



Universiteit
Leiden
The Netherlands

Exploration through video games

Gómez Maureira, M.A.

Citation

Gómez Maureira, M. A. (2023, April 26). *Exploration through video games*. Retrieved from <https://hdl.handle.net/1887/3594721>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3594721>

Note: To cite this publication please use the final published version (if applicable).

4 Video Games That Elicit Curiosity

Prior work discussed in Chapter 2 has illustrated promising links between curiosity and the playing of games. The previous chapter examined one form of curiosity (i.e., conceptual) through a case study. However, more forms of curiosity exist and, as of yet, it is unclear how these various forms may be elicited through gameplay or how differently designed games succeed in stimulating curiosity and encouraging exploration. Thus, it is beneficial to take a broader perspective, moving from a single case to the wider landscape of video games, and examine how they elicit curiosity and exploration.

The research question that guides the work in this chapter is:

What types of games elicit exploration?

This chapter describes the setup, execution, and findings of a study designed to *identify game titles and genres that elicit curiosity in video game players*. In addition, it inquires whether there is a *relation between the games that makes players curious and personality-based curiosity traits*. No published work was found addressing these points before this study.

In the study, data is collected from a wide range of video game players about their subjective experience of curiosity with games they have played in the past. Participants are asked to rank games they have played from a list of 15 pre-selected game titles in order of how curious the game made them feel. They are also asked to suggest and rank additional game titles based on different dimensions of curiosity. The study follows the Five-Dimensional Curiosity Scale (5DC) proposed by Kashdan et al. (2018) and investigates links between the individual dimensions and specific games or genres. Besides this, it identifies correlations between games and curiosity dimensions that are the subject of recent psychological studies (Grossnickle 2016; Reio Jr. et al. 2006).

The rationale for exploring such links is that what triggers curiosity in players might depend on their personality traits; on long-term inclinations regarding what elicits their curiosity. While some games might involve multiple design strategies to stimulate curiosity on different dimensions, such as providing interesting landscapes to explore and intricate puzzles to solve, other games might appeal more narrowly to individual curiosity dimensions.

Participants in the study are not asked to apply a specific definition of curiosity or what does or does not constitute a video game. As a result, the study captures a range of subjective associations of the experience of curiosity in video games. Furthermore, given that video games can invoke curiosity in different ways, the study aims to identify consensus and patterns among video game players rather than specific strategies for eliciting curiosity.

The study is designed to answer three research questions (RQs):

1. What games and genres elicit curiosity as part of their gameplay?
2. Do individual differences in trait curiosity dimensions impact what games and genres make a player curious?
3. Does age or gender impact what games or genres elicit curiosity?

The most important outcome of the study is the systematic and informed selection of games and genres that elicit curiosity, described in detail in section 4.4.1 (addressing RQ1). Statistical assessments of the study results do not show strong evidence for an impact of curiosity dimensions on what games and genres elicit curiosity. Some significant correlations were found, e.g., the extent to which role-playing games stimulate curiosity, and are discussed later in section 4.4.2. Overall, however, the study's results do not provide sufficient evidence to answer whether trait curiosity dimensions systematically impact what games and genres elicit curiosity (addressing RQ2). The survey results show that male players are more likely to be curious about *Strategy* and *Task Simulation* games than female players. Players are found to be more curious within *Puzzle* games with increased age (addressing RQ3). The study does not find evidence for different appraisals of individual game titles based on players' age or gender.

The following section describes the method of the study, including the psychometric instrument used to establish curiosity dimensions, and how rankings are established

between game titles that are not played by all participants in the study. This is followed by a description of the study procedure and results. The end of the chapter discusses the findings of the study and their implications for further research about eliciting curiosity through video games.

Chapter Publications

Work presented in this chapter has been published in these peer-reviewed venues:

- **International Conference on Entertainment Computing (ICEC Conference) – 2018**
“Games that Make Curious: An Exploratory Survey into Digital Games that Invoke Curiosity” (M. Gómez-Maureira and Kniestedt 2018)
- **Entertainment Computing (ENTCOM Journal, Volume 32) – 2019**
“Exploring video games that invoke curiosity” (M. A. Gómez-Maureira and Kniestedt 2019)

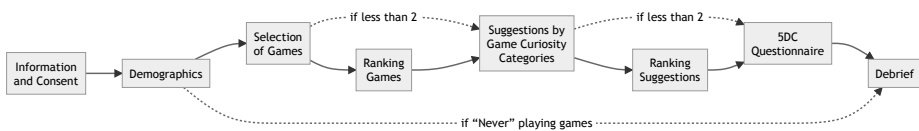
4.1 Materials and Methods

The study uses an online survey format to reach many participants with diverse backgrounds. The survey is specifically aimed at people who have played video games in the past and who are comfortable reading and writing in English. Otherwise, the survey does not have any exclusion criteria. This section describes the individual survey steps and the rationale behind decisions made in the process.

The survey is conceptually separated into the modules: demographics, shared selection of games, suggestions by curiosity dimensions, and curiosity questionnaire.

Within the survey, the modules were presented over separate survey screens. The following subsections describe the individual survey modules and methods used to assess the results.

The survey flow can be seen in the following chart:



4.1.1 Demographics

The demographics module comprises three questions: age, gender, and frequency of playing games.

Data on age and gender are collected to consider whether either impact what game genres elicit curiosity and assess the distribution of the study sample. Psychology research by Giambra et al. (1992) suggests that the need to seek information is not impacted by age, but the need to seek stimulation (i.e., alleviating boredom) is; decreasing with age. They further found differences in genders, with women showing “an increase in impersonal-mechanical curiosity and a decline in interpersonal curiosity” with increased age. In contrast, no such change was found in men. Considering this, there is some evidence that age and gender could have an impact that needs to be logged to analyze the study results.

Furthermore, while this study does not target a specific age range or gender, the sample of the population participating in the survey might be skewed compared to the general (game-playing) population. Data on participants’ age and gender thus indicate whether results can generalize to a broader population or whether caveats need to be considered.

Participants are prompted to provide their age as “year of birth” and their gender by selecting between female and male or entering their gender identification as free-text. A question on playing frequency acts as an exclusion criterium, ending the survey for all participants that do not play video games. It also provides data to assess whether general play frequency impacts what games elicit curiosity in players. Participants are prompted with the statement “I typically play computer or video games ...” that they complete by choosing one of five answers: *Every day*, *Every week*, *Occasionally*, *Rarely*, or *Never*. Choosing *Never* excludes participants from reaching subsequent survey modules and from being part of the study.

4.1.2 Selection of Games

In the second survey module, participants are presented with 15 critically acclaimed game titles and asked to select which of them they have played. The number was picked as a heuristic, balancing the need to involve a sufficiently large number of titles, increasing the chances of participants having played a few of them that they can then rank, and the need to keep participants from becoming overwhelmed by the amount. Especially participants who have played many games from the list would need to spend more time and energy establishing a ranking between the games.

The decision on which games should be included in the list is based on *recency* and *critical acclaim*. The threshold for recency is set to 10 years as an arbitrary threshold but chosen to keep games within the list relatively similar in terms of technical features and game complexity. To illustrate the need for a recency threshold: although *Pong* (Atari 1972) might have been a revolution when it was released, technical capabilities, audio-visual fidelity, and conceptual complexities have increased considerably since then. Ranking of games between different eras could make it more likely to involve nostalgia or recency bias rather than idiosyncratic properties of a game title. At the time of the study, this means only including games that were released between 2007 and 2017.

Critical acclaim is measured based on a game's *Metacritic* score (Metacritic.com 2018), selecting the 15 highest-rated games within the timeframe mentioned earlier (excluding entries from the same game series, as will be discussed below). The resulting selection involved games with a *Metacritic* score of 94 or higher (out of 100). *Metacritic* scores are based on the average scores that are given by a select list of game journalists and critics. Within the video game industry, *Metacritic* scores are often considered a measure of a game's artistic quality and even shape development expectations and hiring practices (GameDeveloper.com 2012). While a high *Metacritic* score is not always predictive of popularity and financial success, there is evidence of a strong correlation (Greenwood-Ericksen, Poorman, and Papp 2013). As such, it is a reasonable proxy measure to establish game titles that are both widely known and known for high production quality. This should make it more likely for participants to have played games on the list and experienced curiosity as part of their experience.

While the selection of games list counts 15 entries, the number of games reflected in it is higher because multiple games within a game series are grouped under a single entry. A *game series* is defined by a shared cast of game characters, a somewhat consistent audio-visual aesthetic, related narrative arcs, and many similar game mechanics, and typically refers to prior entries in some way as part of the title. The decision to group games from a series into a single entry was taken to involve a diverse selection of different games while keeping ranking easy to understand for participants. For example, participants might have played only some titles in a game series that are, however, sufficiently similar to one another for the context of the study. Grouped entries do not stand for all game titles within a series but are limited to those that fit the selection criterium outlined before, i.e., released within the same 10-year timeframe and with similarly high *Metacritic* scores. As such, some entries in the list stand for a single game, while others include a range of game titles. Overall, the 15 entries in the list implicitly include 27 individual game titles (table 4.1).

Table 4.1: List of game selection with corresponding implicitly included game titles (prompted by “Select which of the following games you have played ...”)

| List Entries | Implicitly Included Game Titles |
|---|--|
| Grand Theft Auto IV (or newer) | <i>Grand Theft Auto IV</i> (2008) and <i>Grand Theft Auto V</i> (2013) |
| The Legend of Zelda: Breath of the Wild | Single title, as mentioned (2017) |
| Super Mario Galaxy or Super Mario Odyssey | <i>Super Mario Galaxy</i> (2007) and <i>Super Mario Odyssey</i> (2017) |
| Batman: Arkham City (or Asylum) | <i>Batman: Arkham Asylum</i> (2009) and <i>Batman: Arkham City</i> (2011) |
| The Elder Scrolls V: Skyrim | Single title, as mentioned (2011) |
| Mass Effect (any in the series) | <i>Mass Effect</i> (2007), <i>Mass Effect 2</i> (2010), and <i>Mass Effect 3</i> (2012) |
| Uncharted (any in the series) | <i>Uncharted: Drake's Fortune</i> (2007), <i>Uncharted 2: Among Thieves</i> (2009), <i>Uncharted 3: Drake's Deception</i> (2011), <i>Uncharted 4: A Thief's End</i> (2016) |

| List Entries | Implicitly Included Game Titles |
|--|---|
| BioShock (any in the series) | <i>BioShock</i> (2007), <i>BioShock 2</i> (2010), and <i>BioShock Infinite</i> (2013) |
| Metal Gear Solid V | <i>Metal Gear Solid V: The Phantom Pain</i> (2015) |
| The Last of Us | Single title, as mentioned (2013) |
| Portal (or Portal 2) | <i>Portal</i> (2007) and <i>Portal 2</i> (2011) |
| Red Dead Redemption | Single title, as mentioned (2010) |
| LittleBigPlanet (or LittleBigPlanet 2) | <i>LittleBigPlanet</i> (2008) and <i>LittleBigPlanet 2</i> (2011) |
| Call of Duty: Modern Warfare (or Modern Warfare 2) | <i>Call of Duty 4: Modern Warfare</i> (2007) and <i>Call of Duty: Modern Warfare 2</i> (2009) |
| Street Fighter IV | Single title, as mentioned (2008) |

4.1.3 Ranking Games

After selecting games, participants are presented with the sub-selection of games they have played in the past. Participants are prompted to “Rank the games in order of how much they triggered your curiosity while playing them (most curious on top) – leave out games that did not make you curious at all while playing.” With these instructions, the list of ranked game titles is expected to be not only fewer than 15 (given that few participants will have played all games on the list) but also fewer than all titles they have played, as they can choose not to include titles that have not elicited much curiosity.

It should be noted that the study emphasizes curiosity as part of the gameplay rather than how curious participants might have been to play a game title before doing so. This emphasis is reflected in the phrasing of the prompts in the study.

Asking participants to rank rather than score game titles, for example, on a Likert scale, regarding the curiosity that was experienced brings some benefits. For one, the task is easily explained and requires less time. Furthermore, reporting about affective constructs is challenging, and applying a rating consistently can be particularly difficult (Yannakakis and Martínez 2015). Ranking allows participants to use individual game ti-

titles as points of reference instead, thereby not having to make assumptions about the meaning of choosing, for example, 4 out of 10 on a curiosity scale.

However, ranking introduces some challenges that would not be present with Likert scale ratings. Ranking does not provide the possibility to capture quantitative information. Two game titles might be very close in their impact on a participant or very far apart; the collected data would not provide any evidence as to which is the case. Participants are also forced to rank titles they consider similar regarding experienced curiosity, as they cannot give two titles the same rank. Another challenge comes from the analysis of ranked game titles when not all titles are ranked, and the amount of ranked items is not the same across participants. Not all people play the same game titles. Likewise, participants may rank varying numbers of games, either because they do not play as many games or because they do not consider them to be invoking curiosity.

These challenges are likely, at least in part, reasons for why ratings remain part of many studies to assess affective constructs, despite the criticism that can be leveled against their use. For this study, the benefits of ranking game titles are considered to outweigh the drawbacks. Forced rankings of titles that elicit curiosity equally well should normalize across a large sample size. For the most significant challenge, evaluating rankings across participants, the study took inspiration from a similar challenge in game ranking systems where assessing player strength results from multiple, usually unequal amounts of games played. The following sub-section discusses ranking analysis in more detail.

4.1.4 True Skill Rank Analysis

Ranked results in the study are analyzed using the *TrueSkill* rating system, developed by *Microsoft* for ranking and match-making of players on their *Xbox LIVE* online platform (Herbrich, Minka, and Graepel 2007).

One challenge ranking systems in multiplayer games face is determining the relative strength of players in zero-sum competitions. Simply counting the number of wins neglects the context of how those wins were achieved. A player might have played only against relatively inexperienced opponents and thus notched a high track record that does not reflect their skill level. It should be evident that a hypothetical newcomer playing and winning against a world champion should rank higher than a simple tally

of successes would indicate. Multiple rating systems have been developed to assess the relative strengths of players. Examples include *Elo*, created to rate the performance of chess players (Elo 1978), *Glicko* (Glickman 1999), and *TrueSkill*, which has also been used outside of games in education (Kawatsu, Hubal, and Marinier 2018) and to improve recommendation systems (Quispe and Ochoa Luna 2015). Recent work carried out after the present study has been completed suggests broadly similar accuracy between these rating systems, especially with increasing match-up data points (Dehpanah et al. 2021).

TrueSkill uses a *Bayesian inference algorithm* that updates the score of individual match items (usually representing the skill of players) every time a match is played. Since score points can be lost, participating in a high number of matches does not necessarily result in a higher ranking.

Conceptually, the analysis of ranked game titles in this study is approached as if it were a competition between 15 players. Each game title represents a player that competes against other players, i.e., game titles. For each participant in the study, each ranked game title competes against all other ranked game titles. If a title is ranked higher, the title “wins” and thus increases its *TrueSkill* rating, while the “opponent” has its rating reduced. The higher a title is ranked in curiosity, the more wins it accumulates against other opponents.

After matching up all possible combinations across all participants, the resulting score is a measure of both the rank of a game title and the relative distance to other game titles. While the resulting score is an arbitrary number, it can be used in relation to other scores. Items with relatively similar scores can then be considered closer to equal, while those that differ by wide margins are likely to have won many comparisons.

In addition to analyzing rankings from the pre-established game selection, the same method is used for game titles that participants suggest. Here, participants can suggest up to ten game titles that made them curious and are then asked to rank them. *TrueSkill* scores are calculated for game titles suggested by participants (section 4.1.6) and for game genres that these games are part of (section 4.1.7).

4.1.5 Five-Dimensional Curiosity Scale (5DC) Questionnaire

The 5DC questionnaire, developed by Kashdan et al. (2018), is used to determine the distinctive “trait curiosity” of each participant. Trait curiosity describes the general tendency to be curious rather than whether or not a person is in a curious state (i.e., is currently curious or not). The individual dimensions outlined in the 5DC describe what stimuli (or lack thereof) are most likely eliciting a curious state. Whereas some people are most motivated by the desire to experience novelty, others might be driven by a desire to mitigate a lack of information.

More specifically, the 5DC consists of the following five dimensions:

1. **Joyous Exploration** (JE) — being motivated by novelty
2. **Deprivation Sensitivity** (DS) — experiencing a need to resolve
3. **Stress Tolerance** (ST) — the ability to cope with uncertainty
4. **Social Curiosity** (SC) — wanting to know about others
5. **Thrill Seeking** (TS) — enjoyment of anxiety

The questionnaire has been developed by selecting items of existing measures that evaluate interest and curiosity, openness to experience, need for cognition, boredom proneness, and sensation seeking. Individual questionnaire items were evaluated through three studies with a combined sample size of 3911 participants. The questionnaire was examined regarding test-retest reliability in a 4-month follow-up, with results within the range of stable personality traits.

The 5DC questionnaire consists of 25 statements for which participants indicate agreement on a 7-point Likert scale. The scale ranges from “Does not describe me at all” (1) to “Completely describes me” (7). The 25 statements are grouped into five scoring groups, corresponding to the individual curiosity dimensions, each of which includes five statements. The 5DC questionnaire results in scores for each dimension, calculated by the average score of statements within the scoring group (with items contributing to *Stress Tolerance* being reverse-scored).

The questionnaire is used at the end of the survey. It is discussed here, outside the chronological order of survey modules, to establish terminology used in describing other survey modules.

4.1.6 Suggestions by Game Curiosity Categories

After ranking among the 15 game titles selected for the survey, participants are asked to suggest additional game titles that fit the five curiosity dimensions of the 5DC (section 4.1.5). These suggestions are collected as free text responses and are not limited to a specific release year or other restrictions.

This survey step is meant to collect game titles that can be further analyzed for their ability to invoke curiosity, thus contributing to RQ1. By prompting for suggestions relating to each of the five curiosity dimensions, game titles that are suggested are implicitly grouped into *game curiosity categories* mirroring the respective curiosity dimensions.

Participants are asked to suggest up to two game titles for a total of up to 10 suggestions for each category. Categories for which to suggest games are phrased based on the 5DC questionnaire (table 4.2).

Table 4.2: Prompts used to group suggested game titles in five game curiosity categories.

| Category Prompt | Category Label |
|---|---|
| “Video games that ...” | (based on 5DC) |
| “let me explore or find out new things” | → GEXP (based on <i>Joyous Exploration</i>) |
| “let me solve something” | → GSOL (<i>Deprivation Sensitivity</i>) |
| “let me feel safe and stress-free” | → GSAF (<i>Stress Tolerance</i>) |
| “let me understand or connect to people” | → GCON (<i>Social Curiosity</i>) |
| “make me feel excited and alive ” | → GALI (<i>Thrill Seeking</i>) |

In addition to suggesting game titles for each category, participants also rank their own suggested titles in order of how curious they felt while playing them, similar to the ranking discussed in section 4.1.3. Participants are asked to rank across all games (up to 10) they have suggested rather than per category. In contrast to the ranking of game titles provided by the study, the ranking of suggested game titles is assessed based on the game genres that the suggestions are part of. This is further described in the following sub-section.

4.1.7 Attribution of Game Genres

Game titles suggested by participants in the study are grouped by game genre. It can be expected that many suggestions are mentioned by only 1-2 participants and thus lack a broader consensus. This, however, is likely impacted by how well-known these games are. Even if these titles are ranked high in the curiosity they elicited, they would be part of too few conceptual match-ups to meaningfully increase their *TrueSkill* score (section 4.1.4). To derive useable information, suggested game titles are attributed to two game genres. The resulting grouping of games by their genres can then be assessed through *TrueSkill* scores.

The challenge in attributing game genres is the lack of a broadly shared definition of what constitutes a genre. Genre classifications can originate from multiple motivations, such as easing retrieval of titles, academic efforts to build a taxonomy or marketing considerations. Clarke et al. (2017) argue that genre definitions for video games are a combination of facets, such as gameplay, purpose, presentation, in-game point of view, or theme, to name a few examples. For example, the colloquially common genre of “action games” hints at fast-paced activity in a game but does not indicate a thematic setting or context. Taking control over a sports car in a game is generally a fast-paced activity but would be more commonly referred to as a “racing game”.

Instead of following a specific genre definition, a list of 11 game genres has been devised to describe the most prevalent gameplay activities among suggested game titles. The list is based on commercially used genre labels but modified to be sufficiently general to stand for a range of different games. Instead of a shooter genre, which is typically defined by the competitive use of virtual guns, such a game is labeled as a *Reflex* game to indicate that success in the game is based on fast player reflexes. A racing game is instead labeled as a *Reflex* and *Task Sim* game, the latter of which indicates that players simulate tasks that are associated with a profession. This example also illustrates why game titles are attributed to 2 game genres, as video games frequently involve a combination of activities.

Many games that aim to entertain a large audience involve many activities for the player, optional or mandatory, to play a game to its conclusion. *Grand Theft Auto V* (Rockstar North 2013) lets players shoot virtual characters and race with cars but also allows them to ride a roller-coaster, attend virtual therapy sessions, or solve a mur-

der mystery. In this example, the genres *Reflex* and *Exploration* are attributed to the game. The genres reflect that most activities in the game require fast reflexes and that a significant appeal of the game is in exploring the city and all activities that the game facilitates. While some nuance is lost by converting game titles into game genres, it allows for evaluating which activities in a game can be conducive to eliciting curiosity.

The game genres used in this study are based on collecting commonly used genre labels from online game stores and critic score aggregators. The list has been modified and condensed to a subset that is sufficiently general, descriptive, and independent of other labels. The final list is the result of argumentative discourse between the author and other game researchers and designers.

The following list outlines the genre labels, followed by a descriptor:

- **Reflex** — requires fast reflexes to perform well.
- **Exploration** — provides spatial or conceptual discovery that is not automatically brought to the player's attention.
- **Puzzle** — presents tasks that must be solved through predefined processes.
- **Strategy** — requires players to plan their actions, considering available resources.
- **RPG** — defined by assuming the role of one or more characters and making choices that impact game progression.
- **Story** — game progresses as part of a structured narrative.
- **Task Sim** — asks players to perform tasks associated with professions, emphasizing the nature of the task.
- **Social Sim** — asks players to perform actions associated with social interactions and everyday tasks.
- **Collecting** — is structured around gathering items to gather all or as many items as possible.
- **Frantic** — uses aesthetic elements, concurrent game mechanics, or both to saturate the cognitive capabilities of players.
- **Chance** — progress in the game is largely independent of the actions taken by the player but differs between game sessions.

It should be noted that the attribution of game genres to game titles is not dependent on how many activities a game features; all titles are given precisely two genre labels.

4.2 Procedure

This section describes the logistical details of conducting the survey and analyzing its results. The survey was conducted over one month and was completed by 117 participants. All responses were collected through the online survey system *Qualtrics* provided by the University.

Participants were recruited through *convenience sampling* and *referral sampling* within the University of the author, academic mailing lists, and online recruiting on *Facebook*, *Twitter*, and *Reddit*. The survey was also part of the survey exchange platform *Survey-Tandem*, which promotes a survey in exchange for participation in the survey studies of other researchers. The target demographic included everyone who does or has played video games at some point in their lives. While the nationality of participants was not tracked, the most likely audience was English-speaking people in the Netherlands and, to a lesser extent, in western Europe and the United States.

The survey led participants through 8 modules split over multiple web pages, with the 5DC questionnaire module broken up over several pages to aid readability.

Apart from the modules described in *Materials and Methods* (section 4.1), the survey was prefaced by an information and consent module. On that page, participants were informed about the study's goal and the experimenters' contact information. To proceed, participants had to consent to have their responses recorded and stored on the University's *Qualtrics* server. The last module of the survey was a concluding debrief step that thanked participants for their time and repeated the contact information in case of further questions.

The survey flowchart shows that not all participants went through all modules. For one, participants who indicated that they do not play games were brought to the end of the survey, and their data was not used. The two modules asking participants to rank among previously selected or suggested game titles are only shown if there are titles to rank (i.e., when selecting two or more game titles).

4.2.1 Data Processing

After the survey concluded, the collected data had to be processed for further assessment. Incomplete responses were removed from the dataset.

Initial Ranking Calculation Initially, data processing of rankings was carried out without using the *TrueSkill* algorithm. It was only in assessing the collected data that a more specialized assessment method was sought. Initially, rankings were calculated by tallying how often a game title was played across all participants and using that number to adjust the weight in any ranking it was used in.

For each participant and game title, the following formula was used to calculate a title's curiosity score:

$$\text{curiosity score} = \frac{1 - \left(\frac{\text{rank}}{\text{rankcount}}\right) + \left(\frac{1}{\text{rankcount}}\right)}{\text{playcount}_{\text{all}}}$$

The formula weighs the score based on how many game titles a participant has ranked (rankcount), at which position it is ranked (rank), and how many participants have played the game (playcount_{all}). The score increases, the higher it is ranked, the fewer items a participant ranks, and the fewer participants have played the game title. The rationale for this weighing is that game titles played by many will be featured in more rankings and therefore accumulate a higher score based, at least in part, on popularity rather than curiosity. By using this formula, a game played by many will need to be consistently ranked high to end up with a high curiosity rank score.

The overall curiosity score of a game title is the sum of all scores calculated for each participant.

TrueSkill Calculation For the ranking of selected games provided by the survey, *TrueSkill* ratings were calculated with the help of a Python package (Lee 2015). The 15 game titles had their *TrueSkill* rating initialized to be equal to one another. The script then iterated over all participant responses and, within each response, repeated over all unique pair-wise combinations of ranked game titles. For participants that had ranked among all 15 titles, this would result in 105 unique combinations and, therefore, 105 match-ups, with one game title winning over the other. The actual number of such matches depended on how many titles they ranked; in most cases, that number was lower. Roughly half of all participants ranked four game titles, which results in

only six pair-wise matches for each. The total number of these matches was the sum of all participants that ranked at least two titles. For each match, the *TrueSkill* rating of both competing game titles was adjusted based on the match result. At the end of the process, each game title had a *TrueSkill* rating based on all participants' rankings.

Processing of Free Text Entries Before the *TrueSkill* calculation process could be repeated for game genres based on game titles suggested by participants, the entries had to be checked for duplicates and valid entries. Since game titles could be entered as free text, the resulting list could include different spellings of the same game titles. As a first step, a “fuzzy” text-matching Python package (SeatGeek 2017) was used to create a dictionary of unique entries. The algorithm uses the *Levenshtein distance*, a string metric that assesses the difference between two letter sequences, to calculate a percentage of how closely two strings match each other.

The algorithm identified text strings matching an existing entry with 70% accuracy (arbitrarily set based on experimenting with different values). Successful matches were grouped under a single title, while unsuccessful matches created a new title. All groupings made this way were logged and manually checked. The dictionary was manually expanded for any mismatched items, and the process was repeated until all “fuzzy” matched items were correctly grouped. The dictionary was manually checked for game titles that did not exist. Game titles that could not be identified through a subsequent web search were discarded from the dictionary.

As a next step, entries belonging to the same game series, or referring to the same game by another name, were combined into a single entry, e.g., *Oblivion* (Bethesda Game Studios 2006) was attributed to *Elder Scrolls*. The only exceptions were *The Legend of Zelda: Breath of the Wild* (Nintendo EPD 2017) — shortened to *Zelda:BotW* (distinct from *The Legend of Zelda* game series) and *World of Warcraft* (Blizzard Entertainment 2004) — shortened to *WoW* (distinct from the *Warcraft* game series). This process was carried out manually by consulting online resources such as game descriptions, reviews, and gameplay videos. The decision not to consider *Zelda:BotW* and *WoW* as regular entries in the rest of their respective game series is based on the assessment that these games are sufficiently distinct to warrant their own entries. In the case of *Zelda:BotW*, the game marked a transition to an open-world gameplay environment, whereas other games in the series are more linear. For *WoW*, despite sharing a similar name to other games in

the *Warcraft* game series, *WoW* marked the transition from *real-time strategy* to what is known as a *massively multiplayer online game*. As a result, these exceptions were attributed to different game genres (as defined in section 4.1.7) than other entries in the respective game series.

Combining game titles of a series into a single item removes nuances that individual game titles might have brought to a game series. For the most part, however, games of a series intentionally share characteristics such as theming and gameplay. Making this decision allowed for assessing general patterns in game design, expressed through attributed game genres.

Processing of Game Genres After processing free text entries, the resulting list of game titles was converted to a look-up Python dictionary with corresponding game genres that could be used for further processing. Exactly two genres were attributed to each game title, choosing the most representative labels from a list of 11 genres (outlined in section 4.1.7). The first genre was chosen to be the most representative, with the second indicating the secondary genre. Genres were attributed independently by the author and one of their peers from the game research and design domain. Attributions were made based on personal familiarity with game titles and further desk research involving developer descriptions, reviews, and gameplay videos. Afterward, the two independently created dictionaries were compared for mismatches in choice of genres and their order. Mismatches were resolved item by item by discussing arguments between the author and the supporting peer until they reached an agreement.

Once the dictionary was completed, *TrueSkill* scores were calculated for all game genres. In contrast to the calculation for the selection of 15 game titles, scores for genres were based on a maximum of 10 suggestions per participant (up to 2 per curiosity category). This resulted in a maximum of 20 ranked genres (2 per game title), with individual genres being featured on multiple ranks (e.g., *Exploration* could be ranked 1st, 4th, and 8th simultaneously). Each genre instance was then matched against all other instances. In this case, the same genre pairing could lead to different winners since they would be representing different game titles. Matches between the same genres were not carried out.

Processing Game Curiosity Categories and 5DC Questionnaire A ranking of curiosity categories was calculated in addition to ranking game titles and genres. In this case,

the ranking was established by replacing the ranked game title suggestion with the curiosity category for which the game title was entered. Similar to the ranking of game genres, each participant could implicitly rank a maximum of 10 curiosity categories, which would feature the same entries on up to two rank slots. For example, game titles entered under the prompt “Video games that let me *explore* or find out new things” (GEXP category) could be ranked 1st and 3rd, thus leading to GEXP being featured twice at different ranks.

Scores of the 5DC questionnaire were created for each participant by calculating the mean of Likert scale ratings of questions contributing to the corresponding five dimensions. Ratings were reverse scored for the *Stress Tolerance* (ST) dimension, as required by the questionnaire instructions.

4.3 Results

Out of 117 participants completing the survey, $N=113$ reported playing video games and thus represent the valid sample of the study. The mean age of participants was $Mn=27.6$ ($SD=5.8$), with 44 identifying as female (38.9%) and 69 identifying as male (61.1%). Converted to Likert scale ratings, the mean frequency of playing games was $Mn=2$ ($SD=0.94$), equating to playing games “Every week”. Statistical tests use a significance level of 0.05 in this study.

Playing frequency was found to differ between genders (Mann-Whitney $U=1987$, $p=0.004$, two-tailed), with male participants playing more frequently than female participants.

Selection of Games On average, participants ranked $Mn=6.5$ ($SD=5.9$, $Mdn=3$) of the selection of 15 game titles, with $n=96$ ranking at least two titles and $n=2$ ranking all of them. The most frequently played game title was *Portal* ($n=66$), while the least played was *Metal Gear Solid V* ($n=19$).

As previously discussed, the selection of games was processed first through a weighted sum of rankings and later through the *TrueSkill* algorithm. An overview of the results is listed in table 4.3. The table illustrates that scoring based on the unweighted sum of ranks is closely tied to the playcount. Using a weighted sum as a curiosity score introduces several changes in the ranking, most notably by pushing *Zelda:BotW* to the

first rank by a considerable margin. The ranking results of the *TrueSkill* algorithm shifts some of the middle ranks around, but is otherwise close to the weighted scores.

Given that the ranking results of the *TrueSkill* scores are methodologically more sound (see above), assessment of rankings is based on these.

Table 4.3: Comparison of ranking results based on playcount (how many participants have played the game), sum of unweighted ranks, sum of weighted ranks, and *TrueSkill* score ranking. All values are normalized to a 0 to 1 scale for better comparison. A value of 1.00 indicates the highest value of a given column, while 0.00 indicates the lowest.

| Game Title | Playcount | Unweighted | Weighted | TrueSkill |
|--------------------------------|-----------|------------|----------|-----------|
| Portal (1 & 2) | 1.00 | 1.00 | 0.72 | 0.64 |
| Elder Scrolls: Skyrim | 0.94 | 0.97 | 0.79 | 0.71 |
| Grand Theft Auto IV (or newer) | 0.79 | 0.63 | 0.38 | 0.37 |
| BioShock (any) | 0.66 | 0.66 | 0.55 | 0.52 |
| Mass Effect (any) | 0.57 | 0.54 | 0.60 | 0.62 |
| Super Mario Galaxy / Odyssey | 0.53 | 0.48 | 0.47 | 0.58 |
| Call of Duty: MW (1 & 2) | 0.43 | 0.22 | 0.11 | 0.14 |
| Uncharted (any) | 0.32 | 0.32 | 0.40 | 0.37 |
| LittleBigPlanet (1 & 2) | 0.30 | 0.23 | 0.29 | 0.42 |
| Batman: Arkham Asylum / City | 0.28 | 0.31 | 0.50 | 0.51 |
| The Last of Us | 0.28 | 0.35 | 0.59 | 0.53 |
| Red Dead Redemption | 0.19 | 0.24 | 0.47 | 0.51 |
| Zelda: Breath of the Wild | 0.15 | 0.35 | 1.00 | 1.00 |
| Street Fighter IV | 0.06 | 0.00 | 0.00 | 0.00 |
| Metal Gear Solid V | 0.00 | 0.09 | 0.47 | 0.54 |

Suggestions by Game Curiosity Categories For the game suggestions per category module, a total of 301 unique game titles were mentioned, out of which 136 were suggested by at least 2 participants. The top 10 suggestions were: *Elder Scrolls* (suggested

Chapter 4. Video Games That Elicit Curiosity

by **39**), *Portal* (**36**), *Zelda:BotW* (**33**), *WoW* (**28**), *Fallout* (**22**), *Minecraft* (**18**), *GTA* (**18**), *Horizon: Zero Dawn* (**17**), *Final Fantasy* (**17**), and *The Sims* (**16**). A breakdown of suggested game titles per game category is shown in table 4.4.

Table 4.4: Game titles mentioned for each of the five curiosity categories (showing titles with at least 5 mentions). Titles in bold appear in multiple categories.

| Game Category | Game Titles (Number of Mentions) |
|--|---|
| GEXP (“Explore”, “find out”) 92 unique titles | Elder Scrolls (17) , Fallout (14), Minecraft (11), Zelda: BotW (9), Dark Souls (8), Horizon: Zero Dawn (8) , The Witcher (8), Subnautica (7), World of Warcraft (7) , Final Fantasy (5) , Assassin’s Creed (5), Zelda (5) |
| GSOL (“Solve things”) 113 unique titles | Portal (29), The Witness (8), Elder Scrolls (7) , Myst (5), The Talos Principle (5) |
| GSAF (“Safe”, “stress-free”) 100 unique titles | Sims (8) , Stardew Valley (7), Elder Scrolls (6) , Cities: Skylines (5) |
| GCON (“Connect to people”) 84 unique titles | World of Warcraft (13) , Final Fantasy (7) , Journey (5), Sims (5) |
| GALI (“feeling alive”) 108 unique titles | GTA (7), PlayerUnknown’s Battlegrounds (6), World of Warcraft (5) , Horizon: Zero Dawn (5) , Elder Scrolls (5) |














Ranking of the attributed game genres by *TrueSkill* scores results in the following normalized ranking (curiosity rank, count rank): *Social Sim* (**1.00, 0.08**), *Collecting* (**0.87, 0.10**), *RPG* (**0.80, 0.42**), *Exploration* (**0.76, 0.75**), *Story* (**0.58, 0.33**), *Task Sim* (**0.54, 0.17**), *Chance* (**0.42, 0.00**), *Strategy* (**0.36, 0.49**), *Reflex* (**0.29, 1.00**), *Puzzle* (**0.18, 0.52**), and *Frantic* (**0.00, 0.02**). Counts of suggestions per game genre ranged from **10** for *Chance* (normalized to **0.00**) to **318** for *Reflex* (normalized to **1.00**).




TrueSkill scores of game categories result in the following normalized ranking (curiosity rank, count rank): **GEXP (1.00, 1.00)**, **GALI (0.32, 0.17)**, **GCON (0.16, 0.01)**, **GSOL (0.02, 0.74)**, and **GSAF (0.00, 0.00)**. Counts ranged from **103** for **GSAF** (normalized to **0.00**) to **180** for **GEXP** (normalized to **1.00**). Of note is that both **GCON** and **GEXP** had fewer unique suggestions than **GSAF**, even if **GSAF** had the lowest total suggestions.

Ranking of game genres was found to differ by gender for the genres *Strategy* (Mann-Whitney $U=1911$, $p=0.002$, two-tailed, lower ranking in females) and *Task Sim* ($U=1714$,

$p=0.036$, two-tailed, lower ranking in females). Age was found to correlate with a higher ranking of the *Puzzle* genre (Spearman's $\rho=0.226$, $p=0.019$) and a lower ranking of the game curiosity category GSAF ($\rho=-0.231$, $p=0.018$).

Table 4.5: Two-tailed Spearman's rank correlations between 5DC dimensions and other measures. VS-MPR shows the maximum possible odds in favor of a hypothesis; i.e. of correlations between measures.

| 5DC Dimension | — | Measure | ρ | p | VS-MPR |
|------------------------------|---|--|--------|-------|--------|
| JE (Joyous Exploration) | — |  RPG | 0.2 | 0.037 | 3.019 |
| | — |  GCON ('connect with people') | 0.234 | 0.016 | 5.617 |
| DS (Deprivation Sensitivity) | — |  Collecting | -0.193 | 0.045 | 2.655 |
| | — |  GTA (IV+) | 0.252 | 0.007 | 10.607 |
| ST (Stress Tolerance) | — |  Zelda: BotW | -0.239 | 0.011 | 7.499 |
| | — |  Call of Duty | 0.214 | 0.023 | 4.295 |
| | — |  RPG | 0.293 | 0.002 | 29.545 |
| | — |  Puzzle | -0.226 | 0.018 | 5.037 |
| | — |  GSOL ('solve something') | -0.222 | 0.022 | 4.310 |
| SC (Social Curiosity) | — |  Social Sim | 0.220 | 0.022 | 4.414 |
| | — |  Frantic | -0.212 | 0.03 | 3.535 |
| TS (Thrill Seeking) | — |  GTA (IV+) | 0.279 | 0.003 | 22.456 |
| | — |  RPG | 0.230 | 0.016 | 5.515 |

 Individual Game  Genre  Game-based Curiosity Dimension

Assessment of the 5DC Questionnaire The aggregated results of the 5DC questionnaire were: *Joyous Exploration* (JE) $Mn=5.38$, $SD=0.86$; *Deprivation Sensitivity* (DS) $Mn=4.98$, $SD=1.15$; *Stress Tolerance* (ST) $Mn=4.36$, $SD=1.42$; *Social Curiosity* (SC) $Mn=5.11$, $SD=1.14$; and *Thrill Seeking* (TS) $Mn=4.20$, $SD=1.34$ — each based on Likert scale ratings from 1 to 7.

Demographic differences were found by gender, with ST being significantly higher in male participants (Mann-Whitney $U=978$, $p=0.001$, two-tailed), while SC was sig-

nificantly higher in female participants ($U=1988$, $p=0.006$, two-tailed). Participants' age was found to be correlated with a lower score of SC (Spearman's $\rho=-0.297$, $p=0.001$).

Significant correlations between 5DC dimensions and rankings are shown in table 4.5. For clarity, ρ was inverted to match the meaning of an increase in the score of 5DC dimensions. This means that a rating of 1 in a ranking is higher than a 2, but 1 is lower than 2 in the 5DC questionnaire.

4.4 Discussion

This section interprets the survey results in the context of the study's three research questions. The most relevant results are visualized and repeated.

4.4.1 Game Titles and Genres That Elicit Curiosity

The primary goal of this study was to establish game titles and genres that have elicited curiosity as part of their gameplay. Through the responses of 113 participants playing video games, a list of popular games was ranked in terms of the curiosity they elicited. Additional game titles were collected and attributed to genres that were defined for this study.

Assessment of Game Titles The ranking of the 15 game titles selected for the survey is summarized in figure 4.1. Among these, *Zelda:BotW* stands out as being ranked highest in eliciting curiosity, by a margin of 40% above the 2nd rank, *Elder Scrolls: Skyrim*. This is despite the fact that the game has been played by only 23% of all participants, compared to 56% for *Elder Scrolls*. This means that when participants did play the game, it was likely to rank high in eliciting their curiosity.

Games in the list can be roughly positioned on a spectrum between games in which players have a great deal of freedom in choosing how to navigate the game world and games that restrict the movement options more narrowly. Describing the games from this perspective, *Zelda: BotW*, *Elder Scrolls*, *Metal Gear Solid V*, *Red Dead Redemption*, and *GTA IV+* give the player an enormous game world to freely explore. As part of the narrative, movement options might open up over time or are restricted for a predefined duration. However, overall, these games can be considered part of "open-world

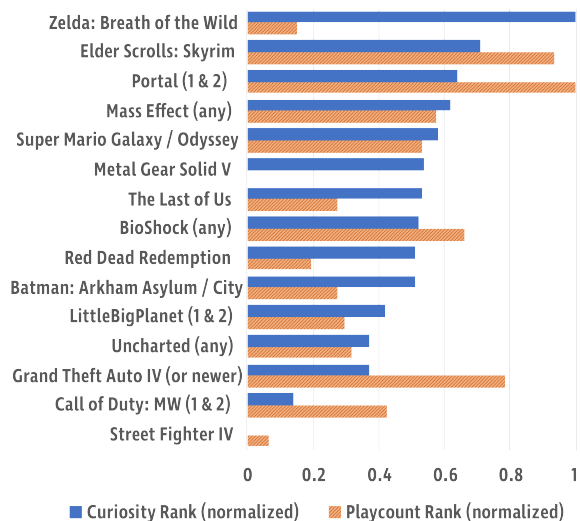


Figure 4.1: Shared selection of games ranked by how curious participants felt and how many participants had played them. Values are normalized to 0-1 for comparison (0 for lowest rank and 1 for highest rank).

games”, a genre label that is not strictly defined by specific gameplay but rather a combination of activities and environment properties. *Zelda: BotW* is unique among other open-world games in that players face almost no boundary that cannot be overcome. For example, players can climb on virtually all surfaces, and most of the activities in the game are optional and non-linear: players could choose to engage with the game’s final challenge (combat against the series’ adversary, “Ganon”) almost from the beginning of the game.

In the middle of the spectrum are games with free-roaming sections (game levels) but requiring players to fulfill objectives within that section to progress to new areas. The exploration space within such levels can be reasonably large, but movement between levels tends to be limited or facilitated through a hub, such as a menu screen or overview environment (e.g., a game’s world map). This is the case for the games *Portal*, *Mass Effect*, *Super Mario Galaxy / Odyssey*, *The Last of Us*, *BioShock*, *LittleBigPlanet*, *Uncharted*, and *Call of Duty*. *Batman: Arkham Asylum / City* can be considered an edge

Chapter 4. Video Games That Elicit Curiosity

case, with *Arkham Asylum* being closer to segmented free-roaming and *Arkham City* being closer to an open-world environment.

On the other end of the spectrum is *Street Fighter IV*, a martial arts fighting game in which players face another opponent within a tightly spaced area.

Considering the ranking, it appears that curiosity in the context of gameplay was primarily assessed in terms of the ability to navigate the game environment; or, in other words, spatial exploration. It is essential to note that this interpretation of the results cannot be confirmed through the data collected in the study. It is a possible explanation for why some games have ranked higher than others, but it requires further investigation.

StreetFighter could have been interpreted as a source of curiosity for how other people react, i.e., social curiosity. However, this form of curiosity did not appear to have been at the forefront of participants' consideration when ranking the 15 game titles. Notably, *RPG* games such as *Elder Scrolls* and *Mass Effect* provide freedom not only in navigating space but also by impacting how a game's narrative plays out based on their interactions with other (non-player) characters. Here, the source of curiosity might well be based on spatial, conceptual, and social exploration. To recall, as mentioned in section ??, each of these three describes different domains of exploration: either traversal of physical space (made virtual within video games) in spatial exploration, cognitive interpretation of information in conceptual exploration, and investigation of the intentions and behaviors of others in social exploration. To what extent each of these domains is represented in *RPGs* is difficult to determine, and perhaps it is, in fact, the involvement of all three domains that make the genre rank high in eliciting curiosity (as will be discussed in the next section).



Figure 4.2: Screenshots from the games *Zelda: Breath of the Wild*, *Elder Scrolls: Skyrim*, and *Portal*, the three highest ranked game titles.

Portal games are noteworthy in their ranking in the study, as the gameplay is confined to a relatively claustrophobic series of rooms in which players need to solve spatial puzzles. Curiosity in *Portal* is likely elicited by a combination of spatial and conceptual exploration. Despite the relatively small game environments, the game allows players to connect spaces with the “portal gun”, a device that creates wormhole-like connections between two locations in the game. Through this device, the nature of spatial exploration is very different compared to other games. It is tightly integrated into a series of cognitive challenges that must be overcome to progress in the game. The game also stands out in its narrative, which surprises the player with changes about the nature of the game, ultimately provoking them to break out of the confines of the series of puzzle rooms. In the process, players can see behind the game spaces that appear to make up the boundary of the game. This is perhaps best compared to a movie that, halfway through its runtime, pulls back from the apparent narrative and reveals behind-the-scene stages of earlier scenes as part of the actual narrative.

The relatively low ranking of *GTA IV+* suggests that the ability to explore large game environments is not directly correlated to how curious players feel when playing. Games in the *GTA* series involve vast open worlds and various activities that players can pursue (e.g., car racing, bank heists, riding a roller coaster, and even participating in a virtual yoga session). A possible explanation for a lower ranking could be that both *GTA IV* and *GTA V* have tightly scripted narratives, with freedom given primarily between tasks that players are supposed to follow to progress. The majority of these tasks and other activities in the game are further focused on gun combat and car racing. This explanation of why it did not elicit much curiosity might be confirmed by the observation that *Red Dead Redemption*, made by the same developer, ranked higher. This game is structured very similarly to *GTA* games but provides a novel take on the *GTA* game structure. In *Red Dead Redemption*, players find themselves in a wild west environment, traversing the world on horseback rather than in a car. The game emphasizes a romanticized gunslinger atmosphere and uses a morality system in which the player’s actions influence how other characters interact with the main character, possibly translating into a higher degree of perceived freedom. However, whether or not these explanations account for how these games elicit curiosity can not be based on the data gathered within the survey.

Participants submitted a large number of game titles as having elicited their curiosity across the five curiosity categories. However, less than half have been played by at least two participants, and only a handful of game titles having been submitted by more than 10 participants. Of the top 10 suggestions, four titles were already part of the selection of 15 game titles. Those not on the list included: *WoW*, *Fallout*, *Minecraft*, *Horizon: Zero Dawn*, *Final Fantasy*, and *The Sims*.

Of these, *WoW*, *Minecraft*, and *The Sims* are game titles that stand out as instances of eliciting curiosity in unique ways. *WoW* provides players with a vast world, but crucially, does so in a social setting. Players compete and collaborate in a shared virtual world, thus offering opportunities for spatial exploration and space for developing social curiosity. This is also true for *Minecraft*, with the addition that the game environment of *Minecraft* is practically infinite in explorable space. The algorithms that procedurally create the game environment are based on real-world principles and, as a result, are capable of shaping interesting landscapes. These can be mined for resources and used as the foundation for player-driven landscaping.

Finally, *The Sims* represents a game that does not feature much spatial exploration. Instead, it provides players with a canvas for building houses and playing out social scenarios through virtual avatars. In the case of *The Sims*, curiosity is most likely elicited by social curiosity; as interest in the relationships and circumstances of the virtual characters.

Assessment of Game Genres and Curiosity Categories Within the initial selection of games, the three highest ranked titles include the genres *Exploration & Puzzle* (*Zelda:BotW*), *Exploration & RPG* (*Elder Scrolls: Skyrim*), and *Puzzle & Reflex* (*Portal*). For these games, the genre labels *Exploration*, *Puzzle*, and *RPG* are most indicative of involving curiosity, with *Reflex* being more related to how success is achieved in the game.

The resulting *TrueSkill* ranking of genres among suggested game titles is shown in figure 4.3. While *Exploration* and *RPG* are ranked relatively high, the *Puzzle* genre is ranked second to last. This suggests that participants did not associate puzzle games with eliciting curiosity. *Portal* is not a typical puzzle game, as the game involves surprising moments in its narrative and gameplay. In contrast to other puzzle games, *Portal* sets ex-

pectations about the nature of the challenges that players will face, only to subvert them later in the game.

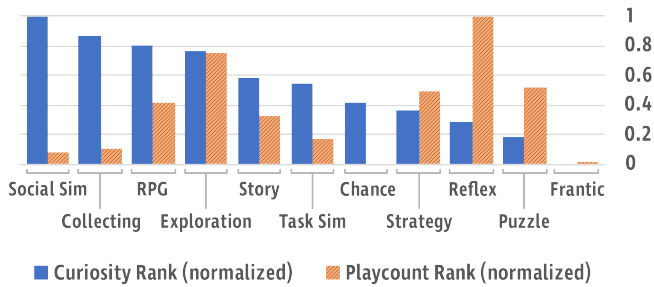


Figure 4.3: Game genres ranked by how curious participants felt in games tagged with the genre (blue), and how many games were mentioned for the genre (orange-striped). Values are normalized between 0 and 1.

A surprising outcome of the survey are the genres that ranked first and second: *Social Sim* and *Collecting*. A relatively low number of game titles carried one of these genres, with only *Chance* and *Frantic* being used less often. Due to the high amount of game titles suggested, and most of those not having been suggested by many participants, it was not practical to create a ranking among all suggested titles. However, a few examples from these genres can be instructive in how curiosity is elicited.

In *The Sims* (*Social Sim* & *Task Sim*) and *Animal Crossing* (*Social Sim* & *Collecting*), most activities are centered around the day-to-day routine of virtual characters.

Within *The Sims*, players shape the homes and relationships of several characters. Even if they have considerable control over their characters' actions, the game involves frequent opportunities for emergent behaviors and interactions between characters. Following these characters' lives can elicit curiosity, similar to watching TV soap operas. Although the game does not explicitly create a narrative for players, Simlish, the fictional language used by characters in the game, conveys a sense of different emotions. Characters further indicate the topics of their interactions through iconographic thought bubbles and inform the player of their valence due to the interaction (ranging from very negative to very positive). As a result, players can get a vague idea about a dialogue between characters but are required to fill in more specific details mentally.

In *Animal Crossing*, players only control their player character. Other characters in the game go on about their own lives and have distinct personalities and hobbies they pursue. Activities in the game focus on acquiring various items, such as fishing or farming, that can be sold or used to improve the player's home or help other characters. Curiosity is likely elicited by getting to know the many different neighbors with which players share their village. Those may change over time, drawing from a roster of 413 distinct characters in the latest installment of the game series. *Animal Crossing* is also an example of the *Collecting* genre, which elicits curiosity by introducing uncertainty in activities connected to acquiring items. At first, many activities in the game yield seemingly random results. The act of fishing, for example, can yield 80 different fish species and various items (such as lost keys that can be returned to villagers). Over time, players can discover patterns that make specific catches more likely, such as discerning the shadow size of certain fish, preferred locations, and times at which they are most active. A similar depth of options is found in other activities, such as catching bugs or digging for fossils. In these cases, it is likely the uncertainty of what specific item will be acquired that elicits curiosity in players.

Another surprise of the resulting ranks is that the *Puzzle* genre ranked low, even though it appears to be a fitting genre to engage the *Deprivation Sensitivity* dimension of curiosity. To recall, the 5DC describes this dimension as the need for resolution, a need that puzzles seek to elicit as part of their design. Games suggested under the corresponding category (GSOL) are mentioned frequently but ranked low in curiosity (figure 4.4). It could be that this dimension of curiosity does not strike players as an essential component of curiosity. Interestingly, both *Zelda:BotW* and *Portal* rank high in the shared game list, despite carrying the *Puzzle* genre. For these games, it may not be that they include puzzles that elicit curiosity in players. Instead, exploration might be the most defining genre in *Zelda:BotW*, whereas *Portal* stands out with its unusual game mechanic and surprising narrative components.

Looking further at the ranking of curiosity categories, the GEXP category ranked far above other categories (figure 4.4). This suggests, again, that participants consider “finding out new things” dominant aspects of what elicits their curiosity in a game, compared to “solving” (GSOL) or “connecting to people” (GCON).

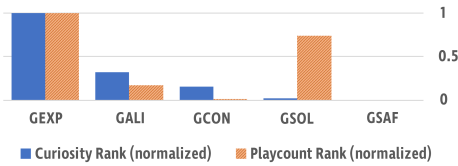


Figure 4.4: Curiosity categories based on the 5DC model, ranked by how curious participants felt in games provided under the category (blue), and how many games were mentioned for the category (orange-striped). Values are normalized between 0 and 1.

4.4.2 Impacts of Trait Curiosity and Demographics

Considering the impact of trait curiosity dimensions (as measured by the 5DC) on other measures, a few significant correlations related to the research questions were found (table 4.5). To recall, the 5DC questionnaire is part of the study to answer RQ2 (“Do individual differences in trait curiosity dimensions impact what games and genres make a player curious?”) and RQ3 (“Does age or gender impact what games or genres elicit curiosity?”) through statistical assessments.

In terms of impact on the ranking of individual games, *GTA* was ranked higher by participants with increased *Stress Tolerance* (ST) and *Thrill Seeking* (TS), while *Call of Duty* (*CoD*) was ranked higher with increased ST. Given that both *GTA* and *CoD* were ranked low overall, this suggests that players do not consider these dimensions as defining what elicits their curiosity in a game. High ST and TS might impact what kind of games these participants play but not necessarily impact what stands out as stimulating curiosity. *Zelda:BotW* was ranked higher with decreasing ST. Here also, given the high rating of *Zelda:BotW*, ST does not seem to predict overall curiosity. It can be speculated that, despite having combat and potentially stressful elements, *Zelda:BotW* allows players that are easily stressed to express their curiosity still. On the other hand, to express curiosity in *GTA* or *CoD*, players require a higher stress tolerance.

Regarding genres, *RPG* was found to correlate positively with *Joyous Exploration* (JE), ST, and TS. What stands out, in this case, is that such a correlation is perhaps even more expected for the *Exploration* genre, which did not show any significant correlation. Role-playing also seems to involve social aspects, yet *RPG* was not found to correlate with *Social Curiosity* (SC). The best effort of interpreting this result is that *RPG* games elicit

curiosity through spatial exploration and moments of high intensity, but are not solely defined by those, as would be the case for *Exploration* or *Reflex*. It might be that *RPG* represents somewhat of a balance between these genres and elicits curiosity more effectively as a result. This interpretation is supported by *RPG* having ranked higher than either of the two (figure 4.3).

Collecting was found to be inversely correlated with *Deprivation Sensitivity* (DS). This might suggest that a high need for resolution is more likely met with frustration than curiosity when faced with a game defined by extensive foraging. High ranking of *Puzzle* games inversely correlated with ST, suggesting that easily stressed players are more drawn to games that elicit their curiosity through cognitive challenges. For SC, *Social Sim* was correlated positively, and *Frantic* was inversely correlated. The impact of SC on ranking *Social Sim* makes intuitive sense and is most interesting in the absence of other such correlations (e.g., no correlation between JE and *Exploration*). Having little curiosity for *Frantic* gameplay might point at such games rarely focusing on game characters.

Notably, only two game curiosity categories were found to correlate with 5DC dimensions; and those did not correlate with what would be considered their respective dimensions. Games that emphasize the connection with people (GCON) correlated with JE, suggesting that games provided under that category might have involved exploration together with other characters or players. Games that let players “solve something” (GSOL) are inversely correlated with ST, similar to the *Puzzle* genre, and likely follows the same explanation.

In demographics, the study found only a few indications for impacts based on gender, age, or play frequency. Gender differences were found for the genres *Strategy* and *Task Sim*, which were ranked higher in eliciting curiosity by male participants. Age correlated with ranking *Puzzle* higher, although it should be noted that most participants were between 20 and 40 years old. It also inversely correlated with the ranking of games that elicit a feeling of safety (GSAF), suggesting that younger participants might still experience a higher degree of novelty when playing “stress-free” games such as *The Sims*.

4.4.3 Limitations

A few study design limitations should be noted when basing further work on the results and subsequent interpretations.

A fundamental limitation of the study is that the affective concept of curiosity and the definition of what constitutes a video game is not explained to participants. Both are frequently used terms in everyday language use but do not follow agreed-upon definitions. Therefore, the study's results reflect a wide range of interpretations that were not further captured in detail. The study did not ask participants to explain how they understood either of these terms; thus, it is impossible to assess how different interpretations affect the results. This limitation was deliberately accepted as part of the study design to assess curiosity through the lens of the emerging interpretations of video game players. However, it means that findings regarding the experience of curiosity are not instructive for a better understanding of curiosity as an affective construct.

The survey and all promotional material related to it were presented in English. This, in addition to the choice of online recruitment channels and the utilization of personal and professional networks in the Netherlands, makes it likely that most participants were English-speaking Dutch residents. Demographic data did not include information on the participants' residence and, therefore, cannot provide information about whether this impacted the results. It is reasonable to assume that the same survey conducted in other parts of the world would have generated different results, as familiarity with individual games and genres will differ.

Another essential aspect to emphasize is the use of the *TrueSkill* algorithm and the decision to have participants provide a ranking of games rather than evaluate them on a Likert scale. While the rationale for its use has been discussed in prior sections of this study, it remains an untested measure for evaluating player curiosity. Cross-evaluating the results of this study with other measures would help to strengthen the findings and solidify the viability of the *TrueSkill* method for ranking affective appraisals by participants, such as their curiosity.

It is likely that detail was lost by the decision to combine individual game titles of a series into single entries. This decision was taken to examine game design patterns rather than focus on specific, individual differences between titles. However, it is possible that

doing so removes essential information, as the nature of what elicits curiosity within individual game titles might be very different.

For future studies, the game genre labels devised as part of the study require further examination. Depending on the game, using only two genre labels is not enough to describe what activities could elicit curiosity in players. Further investigations should explore the individual design aspects of the suggested titles and how they stimulate curiosity.

Finally, the interpretation of this study should not be that an *Exploration* game automatically elicits curiosity in players. A game might involve many aspects of what constitutes an *Exploration* game without successfully eliciting much curiosity. A gap remains between the intention of a game and its ability to realize that intention.

4.5 Conclusion

This study aimed to provide a starting point for considering what game titles and genres should be analyzed regarding their potential to elicit curiosity. Through the suggestions of survey participants, a list of 15 selected game titles was ranked by this criterium and assessed through the *TrueSkill* score algorithm. Suggestions of participants extended on that list and allowed for the ranking of game genres defined for the study. As a result, the study was able to address RQ1 by creating a corpus for further investigation.

The study found that the games *Zelda: Breath of the Wild*, *Elder Scrolls: Skyrim*, and *Portal* were ranked as the most successful in eliciting curiosity. Within these, the genre labels *Exploration*, *RPG*, and *Puzzle* are most representative of what activities in the game elicit curiosity. Among games suggested by participants, the genres *Social Sim*, *Collecting*, *RPG*, and *Exploration* ranked the highest; thus, providing evidence that the potential of *Puzzle* games to elicit curiosity is highly dependent on the game.

The study's results suggest that what makes players curious in a game does not systematically correlate to their scores on the 5DC questionnaire and, thus, their trait dimensions. Although some individual correlations with curiosity dimensions were found and described as part of the discussion, the results do not provide a sufficient basis for answering RQ2.

Results of the study further show that male participants ranked *Strategy* and *Task Sim* games higher in curiosity than female participants. Age was found to correlate with the ranking of *Puzzle* games, with older players ranking these games higher. The study answers RQ3 by revealing differences in what genres elicit curiosity based on age and gender. However, no significant differences were found in the ranking of individual game titles.

Overall, the results of the study provide evidence for the theory that games that strike a balance between uncertainty and structure tend to rank high. In contrast, highly deterministic games (requiring only cognitive or physical aptitude) or based on chance tend to rank lower in curiosity. How to strike that balance and whether that theory holds true will need to be assessed as part of future investigations.

The next step, based on the study's outcome, is to look at the individual games and genres that were most successful in eliciting curiosity. More specifically, it means investigating what design interventions are most likely to contribute to curious behavior in a game.

