism, EXT = Extraversion

Table 7: Means WTST ranks (and Standard Deviations) for the 7 different game genres across the 5 FFM traits

	Challenging	Shooter	Puzzle	Strategy	Role-Play	Simulation	Multiplayer	χ^2 (p-value)
OPE	3.90[a] (1.94)	5.12[a,b,c,d] (1.76)	4.44[e] (1.85)	2.94[b] (1.80)	3.56[c,e] (1.83)	3.20[d] (1.78)	3.92 (2.01)	35.93 (< 0.001)
AGR	3.95[a] (1.72)	5.29[a,b,c,d] (1.93)	4.83[e,f] (1.84)	2.82[g] (1.63)	4.00[b,e,g] (1.53)	3.07[c,f] (1.68)	3.54[d] (1.83)	55.54 (< 0.001)
CON	3.84 (1.64)	4.70[a] (1.94)	4.73[b] (1.95)	2.93 (1.81)	3.86[a,b] (2.00)	3.52 (1.97)	3.41 (1.85)	25.53 (< 0.001)
NEU	3.71[a] (1.79)	5.17[a,b,c] (1.66)	4.63[d,e] (2.09)	2.39[f] (1.48)	4.20[b,d,f,g] (1.79)	3.07[c,e] (1.68)	3.85[g] (1.93)	48.84 (< 0.001)
EXT	3.88 (1.79)	4.53 (2.03)	4.35 (2.09)	3.26 (1.85)	3.70 (1.92)	3.70 (1.90)	3.77 (2.07)	10.12 (0.11)

Notes: the χ^2 Friedman test corresponds to rejecting the null hypothesis that repeated ranks of the same individuals have the same distribution. Personality traits rejecting this null hypothesis at the 0.05 level are reported in bold. Topics that share the same subscripts row-wise indicates a significant difference at the < 0.05 level through a Nemenyi post-hoc test.

WTST = Willingness-To-Spend-Time, Multi = Multiplayer, OPE = Openness To Experience, AGR = Agreeableness, CON = Conscientiousness, NEU = Neuroticism, EXT = Extraversion

Similar to the WTPlay ranking, we found that all personality traits ranked game topics differently, with the exception of the Extraversion trait $\left(\chi^2(6) = 10.12,\ p = 0.11\right)$. Participants with a high-level Openness To Experience ranked games differently $\left(\chi^2(6) = 35.93,\ p < 0.001\right)$, and a Nemenyi post-hoc test revealed that Open participants ranked the Shooter game topic significantly lower than Challenging, Strategy, Role-Play and Simulation game topics. Additional post-hoc tests indicated that Open subjects ranked Role-Play game topic significantly higher than games featuring Puzzle elements. Friedman's test also revealed that Agreeable subjects ranked game topics differently $\left(\chi^2(6) = 55.54, p < 0.001\right)$. Nemenyi post-hoc indicated that Agreeable participants ranked Shooter games significantly lower when compared to Challenging, Role-Play, Simulation and Multiplayer game topics. Furthermore, a Friedman test showed that Agreeable subjects ranked Puzzle game topic lower when compared to games with Role-Play and Simulation features. Additionally, Role-Play game topic was ranked significantly higher when compared to games with Simulation elements. Regarding Neurotic subjects, the Friedman test revealed that game topics were ranked differently $\left(\chi^2(6) = 25.53,\ p < 0.001\right)$, and post-hoc tests indicated that the Shooter game topic was ranked significantly lower in comparison with Challenging, Role-Play and Simulation type of games. Similarly, Neurotic participants ranked Role-Play game topic lower when compared to Puzzle, Strategy and Multiplayer game topics, whilst ranking Simulation game topic significantly higher than a game containing Puzzle elements. Finally, Conscientious subjects also significantly ranked game topics differently $\left(\chi^2(6) = 48.84, p < 0.001\right)$, and a Nemenyi post-hoc revealed that this difference was due to ranking the Role-Play game topic significantly higher than Shooter or Puzzle genres of game.

Continuing on the same process that we did for WTPlay rankings, we investigated the relationship between personality traits and game topic preferences here. As such, we ran 7 binary logit regressions using the personality traits again as predictors variables, controlling for age, gender and game use. Likewise, we used the ranking of a game topic as our exogenous variable, coded 1 if the game topic was ranked first, 0 otherwise.

Results are displayed on Table 8[22]. It was found that Age[23] was negatively associated with ranking Challenging ($\beta = -0.138$, OR$= 0.871$, $p < 0.05$) and Shooter ($\beta = -0.181$, OR$= 0.834$, $p < 0.1$) game topics first, whereas it was positively associated with ranking games featuring Multiplayer ($\beta = 0.086$, OR$= 1.091$, $p < 0.05$) elements first. Being a female subject was a significant positive predictor of ranking Puzzle ($\beta = 1.600$, OR$= 4.955$, $p < 0.05$) and Multiplayer ($\beta = 1.870$, OR$= 6.488$, $p < 0.05$) game topics first. Being a gamer had no impact on ranking any game topic first.

Regarding personality traits, a one-unit increase on the Agreeableness sub-scale corresponded to a decrease of 9.3% in the odds of ranking the Strategy game topic first ($\beta = -0.097$, OR$= 0.907$, $p < 0.05$), whereas an increment on the Extraversion sub-scale increased the likelihood of 13% to rank the Shooter game topic first ($\beta = 0.128$, OR$= 1.137$, $p < 0.05$). For Neurotic subjects, a one-unit increase on the sub-scale led to a decrease of 9% in the likelihood of ranking Puzzle game topic ($\beta = -0.092$, OR$= 0.911$, $p < 0.05$) first, whilst it increased the odds of 5% to rank games with Role-Play elements first ($\beta = 0.050$, OR$= 1.050$, $p < 0.1$). In addition, it appeared that Conscientiousness scores were not significantly related with any of the game topics, nor were Openness To Experience scores.

6 Discussion

The purpose of our study was to investigate how personality traits might affect and shape individuals' preferences in terms of video game elements. Although we are not

[22] Similarly to the WTPlay measure, we conducted ordinal logistic regression, to investigate which predictors influenced the ranking of a given game topic. Results found in Table 12 of the Appendix show consistent results with those obtained here.

[23] Each variable effect is given *ceteris paribus*.

Table 8: Results of logistic regressions on personality traits predicting preference for video game topics through Willingness-To-Spend-Time, controlling for Age, Gender and video game use ($N = 150$)

	WTST_Challenging		WTST_Shooter		WTST_Puzzle		WTST_Strategy		WTST_RolePlay		WTST_Simulation		WTST_Multiplayer	
	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)
Constant	-1.320 (2.904)	-	-4.097 (-4.409)	-	6.863 (3.078)	-	0.637 (2.527)	-	-5.016 (2.057)	-	-4.511 (2.612)	-	2.315 (3.875)	-
Age	-0.138** (0.062)	0.871	-0.181* (0.101)	0.834	-0.039 (0.032)	0.961	0.035 (0.027)	1.036	-0.016 (0.021)	0.984	0.016 (0.027)	1.016	0.086*** (0.029)	1.091
OPE	0.027 (0.049)	1.028	-0.068 (0.074)	0.934	-0.079 (0.053)	0.923	0.059 (0.042)	1.062	0.049 (0.035)	1.050	-0.030 (0.041)	0.970	-0.095 (0.063)	0.909
AGR	0.048 (0.054)	1.050	0.100 (0.081)	1.106	-0.065 (0.049)	0.937	-0.097** (0.047)	0.907	0.052 (0.037)	1.054	0.019 (0.041)	1.020	-0.049 (0.068)	0.951
CON	0.036 (0.051)	1.037	0.029 (0.073)	1.029	-0.008 (0.050)	0.991	-0.059 (0.044)	0.942	0.032 (0.035)	1.033	0.014 (0.042)	1.015	-0.083 (0.061)	0.920
NEU	0.035 (0.043)	1.036	0.073 (0.063)	1.076	-0.092** (0.045)	0.911	-0.056 (0.037)	0.945	0.050* (0.029)	1.052	0.032 (0.036)	1.033	-0.070 (0.056)	0.932
EXT	-0.003 (0.041)	0.996	0.128** (0.065)	1.137	-0.044 (0.039)	0.957	-0.002 (0.034)	0.998	-0.038 (0.028)	0.962	0.026 (0.035)	1.027	0.039 (0.048)	1.040
Gender (0=Man, 1=Woman)	-22.53 (1.64×10^{10})	0.00	-1.908 (1.519)	0.148	1.600** (0.677)	4.955	0.021 (0.647)	1.022	-0.170 (0.496)	0.844	0.009 (0.601)	1.009	1.870** (0.828)	6.488
Gamer (0=No, 1=Yes)	-1.102 (0.982)	0.332	-0.592 (1.452)	0.553	-0.155 (0.694)	0.856	0.687 (0.766)	1.988	-0.266 (0.571)	0.766	1.192 (0.0877)	3.293	-0.080 (0.786)	0.923
Pseudo-R² (McFadden)	.198		.190		.146		.089		.054		.031		.264	
Log-Likelihood	-45.739		-27.571		-47.014		-58.57		-84.128		-63.881		-34.241	

Notes: For each type of game, the dependent variable was a dummy variable coded 1 if the game was ranked first, 0 else.
*** $p < 0.01$ two-tailed, ** $p < 0.05$ two-tailed, * $p < 0.1$ two-tailed
WTST = Willingness-To-Spend-Time, SE = Standard Error, OPE = Openness To Experience, AGR = Agreeableness, CON = Conscientiousness, NEU = Neuroticism, EXT = Extraversion

the first study to attempt such a thing (Braun et al., 2016, Chory and Goodboy, 2011, Potard et al., 2019, Peever et al., 2012, Worth and Book, 2015), this study is distinguished by the use of its own game genre nomenclature directly derived from the words of the players themselves. In this light, we differ from other studies that have used relevant, but top-down imposed classifications based on characteristics chosen by the authors manually (Apperley, 2006, Wolf, 2001). To achieve this, we first used topic modeling and millions of reviews to build a new data-driven nomenclature. This new classification comprised 7 topics, whose words in each topic allowed us to formulate singular conceptualizations of video games, closer to the players because topics are grounded on their own feelings and views regarding video games. We then attempted to investigate the relationship between FFM personality traits and video games, hypothesizing that the specific characteristics related to these traits would shape individuals' preferences toward video games. In the next section, we will discuss in more depth the relationship between traits and games, and how our findings relate to those in the literature.

6.1 Personality traits and game preferences

This paper was driven by the question: can personality traits be used to predict preference towards specific game genres? Such a question is particularly relevant for video game firms, that would focus their resources on the personality segment more likely to buy their game. We found convincing results that personality did influence both the ranking of game topics and the preference towards specific game topics.

In line with our first hypothesis, we found that Extraversion was positively associated with the shooting topic, which is consistent with the results of Braun et al. (2016) and Chory and Goodboy (2011). As stated before, individuals with a high Extraversion score are stimulation-seekers and thus, tend to prefer intense activities, which is highlighted by the presence of words such as "*jump*" and "*run*" in the Shooter game topic. This interpretation could also explain why subjects with a high Extraversion score expressed a preference for pro-social elements of the Multiplayer topic, thus being consistent with previous findings (Graham & Gosling, 2013) and the Extraverted pro-social nature. In addition, we also found that Conscientious individuals favored Shooter games. Such a result is surprising, as Conscientiousness often characterizes meticulous and less spontaneous individuals, thus contrasting with the very action-oriented words of the Shooter game topic. This finding is also inconsistent with those obtained by Potard et al. (2019), who found that Conscientious individuals tended to avoid action-oriented games. Furthermore, we fail to confirm the hypothesis that individuals with high Conscientiousness scores would prefer Strategy games.

We also hypothesized that individuals with high scores in Neurotic or Openness to Experience personality traits would prefer games with strong role-play and narrative elements. We confirmed our hypothesis for the former, thus being in line with previously obtained results (Braun et al., 2016), but failed to do so for the latter. In addition, we found that Neuroticism was positively associated with the Multiplayer topic, thus being consistent with Potard et al.'s (2019) results. As outlined in the hypothesis, this preference towards Role-Play and Multiplayer elements might be imputed to the fact that Neurotic individuals can "escape" their negative feelings and the real world for a virtual one through those games, by creating a second virtual and anonymous character. In light of this interpretation, it would also explain why high Neuroticism scores were negatively associated with puzzle elements, as they do not properly allow Neurotic individuals to escape their reality. Regarding Openness To Experience, while we failed to confirm our hypothesis regarding the link between this trait and role-play elements, we found that individuals with high scores in this trait were more likely to favor games containing strategy and simulation components. One explanation could be that both game topics included words suggesting for creativeness, which is one of the characteristics defining individuals with high Openness.

Finally, we hypothesized that a high score in Agreeableness would translate to a preference towards multiplayer games containing opportunities to help others (either human or robots). We were unable to confirm this hypothesis, but found a significant relationship with other types of games. For instance, we found that Agreeable subjects were likely to be attracted to games with role-play elements, whilst avoiding games with strategy elements. One interpretation for these findings is that the strategy topic did not contain words related to pro-social behavior, while the role-play topic did. As stated above, games with role-playing elements offer wider freedom of interaction and choice, thus attracting Agreeable individuals in their willingness to express pro-social behavior.

6.2 Practical implications and limitations

This paper provides substantial practical implications. First, as we said, we developed a new nomenclature of game topics which was developed using millions of reviews written by millions of different users. As a result, this nomenclature provides game topics that are much closer to the players' vocabulary than a manually made classification would be. Besides, because LDA modeling allows for flexibility, this nomenclature can be easily updated as new game concepts appear in the gamers' words. Hence, such topics could be of great interest to academics, aiming to study the influence of psychological factors on video game genres, as well as for companies wanting to use adequate terms in their game advertisements. Second, these topics were used to study their relationship with the personality traits of the FFM. Given the many evidential results of this paper, we contribute to the relevant literature by providing additional answers to how personality can structure preferences for video game elements.

A more critical practical implication of this paper lies in its use by video game firms and its relevance within the realm of consumer behavior. By providing evidence on which personality traits segments are more likely to be receptive to a given video game and its advertising, firms could use and implement these findings to be more competitive, optimize their resources, develop a more effective marketing campaign and build more profound customer affinity. For instance, by framing their advertisements using a vocabulary in line with personality-segmented consumers, marketers could expose consumers to games they would be more likely to enjoy, thus increasing their pleasure.

In addition to personality traits, this paper provides interesting insights on the relationship between age and gender, and game genres preferences. Indeed, we found that females were more likely to play games with Puzzle and Multiplayer elements, whereas males were more likely to play games with strategy elements, which is in line with results found in the relevant literature (Phan et al., 2012). Regarding age, we found that as individuals get older, they tend to prefer multiplayer games, and to shift away from games with violent or role-play elements. It should be noted that while some studies investigated age-related preferences on video games in pre-teenagers and adolescents (Homer et al., 2012) or university students (Greenberg et al., 2008), findings in the literature remain scarce for adults and elderly. Thus, while not the primary focus of our study, this paper provides new insight into how some demographics can influence game preferences, thus allowing firms to segment their market according to these factors as well.

Despite those interesting findings, our study is not without bias. First, topics were named manually, using the aggregated semantic meaning of words contained in it to do so. While not an issue for our experiment as we displayed only salient top words to subjects (thus letting them free to have their own interpretation), it becomes problematic if firms want to use those labels. To alleviate this issue, we believe, like Faisal and Peltoniemi (2015), that a machine learning approach using a probabilistic framework should be preferable to interpret the general meaning of a topic and name it accordingly. Furthermore, only English reviews were used to construct this new classification, thus possibly affecting what should be important for players in video games during the topic modeling process. While English is the most used language on the Steam platform with 41.13% of users reporting using it (Clement, 2021a), examining how main video game themes might evolve across languages could be an interesting line of investigation.

Another possible limitation that could be raised lies in the use of the Big Five Inventory. The use of such an instrument stemmed from the desire to have a short questionnaire to complete (44 items), as to prevent drop-out. However, a trade-off of this scale is that it only allowed us to collect the main personality traits, and not more in-depth facets. A study similar to ours using a lengthier, but also more precise scale, such as the NEO-PI (Costa & McCrae, 2008) or the IPIP (Goldberg et al., 2006) would therefore be an interesting lead to pursue. Additionally, we used the upper quartile to determine the predominant personality trait(s). Although this approach is valid to distinguish strong personality profiles (Fujiwara and Nagasawa, 2015; Rose et al., 2002; Segalin et al., 2017), it can omit subjects that do not have any trait(s) score(s) lying in the upper quartile, thus complicating the process for firms to segment those individuals according to their personality. Using the proper metric to assign predominant personality traits is relevant, particularly because firms need such a metric to segment their consumers.

In the same vein, while the focus of this study was on personality, the influence of alternative psychographic variables, such as the values and lifestyle of individuals, on game topic preferences may be equally relevant. These variables are indeed known to play a role in consumer preferences (Liu et al., 2019, Schwartz and Bilsky, 1987), making them prime candidates alongside personality for understanding what affects consumers' video game choices and opens up avenues for further research.

Finally, the results presented here are substantial, but limited by the fact that they demonstrate only correlations, not causal relationships. Therefore, one could build on this paper to construct a laboratory experiment, using a randomization process to establish a genuine causal link between consumers' personality traits and their video games preferences.

7 Conclusion

Properly segmenting their market and proposing the relevant products, understood and recognized by consumers, is a constant challenge for firms. As the video game market is becoming one of the main cultural mediums in current times, knowing whether personality traits influence preferences towards main game genres could be a critical advantage for video game companies. However, most of the studies investigating such a relationship relied on genres manually established, thus possibly not being recognized by gamers themselves. In this paper, we tackle this limitation by developing first a new nomenclature of video game genres, using a machine learning approach and millions of reviews written by thousands of unique users. This nomenclature proposes seven main genres composing the video game landscape, with the advantage of being more flexible than manual classifications. We use subsequently those novel data-driven topics to investigate their relationship with the personality traits of the Five-Factor Model.

Convincing results were found. As such, it was found that individuals with high Extraversion scores would be more likely to prefer games with multiplayer and shooter-oriented components, with the latter being also favored by highly Conscientious individuals. Additionally, we found that individuals with high Neuroticism scores would select games with multiplayer or role-play elements, while avoiding puzzle elements. Findings indicated that highly Agreeable individuals would prefer to play games with role-play elements, but would be more likely to avoid games involving strategy. Finally, results showed that high Openness To Experience scores would translate to a preference towards games containing strategy and simulation elements.

Consequently, besides developing a new data-driven game genre nomenclature, this paper provides new insights into how personality traits shape preferences towards specific elements. Specifically, this paper contributes to give a practical understanding for firms on the games to advertise to their personality-segmented consumers to maximize their sales and their consumers' enjoyment efficiently.

Nevertheless, further research is required to completely understand and explain what guides consumers in their purchase of video games, apart from personality traits. While this paper offers substantial contributions to this matter, we believe that conducting a laboratory experiment allowing to establish a causal relationship and implementing other psychological characteristics such as values or lifestyle are interesting and promising avenues to pursue.

Bibliography

Allport, G. W. (1938). *Personality: A Psychological Interpretation*. H. Holt.

Allport, G. W., & Oldbert, H. (1938). Trait-Names: A Psycho-lexical Study. *Psychological Review Monographs*, *211*.

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality Psychology and Economics. *NBER*. https://doi.org/10.3386/w16822

Aluja, A., García, Ó., & García, L. F. (2003). Relationships among extraversion, openness to experience, and sensation seeking. *Personality and Individual Differences*, *35*(3), 671–680. https://doi.org/10.1016/S0191-8869(02)00244-1

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H. R., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. *Psychological Bulletin*, *136*(2), 151–173. https://doi.org/10.1037/a0018251

Apperley, T. H. (2006). Genre and game studies: Toward a critical approach to video game genres. *Simulation & Gaming*, *37*(1), 6–23. https://doi.org/10.1177/1046878105282278

Bell, Y. W., & David, R. (2008). Market segmentation. *The Marketing Book* (pp. 260–282). Routledge. https://doi.org/10.4324/9780080942544-19

Blattberg, R. C., Peacock, P., & Sen, S. K. (1976). Purchasing Strategies across Product Categories. *Journal of Consumer Research*, *3*(3), 143–154. https://doi.org/10.1086/208661

Blei, D., Andrew, Y., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Boyce, C. J., & Wood, A. M. (2011). Personality and the marginal utility of income: Personality interacts with increases in household income to determine life satisfaction. *Journal of Economic Behavior & Organization*, *78*(1), 183–191. https://doi.org/fj4rdv

Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics*, *46*(4), 1171–1178.

Braun, B., Stopfer, J. M., Müller, K. W., Beutel, M. E., & Egloff, B. (2016). Personality and video gaming: Comparing regular gamers, non-gamers, and gaming addicts and differentiating between game genres. *Computers in Human Behavior*, *55*, 406–412. https://doi.org/10.1016/j.chb.2015.09.041

Brody, R. P., & Cunningham, S. M. (1968). Personality Variables and the Consumer Decision Process. *Journal of Marketing Research*, *5*(1), 50–57.

Carver, C. S., Sutton, S. K., & Scheier, M. F. (2000). Action, Emotion, and Personality: Emerging Conceptual Integration. *Personality and Social Psychology Bulletin*, *26*(6), 741–751. https://doi.org/10.1177/0146167200268008

Cattell, R. B. (1945). The Description of Personality: Principles and Findings in a Factor Analysis. *American Journal of Psychology*, *58*(1), 69–90.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Chory, R. M., & Goodboy, A. K. (2011). Is Basic Personality Related to Violent and Non-Violent Video Game Play and Preferences? *Cyberpsychology, Behavior,* *and Social Networking*, *14*(4), 191–198. https://doi.org/10.1089/cyber.2010.0076

Clarke, R. I., Lee, J. H., & Clark, N. (2015). Why Video Game Genres Fail: A Classificatory Analysis. *Games and Culture*, *12*(5), 445–465. https://doi.org/10.1177/1555412015591900

Clement, J. (2021a). Language spoken by global Steam users 2020 | Statista [[Online; accessed 23 Apr. 2021]]. https://www.statista.com/statistics/957319/steam-user-language

Clement, J. (2021b). Number of Steam MAU 2020 | Statista [[Online; accessed 8. Apr. 2021]]. https://www.statista.com/statistics/733277/number-stream-dau-mau

Costa, P. T., Jr., & McCrae, R. R. (2008). The five- factor theory of Personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *In book: Handbook of personality: Theory and research* (pp. 159–181). Guilford.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: a meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin*, *124*(2), 197–229.

Esuli, A. (2018). Steam-crawler [[Online; accessed 29. Dec. 2020]]. https://github.com/aesuli/steam-crawler

Faisal, A., & Peltoniemi, M. (2015). Establishing Video Game Genres Using Data-Driven Modeling and Product Databases. *Games and Culture*, *13*(1), 20–43. https://doi.org/ggb5gp

Fernàndez-Tobías, I., Braunhofer, M., Elahi, M., Ricci, F., & Cantador, I. (2016). Alleviating the new user problem in collaborative filtering by exploiting personality information. *User Modeling and User-Adapted Interaction*, *26*(2), 221–255. https://doi.org/ghchkc

Fujiwara, K., & Nagasawa, S. (2015). Relationships among Purchase Intentions for Luxury Brands and Personality Traits Based on the Big Five. *American Journal of Industrial and Business Management*, *5*(11), 631–639. https://doi.org/10.4236/ajibm.2015.511063

Furnham, A., & Fenton-O'Creevy, M. (2018). Personality and political orientation. *Personality and Individual Differences*, *129*, 88–91. https://doi.org/10.1016/j.paid.2018.03.020

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., Shibuya, A., Liau, A. K., Khoo, A., Bushman, B. J., Rowell Huesmann, L., & Sakamoto, A. (2009). The Effects of Prosocial Video Games on Prosocial Behaviors: International Evidence From Correlational, Longitudinal, and Experimental Studies. *Personality and Social Psychology Bulletin*, *35*(6), 752–763. https://doi.org/10.1177/0146167209333045

Giluk, T. L., & Postlethwaite, B. E. (2015). Big Five personality and academic dishonesty: A meta-analytic review. *Personality and Individual Differences*, *72*, 59–67. https://doi.org/10.1016/j.paid.2014.08.027

Goldberg, L. (1981). *Language and Individual Differences The Search for Universals in Personality Lexicons*. L. Wheeler.

Goldberg, L. (1990). An alternative "description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6).

Goldberg, L., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/dwwn9d

Graham, L. T., & Gosling, S. D. (2013). Personality Profiles Associated with Different Motivations for Playing World of Warcraft. *Cyberpsychology, Behavior, and Social Networking*, *16*(3), 189–193. https://doi.org/10.1089/cyber.2012.0090

Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A Dimension of Personality. *Handbook of Personality Psychology* (pp. 795–824). Academic Press.

Green, C. S., & Bavelier, D. (2012). Learning, Attentional Control, and Action Video Games. *Current Biology*, *22*(6), R197–R206. https://doi.org/10.1016/j.cub.2012.02.012

Greenberg, B. S., Sherry, J., Lachlan, K., Lucas, K., & Holmstrom, A. (2008). Orientations to Video Games Among Gender and Age Groups. *Simulation & Gaming*, *41*(2), 238–259.

Greitemeyer, T., & Mügge, D. O. (2014). Video Games Do Affect Social Outcomes: A Meta-Analytic Review of the Effects of Violent and Prosocial Video Game Play. *Personality and Social Psychology Bulletin*, *40*(5), 578–589. https://doi.org/f53ng6

Grover, R., & Srinivasan, V. (1989). An Approach for Tracking Within-Segment Shifts in Market Shares. *Journal of Marketing Research*, *26*(2), 230–236.

Hirsh, J. B., Kang, S. K., & Bodenhausen, G. V. (2012). Personalized Persuasion: Tailoring Persuasive Appeals to Recipients' Personality Traits. *Psychological Science*, *23*(6), 578–581. https://doi.org/10.1177/0956797611436349

Homer, B. D., Hayward, E. O., Frye, J., & Plass, J. L. (2012). Gender and player characteristics in video game play of preadolescents. *Computers in Human Behavior*, *28*(5), 1782–1789. https://doi.org/10.1016/j.chb.2012.04.018

Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*. https://spacy.io/

Hunt, S. D., & Arnett, D. B. (2004). Market Segmentation Strategy, Competitive Advantage, and Public Policy: Grounding Segmentation Strategy in Resource-Advantage Theory. *Australasian Marketing Journal (AMJ)*, *12*(1), 7–25. https://doi.org/cqm5ks

John, O., Naumann, L., & Soto, C. (2008). Paradigm shift to the integrative Big Five taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *In book: Handbook of personality: Theory and research* (pp. 114–158). Guilford.

Judge, T., Bono, J., Ilies, R., & Gerhardt, M. (2002). Personality and Leadership: A Qualitative and Quantitative Review. *Journal of applied psychology*, *87*(4), 765–80. https://doi.org/10.1037/0021-9010.87.4.765

Kamakura, W. A., & Mazzon, J. A. (1991). Value Segmentation: A Model for the Measurement of Values and Value Systems. *Journal of Consumer Research*, *18*(2), 208–218.

Karumur, R. P., Nguyen, T. T., & Konstan, J. A. (2016). Exploring the Value of Personality in Predicting Rating Behaviors: A Study of Category Preferences on MovieLens. *RecSys '16: Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 139–142). Association for Computing Machinery. https://doi.org/10.1145/2959100.2959140

Larsen, R., & Buss, D. (2017). *Personality Psychology: Domains of Knowledge About Human Nature*. McGraw-Hill Education.

Lee, J., Clarke, R. I., & Perti, A. (2015). Empirical evaluation of metadata for video games and interactive media. *Journal of the Association for Information Science and Technology*, *66*(12), 2609–2625. https://doi.org/10.1002/asi.23357

Lee, K., & Ashton, M. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, *39*(2), 329–358. https://doi.org/btqwqw

Li, X. (2020). Towards Factor-oriented Understanding of Video Game Genres using Exploratory Factor Analysis on Steam Game Tags. *2020 IEEE International Conference on Progress in Informatics and Computing (PIC)* (pp. 207–213). IEEE. https://doi.org/ggzm

Lin, C.-F. (2002). Segmenting customer brand preference: demographic or psychographic. *Journal of Product & Brand Management*. https://doi.org/10.1108/10610420210435443

Lin, D., Bezemer, C.-P., Zou, Y., & Hassan, A. E. (2019). An empirical study of game reviews on the Steam platform. *Empirical Software Engineering*, *24*(1), 170–207.

Liu, H., Huang, Y., Wang, Z., Liu, K., Hu, X., & Wang, W. (2019). Personality or Value: A Comparative Study of Psychographic Segmentation Based on an Online Review Enhanced Recommender System. *Applied Sciences*, *9*(10). https://doi.org/10.3390/app9101992

Lucas, R., Diener, E., Grob, A., Suh, E., & Shao, L. (2000). Cross-Cultural Evidence for the Fundamental Features of Extraversion. *Journal of personality and social psychology*, *79*(3), 452–68. https://doi.org/d8gp58

McCallum, A. K. (2002). *Mallet: A machine learning for language toolkit*. http://mallet.cs.umass.edu

McCarthy, E. J. (1960). *Basic Marketing - A Managerial Approach*. Richard D. Irwin, Inc.

McCrae, R. R., & Costa, P. T. (1997). Conceptions and Correlates of Openness to Experience. *Handbook of Personality Psychology* (pp. 825–847). Academic Press. https://doi.org/10.1016/B978-012134645-4/50032-9

Mcdonald, J. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire*, *1*.

Monahan, S. (2021). Video games have replaced music as the most important aspect of youth culture. *the Guardian*. https://www.theguardian.com/commentisfree/2021/jan/11/video-games-music-youth-culture

Myers, I. B. (1962). *The Myers-Briggs Type Indicator: Manual*. https://doi.org/ghh3sm

Nakatani, S. (2010). Language detection library for java. https://github.com/shuyo/language-detection

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: replicated factors structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, *66*(6). https://doi.org/cqx8f6

Nyhus, E. K., & Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, *26*(3), 363–384. https://doi.org/10.1016/j.joep.2004.07.001

Peever, N., Johnson, D., & Gardner, J. (2012). Personality & video game genre preferences. *IE '12: Proceedings of The 8th Australasian Conference on Interactive*

Entertainment: Playing the System (pp. 1–3). Association for Computing Machinery.

Phan, M. H., Jardina, J. R., Hoyle, S., & Chaparro, B. S. (2012). Examining the Role of Gender in Video Game Usage, Preference, and Behavior. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1496–1500. https://doi.org/ggnv

Pitt, C. S., Bal, A. S., & Plangger, K. (2020). New approaches to psychographic consumer segmentation: Exploring fine art collectors using artificial intelligence, automated text analysis and correspondence analysis. *European Journal of Marketing*. https://doi.org/gh97r4

Plaisant, O., Courtois, R., Réveillère, C., Mendelsohn, G. A., & John, O. P. (2010). Validation par analyse factorielle du Big Five Inventory français (BFI-Fr). Analyse convergente avec le NEO-PI-R. *Annales Médico-psychologiques, revue psychiatrique*, *168*(2), 97–106. https://doi.org/10.1016/j.amp.2009.09.003

Potard, C., Henry, A., Boudoukha, A., Courtois, R., Laurent, A., & Lignier, B. (2019). Video Game Players' Personality Traits: An Exploratory Cluster Approach to Identifying Gaming Preferences. *Psychology of Popular Media Culture*, *9*(4). https://doi.org/ggzn

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, *58*(1), 103–139. https://doi.org/dhvmjf

Rose, C. L., Murphy, L. B., Byard, L., & Nikzad, K. (2002). The role of the Big Five personality factors in vigilance performance and workload. *European Journal of Personality*, *16*(3), 185–200. https://doi.org/10.1002/per.451

Russoniello, C. V., O'Brien, K., & Parks, J. M. (2009). EEG, HRV and Psychological Correlates while Playing Bejeweled II: A Randomized Controlled Study. *Annual Review of Cybertherapy and Telemedicine 2009* (pp. 189–192). IOS Press. https://doi.org/10.3233/978-1-60750-017-9-189

Rustichini, A., DeYoung, C. G., Anderson, J. E., & Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, *64*, 122–137. https://doi.org/10.1016/j.socec.2016.04.019

Schwartz, S., & Bilsky, W. (1987). Toward A Universal Psychological Structure of Human Values. *Journal of Personality and Social Psychology*, *53*(3), 550–562. https://doi.org/b7v7gj

Segalin, C., Celli, F., Polonio, L., Kosinski, M., Stillwell, D., Sebe, N., Cristani, M., & Lepri, B. (2017). What your Facebook Profile Picture Reveals about your Personality. *MM '17: Proceedings of the 25th ACM international conference on Multimedia* (pp. 460–468). Association for Computing Machinery. https://doi.org/10.1145/3123266.3123331

Shaffer, J. A., & Postlethwaite, B. E. (2012). A Matter of Context: A Meta-Analytic Investigation of the Relative Validity of Contextualized and Noncontextualized Personality Measures. *Personnel Psychology*, *65*(3), 445–494.

Smith, W. R. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing*, *21*(1), 3–8. https://doi.org/gf96vc

Somers, R. H. (1962). A New Asymmetric Measure of Association for Ordinal Variables. *American Sociological Review*, *27*(6), 799–811.

Swope, K. J., Cadigan, J., Schmitt, P. M., & Shupp, R. (2008). Personality preferences in laboratory economics experiments. *Journal of Socio-Economics*, *37*(3), 998–1009.

Twedt, D. W. (1964). How Important to Marketing Strategy is the "Heavy User"? *Journal of Marketing*, *28*(1), 71–72. https://doi.org/10.1177/002224296402800116

Weber, M. R. (2015). The Relationship Between Personality and Student Learning. *Journal of Hospitality & Tourism Education*, *27*(4), 135–146. https://doi.org/ggzk

Wedel, M., & Kamakura, W. A. (2000). *Market Segmentation: Conceptual and Methodological Foundations* (Vol. 8). https://doi.org/10.1007/978-1-4615-4651-1

Wijman, T. (2020). The World's 2.7 Billion Gamers Will Spend $159.3 Billion on Games in 2020; The Market Will Surpass $200 Billion by 2023 | Newzoo [[Online; accessed 10. Nov. 2020]]. *Newzoo*. https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023

Wolf, M. J. P. (2001). *The Medium of the Video Game*. University of Texas Press.

Wolf, M. J. P. (2007). *The Video Game Explosion: A History from PONG to PlayStation and Beyond*. Greenwood.

Worth, N. C., & Book, A. S. (2015). Dimensions of video game behavior and their relationships with personality. *Computers in Human Behavior*, *50*, 132–140. https://doi.org/gft2p2

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. *WWW 2013 - Proceedings of the 22nd International Conference on World Wide Web*, 1445–1456.

Yankelovich, D. (1964). New criteria for market segmentation. *Harvard Business Review*, *42*, 83–90.

Young, S., Ott, L., & Feigin, B. (1978). Some Practical Considerations in Market Segmentation. *Journal of Marketing Research*, *15*(3), 405–412. https://doi.org/ggtc

A Additional statistical analysis

Variance Inflation Factors (VIF)

Before building logit models, we checked for possible multicollinearity between our variables of interest, by computing variable inflation factor (VIF). The VIFs values of all our explanatory variables displayed on Table 9 were quite low, thus indicating that multicollinearity was not an issue for our models.

Table 9: Variance Inflation Factor values predictor variables

Predictor variable	VIF
Age	1.12
Openness To Experience	1.15
Agreeableness	1.29
Conscientiousness	1.20
Neuroticism	1.22
Extraversion	1.15
Gender (0 = Man, 1=Woman)	1.43
Gamer (0 = No, 1=Yes)	1.34

WTPlay and WTST correlations

Table 10: Pearson's and Spearman's correlations between the WTPlay and the WTST measures for each of the 7 topics rankings

	Pearson's correlation between WTPlay and WTST measures	Spearman's correlation between WTPlay and WTST measures
Challenging topic	0.55****	0.51****
Shooter topic	0.58****	0.54****
Puzzle topic	0.43****	0.42****
Strategy topic	0.46****	0.38****
Role-Play topic	0.36****	0.44****
Simulation topic	0.23***	0.21**
Multiplayer topic	0.37****	0.30****

Notes: **** $p < 0.001$, *** $p < 0.01$, ** $p < 0.05$
WTPlay = Willingness-To-Play, WTST = Willingness-To-Spend-Time

Ordinal logistic regressions

As was previously said, we were mainly interested in the influence of personality traits on the game topic that was preferred (i.e. ranked first), hence the decision to rely on a set of binary logistic regressions. However, investigating whether personality traits influenced the ranking of a given game topic remains relevant. Such a thing can be achieved through ordinal logistic regression.

To conduct such regressions, we had to ensure that two assumptions were not violated. The first assumption is that no multicollinearity is present between our predictors, which was not the case here as assessed by the VIFs computed above. The second assumption is the parallel regression assumption, or proportional regression assumption. This assumption says that the coefficients that describe the odds of being in the lowest rank against all higher ranks of the independent variable have to be similar to those that describe the odds between the second-lowest category and all higher categories, and so on and so forth. This assumption thus indicates that since the relationship between each pair of rankings is the same, there is only one set of coefficients and, therefore, only one model. If this assumption were violated, different models should be needed to explain the relationship between each pair of ranks. We used the Brant test (Brant, 1990) of the *brant* library in the R software to test this assumption.

We conducted a Brant test for each game topic. The parallel regression assumption was respected for each of the 7 game topic, for both the WTPlay and WTST measures. Therefore, we conducted ordinal logistic regressions, for which we used personality traits scores, age, gender and game used as predictors. Game topic ranking was the dependent variable, for which 7 was considered as the lowest rank, whereas 1 was considered the highest. We used Somers' D (Somers, 1962), which measures the strength of association between the predictors and the dependent variable, to evaluate models performance. Results for the WTPlay ranking can be found in Table 11, whereas results for the WTST ranking is displayed in Table 12.

Table 11: Results of ordinal logistic regressions on personality traits predicting video game topics ranking through Willingness-To-Play, controlling for Age, Gender and video game use ($N = 150$)

	WTPlay_Action		WTPlay_Shooter		WTPlay_Puzzle		WTPlay_Strategy		WTPlay_RolePlay		WTPlay_Simulation		WTPlay_Multiplayer	
	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)
Age	-0.065****	0.936	-0.046***	0.954	0.0005	1.000	-0.022	0.978	-0.021	0.978	0.003	1.003	-0.010	0.989
	(0.018)		(0.017)		(0.016)		(0.016)		(0.017)		(0.016)		(0.016)	
OPE	0.027	1.027	-0.046*	0.954	0.024	1.024	0.069**	1.072	0.048*	1.050	0.044	1.045	-0.032	0.968
	(0.027)		(0.026)		(0.026)		(0.027)		(0.027)		(0.027)		(0.027)	
AGR	0.049*	1.050	-0.016	0.984	-0.044	0.956	-0.034	0.966	0.045	1.046	0.010	1.010	-0.0133	0.986
	(0.029)		(0.028)		(0.028)		(0.028)		(0.029)		(0.028)		(0.027)	
CON	0.057**	1.059	0.025	1.025	0.016	1.016	0.017	1.017	0.027	1.028	0.006	1.006	0.036	1.036
	(0.027)		(0.026)		(0.026)		(0.028)		(0.027)		(0.026)		(0.027)	
NEU	0.038	1.038	0.020	1.020	-0.015	0.984	-0.015	0.984	0.020	1.020	0.053**	1.054	-0.002	0.997
	(0.023)		(0.022)		(0.023)		(0.023)		(0.022)		(0.024)		(0.023)	
EXT	0.006	1.006	0.076****	1.079	0.003	1.003	0.024	1.024	-0.036	0.964	-0.014	0.985	0.094****	1.099
	(0.022)		(0.022)		(0.021)		(0.022)		(0.022)		(0.023)		(0.023)	
Gender (0=Man, 1=Woman)	-1.436****	1.196	-1.416****	1.006	-0.144	0.986	-1.375****	1.897	-0.217	1.970	-0.552	1.519	-0.127	1.073
	(0.408)		(0.417)		(0.386)		(0.409)		(0.387)		(0.391)		(0.396)	
Gamer (0=No, 1=Yes)	0.181	0.238	0.007	0.242	-0.013	0.866	0.639	0.252	0.676	0.802	0.418	0.576	0.071	0.879
	(0.461)		(0.447)		(0.436)		(0.442)		(0.451)		(0.444)		(0.425)	
Intercepts:														
7|6	0.349		-2.893		-1.993		-0.815		-0.184		0.185		-0.844	
	(1.586)		(1.571)		(1.583)		(1.526)		(1.667)		(1.629)		(1.561)	
6|5	1.079		-2.054		-1.439		-0.223		0.675		1.401		-0.463	
	(1.570)		(1.556)		(1.576)		(1.521)		(1.625)		(1.603)		(1.559)	
5|4	2.998		-0.562		-0.127		1.530		2.653		3.328		1.166	
	(1.568)		(1.542)		(1.574)		(1.531)		(1.610)		(1.613)		(1.549)	
4|3	3.622		0.094		0.701		2.508		3.324		3.947		2.148	
	(1.580)		(1.546)		(1.579)		(1.538)		(1.617)		(1.623)		(1.551)	
3|2	3.622		0.094		0.701		2.508		3.324		3.947		2.148	
	(1.580)		(1.546)		(1.579)		(1.538)		(1.617)		(1.623)		(1.551)	
2|1	5.330		1.019		1.547		3.621		4.955		5.401		3.336	
	(1.629)		(1.554)		(1.588)		(1.551)		(1.652)		(1.652)		(1.576)	
Somer's D	.342		.365		.117		.312		.233		.207		.255	
Log-Likelihood	-225.180		-239.551		-254.652		-234.497		-222.146		-234.326		-239.630	

Notes: **** $p < 0.001$, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
WTPlay = Willingness-To-Play, SE = Standard Error, OPE = Openness To Experience, AGR = Agreeableness, CON = Conscientiousness, NEU = Neuroticism, EXT = Extraversion

Table 12: Results of ordinal logistic regressions on personality traits predicting video game topics ranking through Willingness-To-Spend-Time, controlling for Age, Gender and video game use ($N = 150$)

	WTST_Action		WTST_Shooter		WTST_Puzzle		WTST_Strategy		WTST_RolePlay		WTST_Simulation		WTST_Multiplayer	
	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)	Coefficient (SE)	Odds-Ratio exp(coef)
Age	-0.056*** (0.017)	0.945	-0.008 (0.016)	0.992	-0.003 (0.015)	0.996	0.029* (0.016)	1.029	-0.007 (0.016)	0.992	-0.0008 (0.015)	0.999	0.042** (0.017)	1.043
OPE	0.014 (0.027)	1.014	-0.032 (0.027)	0.967	-0.008 (0.026)	0.991	0.050* (0.026)	1.051	0.026 (0.026)	1.027	-0.008 (0.027)	0.991	-0.014 (0.027)	0.985
AGR	0.031 (0.028)	1.031	-0.040 (0.029)	0.960	-0.040 (0.027)	0.960	-0.026 (0.029)	0.973	0.038 (0.029)	1.039	0.007 (0.027)	1.007	0.039 (0.027)	1.040
CON	0.032 (0.027)	1.033	0.035 (0.027)	1.036	-0.008 (0.027)	0.991	-0.034 (0.027)	0.965	0.007 (0.027)	1.008	0.001 (0.027)	1.001	-0.009 (0.026)	0.990
NEU	0.063*** (0.023)	1.065	0.014 (0.023)	1.014	-0.038* (0.023)	0.961	-0.044* (0.023)	0.956	0.033 (0.023)	1.034	-0.003 (0.022)	0.996	-0.013 (0.022)	0.986
EXT	0.0005 (0.021)	1.000	0.030 (0.022)	1.031	0.004 (0.022)	1.004	-0.005 (0.021)	0.994	-0.062*** (0.023)	0.939	-0.009 (0.021)	0.990	0.019 (0.021)	1.019
Gender (0=Man, 1=Woman)	-1.064*** (0.377)	1.740	-1.710**** (0.430)	1.027	0.981** (0.398)	0.733	-0.741* (0.390)	0.999	0.343 (0.388)	1.325	0.894** (0.384)	1.538	0.650* (0.391)	0.341
Gamer (0=No, 1=Yes)	0.554 (0.459)	0.345	0.028 (0.439)	0.180	-0.310 (0.448)	2.667	-0.0006 (0.436)	0.476	0.281 (0.467)	1.408	0.430 (0.427)	2.446	-1.073** (0.438)	1.917
Intercepts:														
7|6	-0.191 (1.559)	0.825	-2.631 (1.598)	0.071	-3.787 (1.618)	0.022	-2.694 (1.567)	0.067	-1.190 (1.658)	0.304	-1.960 (1.519)	0.140	-0.629 (1.519)	0.533
6|5	1.489 (1.544)	4.434	-1.259 (1.584)	0.283	-3.040 (1.613)	0.047	-2.005 (1.561)	0.134	-0.072 (1.633)	0.930	-1.134 (1.508)	0.321	0.097 (1.514)	1.102
5|4	3.158 (1.567)	23.528	0.286 (1.580)	1.332	-2.037 (1.607)	0.130	-0.269 (1.560)	0.763	1.3192 (1.623)	3.740	0.281 (1.499)	1.325	1.173 (1.515)	3.233
4|3	3.904 (1.582)	49.638	0.952 (1.586)	2.593	-1.359 (1.602)	0.256	0.557 (1.559)	1.745	1.828 (1.623)	6.224	0.970 (1.498)	2.640	2.150 (1.527)	8.590
3|2	3.904 (1.582)	49.639	0.952 (1.586)	2.593	-1.359 (1.602)	0.256	0.557 (1.559)	1.745	1.828 (1.623)	6.224	0.970 (1.498)	2.640	2.150 (1.527)	8.590
2|1	4.737 (1.596)	114.193	1.920 (1.611)	6.824	-0.682 (1.599)	0.505	1.187 (1.559)	3.280	2.993 (1.636)	19.952	2.009 (1.507)	7.461	3.758 (1.564)	42.873
Somer's D	.339		.355		.197		.223		.225		.127		.264	
Log-Likelihood	-235.170		-228.401		-251.004		-241.445		-237.820		-254.842		-246.774	

Notes: **** $p < 0.001$, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
WTST = Willingness-To-Spend-Time, SE = Standard Error, OPE = Openness To Experience, AGR = Agreeableness, CON = Conscientiousness, NEU = Neuroticism, EXT = Extraversion

B Supplementary information

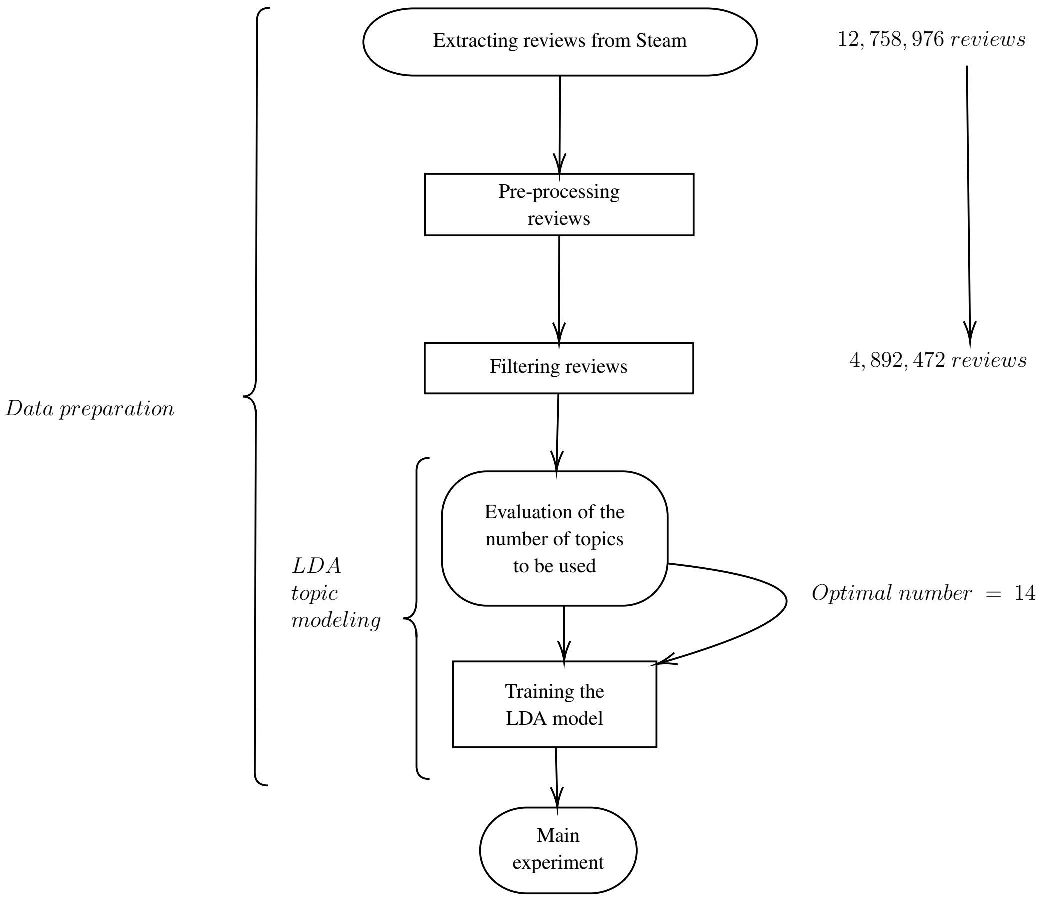


Figure 4: Workflow of the experimental design

The Big Five Inventory instrument (John et al., 2008)

To assess personality traits, we used the Big Five Inventory, developed by John et al. (2008) in our experiment. We used Plaisant et al.'s (2010) translation in the French version of the experiment.

Here are a number of characteristics that may or may not apply to you. Please answer for each statement to indicate the extent to which you agree or disagree with that statement.

Table 13: Big Five Inventory instrument (John et al., 2008)

1 Disagree strongly	2 Disagree a little	3 Neither agree nor disagree	4 Agree a little	5 Agree strongly

I see Myself as Someone Who...

1. Is talkative	23. Tends to be lazy
2. Tends to find fault with others	24. Is emotionally stable, not easily upset
3. Does a thorough job	25. Is inventive
4. Is depressed, blue	26. Has an assertive personality
5. Is original, comes up with new ideas	27. Can be cold and aloof
6. Is reserved	28. Perseveres until the task is finished
7. Is helpful and unselfish with others	29. Can be moody
8. Can be somewhat careless	30. Values artistic, aesthetic experiences
9. Is relaxed, handles stress well	31. Is sometimes shy, inhibited
10. Is curious about many different things	32. Is considerate and kind to almost everyone
11. Is full of energy	33. Does things efficiently
12. Starts quarrels with others	34. Remains calm in tense situations
13. Is a reliable worker	35. Prefers work that is routine
14. Can be tense	36. Is outgoing, sociable
15. Is ingenious, a deep thinker	37. Is sometimes rude to others
16. Generates a lot of enthusiasm	38. Makes plans and follows through with them
17. Has a forgiving nature	39. Gets nervous easily
18. Tends to be disorganized	40. Likes to reflect, play with ideas
19. Worries a lot	41. Has few artistic interests
20. Has an active imagination	42. Likes to cooperate with others
21. Tends to be quiet	43. Is easily distracted
22. Is generally trusting	44. Is sophisticated in art, music, or literature

BFI scale scoring ("R" denotes reverse-scored items):

- Extraversion: 1, 6R, 11, 16, 21R, 26, 31R, 36
- Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42
- Conscientiousness: 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R
- Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39
- Openness: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

Table 14: Three sample reviews and their associated metadata from the review dataset

game_id	1000080	999660	105600
game_Name	Zengeon	SAMURAI SHODOWN NEOGEO COLLECTION	Terraria
Price	16.79	39.99	9.99
game_n_review	97	19	291903
game_review_positive	78	15	24249
game_review_negative	19	4	5510
game_review_score	8.0	6.0	9.0
game_review_score_desc	Very Positive	Mostly Positive	Overwhelmingly Positive
user_ID	76561198077959093	76561198045090676	76561198148464111
user_n_games	146	93	82
user_n_review	57	1	10
user_postdate	2020-07-10 12:14:56	2020-06-21 04:35:17	2020-05-20 20:01:11
user_playtime_hours	0.2	4.6	121.8
user_review_text	I feel furious because this game looked amazing, anime+roguelike elements and when I played it... it was so boring	At first glance, This seems like an amazing collection. However something is very wrong and the game is broken. Difficulty is tuned way too high like it was a rip out of an arcade machine from the 80s.	If you want to pass the time, Terraria is the best game in the world to do that! I highly recommend you get this game!
user_recommended	False	False	True
user_game_purchased	True	True	True
user_review_helpful	3	5	0
user_review_funny	0	6	0
user_review_score	0.402557	0.347539	0.0
comments_for_this_review	0	0	0

C Main experiment protocol

Consent form and introduction

Consent

The LaPEA, laboratory of Paris Descartes University, is responsible of the data processing according to the General Data Protection Regulation (GDPR) UE 2016/679 and according to the modified informatic law of January, 6 1978.

As part of the questionnaire, you agree to transmit us: data about personal opinions, technical data (date and time of your connection via your device). Data collected in this study are anonymous.

These data are necessary to: (i) enable the questionnaire, (ii) perform a statistical averaging of the responses, (iii) ensure the security of our computer system against possible malicious acts.

Only the Laboratory of LaPEA (Paris Descartes University) can process the recollected data.

How long do we keep the data:

- the data relating to the archiving of your results: 1 month
- data relating to archiving your anonymised results for research purposes: 10 years.

Your rights:
You have a right to access, modify, delete and oppose the processing of your personal data, in compliance with the 2016 EU RGPD and the Data Protection Act of 06/01/1978 as amended. You can exercise your rights by sending an email to tod.lubart@parisdescartes.fr. These rights are exclusively personal. We may ask you for proof of identity.

I have read and understand the above statement. I also understand that completion of the questionnaire is entirely voluntary and that I may stop responding if I wish at any time and for any reason by closing this window.

Would you agree to participate in this survey? Yes, I agree (Next) No, I decline (exit survey)

Overview of the experiment

Welcome to this experiment in behavioral economics, conducted as part of a Master 2 thesis. The goal of this study is to investigate preferences in terms of video games. If you would like more information about this study or if you have any questions, please feel free to contact me by email (alexandre.vessereau@etu.univ-paris1.fr).

At the end of the experiment, you will have the opportunity to enter a random draw to win and choose between an **Amazon Gift Card worth €60** or a **Steam gift card worth €60.**

Please click on Next to start the experiment.

Next

Part I: Measuring Willingness-To-Play

In this section, subjects were asked whether they would like to play a game, defined by 10 words. There were 7 pages, one topic per page, and page presentation was randomized.

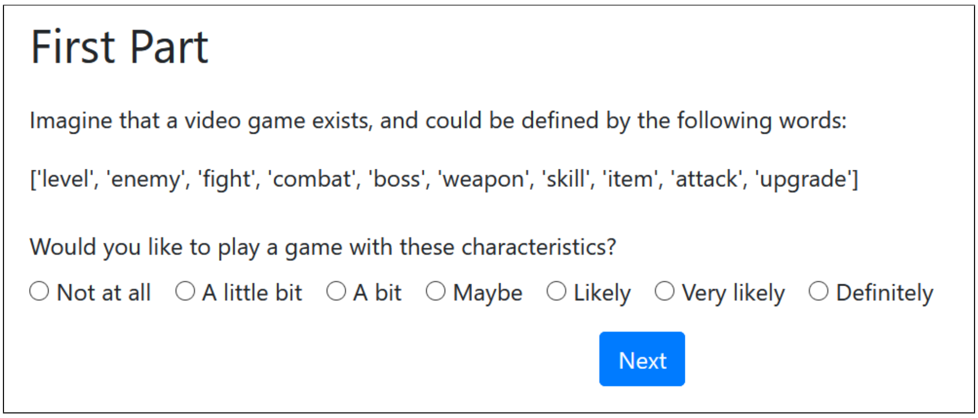


Figure 5: Willingness-To-Play for the *Action* game topic

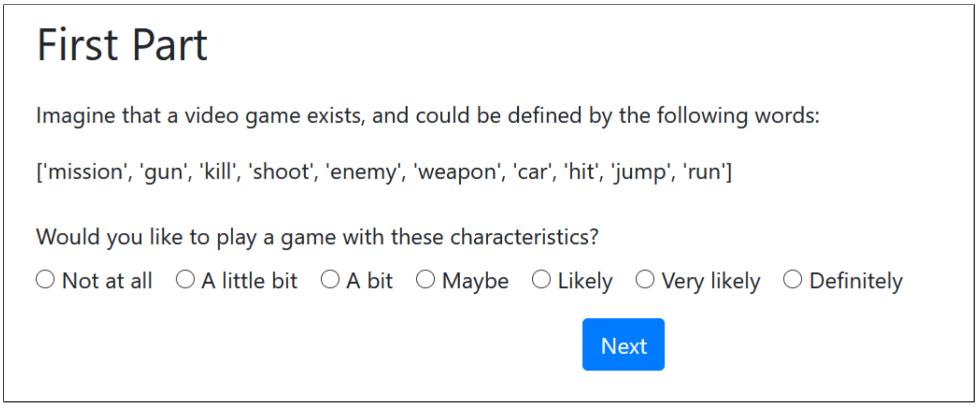


Figure 6: Willingness-To-Play for the *Shooter* game topic

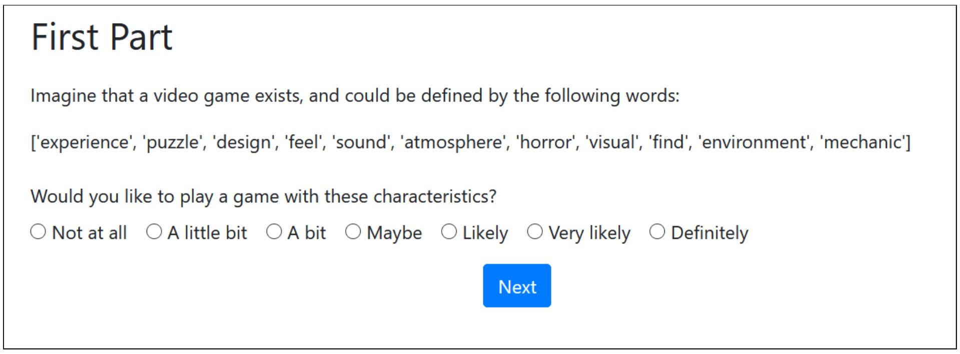


Figure 7: Willingness-To-Play for the *Puzzle* game topic

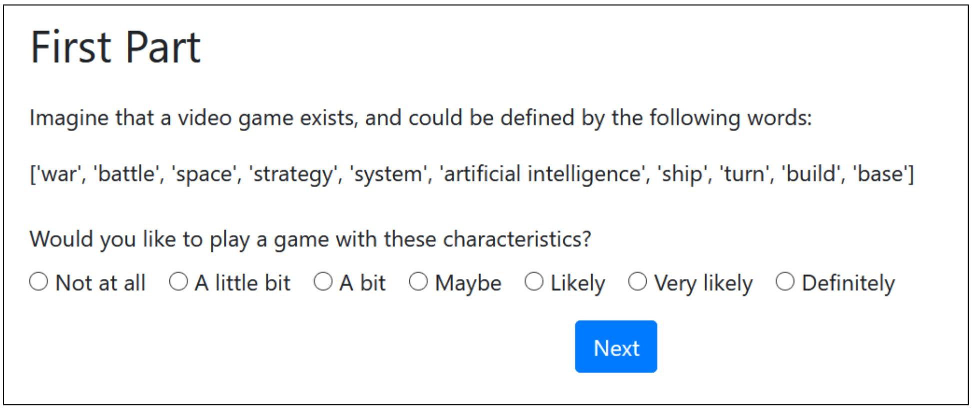


Figure 8: Willingness-To-Play for the *Strategy* game topic

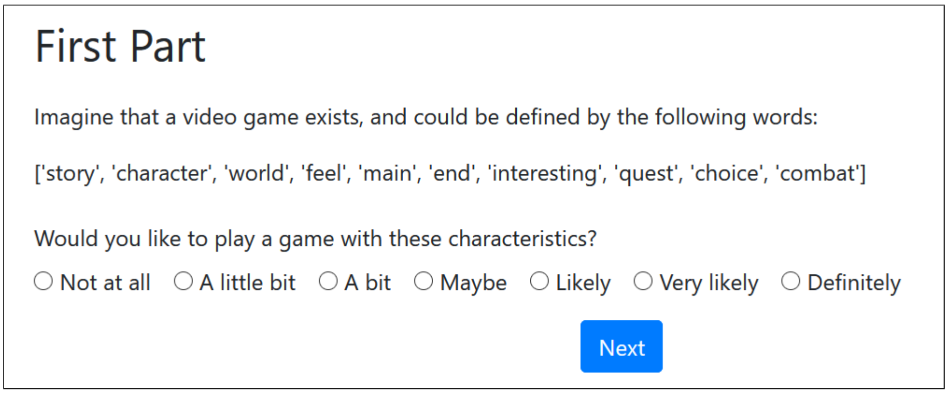


Figure 9: Willingness-To-Play for the *Role-Play* game topic

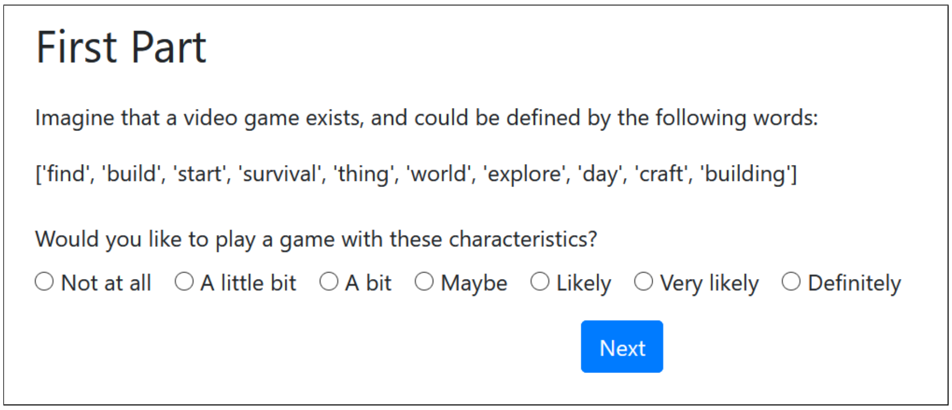


Figure 10: Willingness-To-Play for the *Simulation* game topic

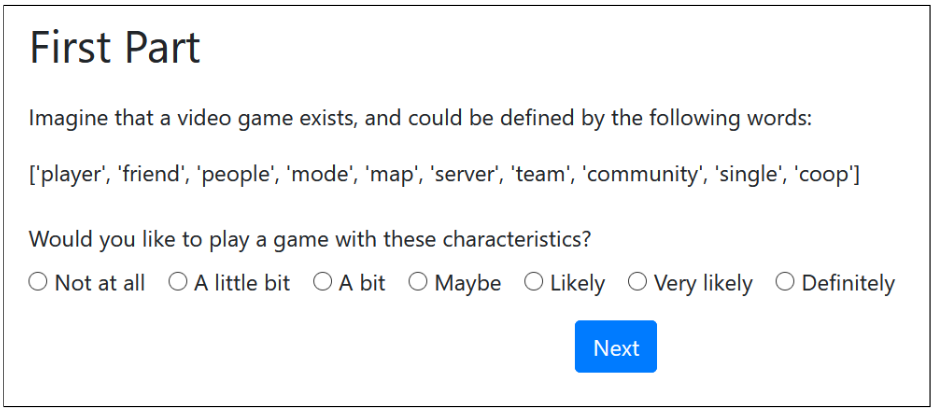


Figure 11: Willingness-To-Play for the *Multiplayer* game topic

Assessing subjects' personality with the Big Five Inventory (John et al., 2008)

Second Part 1/2

Here are a number of characteristics that may or may not apply to you. Please answer for each statement to indicate the extent to which you agree or disagree with that statement.

I see Myself as Someone Who...

	Disagree strongly	**Disagree a little**	**Neither agree nor disagree**	**Agree a little**	**Agree Strongly**
Is talkative	○	○	○	○	○
Tends to find fault with others	○	○	○	○	○
Does a thorough job	○	○	○	○	○
Is depressed, blue	○	○	○	○	○
Is original, comes up with new ideas	○	○	○	○	○
Is reserved	○	○	○	○	○
Is helpful and unselfish with others	○	○	○	○	○
Can be somewhat careless	○	○	○	○	○
Is relaxed, handles stress well	○	○	○	○	○
Is curious about many different things	○	○	○	○	○
	Disagree strongly	**Disagree a little**	**Neither agree nor disagree**	**Agree a little**	**Agree Strongly**
Is full of energy	○	○	○	○	○
Starts quarrels with others	○	○	○	○	○
Is a reliable worker	○	○	○	○	○
Can be tense	○	○	○	○	○
Is ingenious, a deep thinker	○	○	○	○	○
Generates a lot of enthusiasm	○	○	○	○	○
Has a forgiving nature	○	○	○	○	○
Tends to be disorganized	○	○	○	○	○
Worries a lot	○	○	○	○	○
Has an active imagination	○	○	○	○	○
	Disagree strongly	**Disagree a little**	**Neither agree nor disagree**	**Agree a little**	**Agree Strongly**
Tends to be quiet	○	○	○	○	○
Is generally trusting	○	○	○	○	○
Tends to be lazy	○	○	○	○	○
Is emotionally stable, not easily upset	○	○	○	○	○
Is inventive	○	○	○	○	○
Has an assertive personality	○	○	○	○	○

Next

Second Part 2/2

Here are a number of characteristics that may or may not apply to you. Please answer for each statement to indicate the extent to which you agree or disagree with that statement.

I see Myself as Someone Who...

	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree Strongly
Can be cold and aloof	○	○	○	○	○
Perseveres until the task is finished	○	○	○	○	○
Can be moody	○	○	○	○	○
Values artistic, aesthetic experiences	○	○	○	○	○
Is sometimes shy, inhibited	○	○	○	○	○
Is considerate and kind to almost everyone	○	○	○	○	○
Does things efficiently	○	○	○	○	○
Remains calm in tense situations	○	○	○	○	○
Prefers work that is routine	○	○	○	○	○
Is outgoing, sociable	○	○	○	○	○
	Disagree strongly	**Disagree a little**	**Neither agree nor disagree**	**Agree a little**	**Agree Strongly**
Is sometimes rude to others	○	○	○	○	○
Makes plans and follows through with them	○	○	○	○	○
Gets nervous easily	○	○	○	○	○
Likes to reflect, play with ideas	○	○	○	○	○
Has few artistic interests	○	○	○	○	○
Likes to cooperate with others	○	○	○	○	○
Is easily distracted	○	○	○	○	○
Is sophisticated in art, music, or literature	○	○	○	○	○

Next

Part III: Measuring Willingness-To-Spend-Time

Time allocation task

Now, let's imagine you need to devote some time, between 0 and 120 minutes, to **each game** described with the words in the lists below (e.g., 117 minutes to List 1, 102 minutes to List 2, etc...). Please allocate the minutes as you see fit for the 7 lists, making sure to assign a different value for each.

List 1 :
['level', 'enemy', 'fight', 'combat', 'boss', 'weapon', 'skill', 'item', 'attack', 'upgrade']

You devote 35 minutes of your time to this list describing a particular game.

0 120

List 2:
['mission', 'gun', 'kill', 'shoot', 'enemy', 'weapon', 'car', 'hit', 'jump', 'run']

You devote 15 minutes of your time to this list describing a particular game.

0 120

List 3 :
['experience', 'puzzle', 'design', 'feel', 'sound', 'atmosphere', 'horror', 'visual', 'find', 'environment', 'mechanic']

You devote 92 minutes of your time to this list describing a particular game.

0 120

List 4 :
['war', 'battle', 'space', 'strategy', 'system', 'artificial intelligence', 'ship', 'turn', 'build', 'base']

You devote 51 minutes of your time to this list describing a particular game.

0 120

List 5 :
['story', 'character', 'world', 'feel', 'main', 'end', 'interesting', 'quest', 'choice', 'combat']

You devote 27 minutes of your time to this list describing a particular game.

0 120

List 6 :
['find', 'build', 'start', 'survival', 'thing', 'world', 'explore', 'day', 'craft', 'building']

You devote 41 minutes of your time to this list describing a particular game.

0 120

List 7 :
['player', 'friend', 'people', 'mode', 'map', 'server', 'team', 'community', 'single', 'coop']

You devote 46 minutes of your time to this list describing a particular game.

0 120

Please answer the questions below:

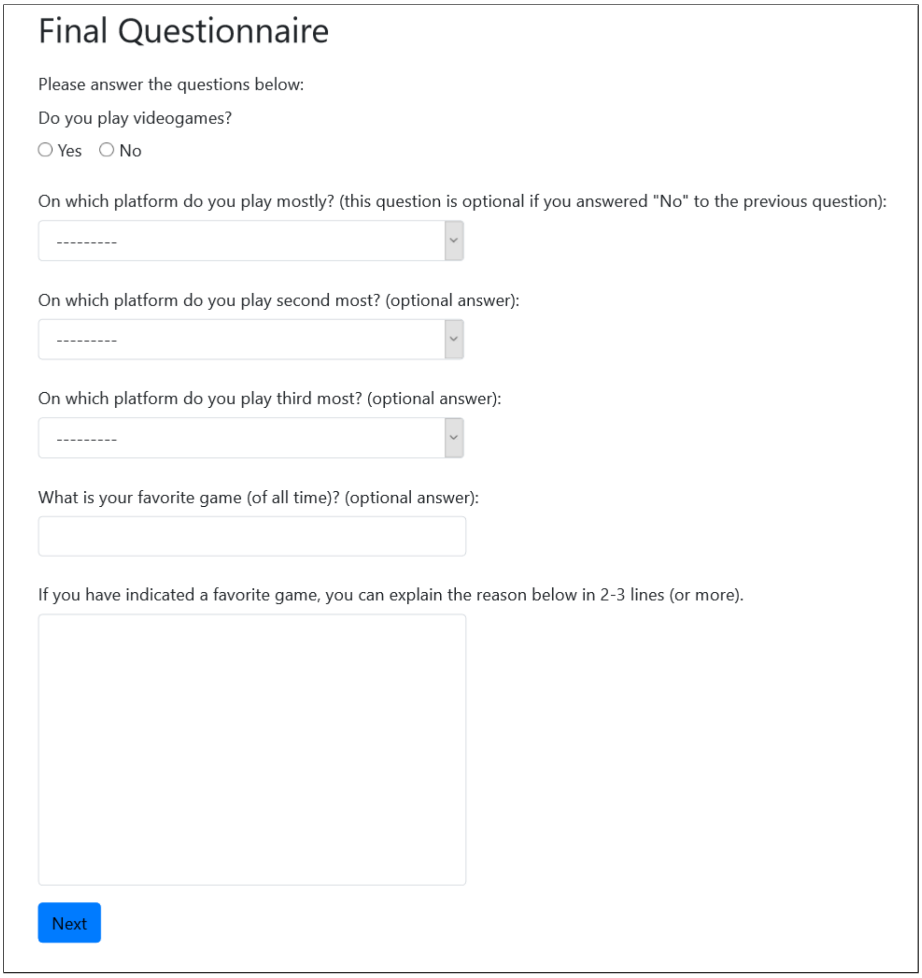
Why did you choose to devote your playing time this way?

○ Because I prefer to devote my time to games that I know little/no about
○ Because I prefer to devote my time playing games that I particularly enjoy
○ I had a hard time choosing and allocating my playing time ○ Other

If you answered Other to the previous question, please justify your choice below:

Next

Exit survey


Final Questionnaire

Please answer the questions below:

Do you play videogames?

○ Yes ○ No

On which platform do you play mostly? (this question is optional if you answered "No" to the previous question):

On which platform do you play second most? (optional answer):

On which platform do you play third most? (optional answer):

What is your favorite game (of all time)? (optional answer):

If you have indicated a favorite game, you can explain the reason below in 2-3 lines (or more).

Next

Final Questionnaire

Please answer the questions below:

How old are you?

What is your gender?

○ Man ○ Woman ○ Other

In which country do you live?

What is your field of study?

○ Mathematics ○ Sciences (biology/physics/chemistry) ○ Social Sciences ○ Sales representative ○ Economy ○ I.T ○ Law ○ Foreign languages ○ Other

What is your socio-professional category?

○ Student ○ Teacher ○ Farmer ○ Blue collar ○ Retail worker ○ Artisan ○ Executive ○ Free-Lance ○ Unemployed ○ Retired ○ Other

What is your educational attainment?

○ Up to secondary ○ Undergraduate ○ Master level ○ PhD Level ○ Other

Would you like to participate in the prize draw (to win either a Amazon or Steam gift card worth 60€)?

○ Yes ○ No

Next

End of the experiment

Prize Draw:

Here is your code to enter the draw: **uyb9ir3d**. Please write it down carefully.

The draw will take place on 05/30/2021 and the link to enter your code to see if you have won will be available at this address: https://docs.google.com/document/d/1-KsZNa-r2gGBK99GgGnllsXDR56-lIPL4pXffee-tSY/edit?usp=sharing

Thank you for participating in this study! You can now leave by closing your browser window, or if you used SurveySwap, by clicking on this link: https://surveyswap.io/sr/sLLjMF9c2qZEmNXM