How do people evaluate virtual goods in social media? The case of Dota 2

Author: Denis Bulygin
Supervisor: Professor Annika Waern

July 8, 2019
Abstract

Virtual purchases are the main source of revenue for developers of F2P games being a market with expected 17.4 billions of dollars volume in 2019. Despite the broad scope of research of virtual purchases, it is still unclear how the player evaluate non-functional goods. Based on analysis of discussions of virtual decorative items this work what experiences nonfunctional items grants players with and how those experiences discussions reflect in the item’s price.

With the use of Structural Topic Modeling framework this work demonstrates the dimensions of players’ experience in their association with price change on the case of Reddit.com subreddit /r/Dota2. Analysis reveals three main categories of discussions: dimensions of hedonic value, dimensions of social value, expectations mismatch. This work contributes to studies of virtual purchases by decomposing each category into experience dimensions and by revealing the relationship between extracted experience dimensions and items price.

Keywords: virtual-goods, virtual-consumption, online games, purchase, evaluation
# Table of Contents

**Abstract**

1

**Table of Contents**

2

**Chapter 1. Introduction**

4

**Chapter 2. Game Description**

6

**Chapter 3. Related works and Theoretical Framework**

9
  - Free-to-play model: market as a monetization tool 9
  - Sociomateriality: why virtual items are valuable 10
  - Three approaches to study virtual items consumption 10
    - Psychological factors of virtual consumption 10
    - Virtual purchases as experience 11
    - Sociological perspective of items' purchase drivers 13
    - Values as singularities of meaning 14
  - Theoretical framework 15

**Chapter 4. Research questions**

16

**Chapter 5. Methodology, data, analysis**

17
  - Methodology: Netnography and digital studies 17
    - Ethics of Netnography 18
  - Methods 19
    - Topic modeling 19
    - Topic modeling application in social sciences 20
    - Structural topic modeling 20
  - Data 21
    - Items dataset 21
    - Text corpus 22
    - Market price dataset 23
  - Analysis 24
    - Research Question 2. Pearson correlation 24
    - Research questions 1 / 3 / 4. Structural Topic model 24

**Chapter 6. Results**

27
  - Research Question 2. Relationship between item price and discussions 27
  - Research questions 1a-1b. Dimensions of players experience 27
    - Hedonic Value 29
      - Cool look 29
      - Favorite cosmetics 29
      - Equipment slot evaluation 29
      - Animations and Effects 30
Items with lore 31
Expectation mismatch 31
  Comparison of Arcana items 31
  Confusion on rarities 32
  Fix the bugs 32
  Bugged Items 33
  Thanks for fixing bugs 33
  Market scam 34
  Copying the Items 34
Social value 36
  Expensive rare items 36
  Chances calculation 36
  Drop chances 37
Other 37
  Store discount sales 37
  The item changes 38
  Event venue 38
Relationship between topics 38
Relationship between topic and price dynamic 43
Results summary 44
  RQ 1a. What dimensions of players' experience occur in the discussions of virtual items? 44
  RQ 1b. What is the relative prevalence of the extracted dimensions in the discussions? 44
  RQ 2. What is the relationship between items' mentions frequency and their market properties? 45
  RQ 3. How do dimensions of players' experience interact with price change of virtual items? 45
  RQ 4. How do dimensions of players' experience interact to each other? 45

Chapter 7. Discussion 46
  Aesthetic value in appearance and visual effects 46
  Social value and scarcity 47
  Expectation mismatch 48
  Judgement devices of items evaluation 49

Chapter 8. Conclusion 50
  Theoretical and practical implications 50

References 52
Chapter 1. Introduction

The development of the gaming industry has led to the emergence of new forms of monetization of games. The situation in which the developer distributes the game for free and earns money on in-game sales of virtual goods has become popular. This model of monetization has become especially popular in online and social games on Facebook. The market for such games is actively developing and by 2025 its volume is expected to reach 189 billion US dollars according to Adroit Market Research.

The growth of such markets leads to increased interest in the understanding of the processes that can be observed there. Industry and academia are beginning to wonder: why do players spend real money on virtual goods? How do the players agree on the price? How to evaluate the goods?

The manifestation of such interest can be attributed to the fact that gaming markets are becoming an important part of the economy and involve more people annually. Volumes of virtual goods markets are growing, therefore markets require regulation. However, there is no understanding of the processes that occur in virtual markets and therefore any regulations will possibly harm the gaming industry.

The recent debates on this topic are mostly related to the problem of loot boxes and gambling which meet legal regulations from the authorities. Nevertheless, it is important to understand the processes in other aspects of virtual consumption such as trading and evaluation. Trading and consumption of virtual goods is the primary source of revenue for most players (Alha et al. 2014), that is why it is important to study the experiences of players that can be interconnected with assets purchase. Game design affects players' perception of consumption experiences which can lead to situations when the experience of players could be improved by design which would result in an increase of developer's revenue.

Goods become valuable due to their ability to make the game easier, i.e. "functional" role (Lehdonvirta 2009); or due to "decorative" benefits (Gyuhwan and Taiyoung 2007) which could be expressed in "social" or "emotional" (hedonic) attributes (Lehdonvirta 2009). In a case of mainly functional items, purchase intent is mostly associated with item's utility (use value) in the game. As for decorative items, it is important to consider the status and aesthetic qualities, i.e. visual and sound representation in the virtual world (Lehdonvirta 2009).

Current studies of virtual consumption mostly focus on functional goods (Hamari and Keronen 2016) possibly due to the fact that in most of online-games, both “functional” and “decorative” values are highly interwoven in the same item (Lehdonvirta 2009), which provides predominance of functional value in consumer motivations to purchase an item. Focus on functional goods leads to lack of understanding of how the value of decorative goods is perceived and constructed.

Current research of cosmetic items is mostly related to psychological factors of consumption (Hamari and Keronen 2016). The works in the context of players experience (Toups et al. 2016; Keronen 2016).

---

Marder et al. 2019) focus on revealing the practices important for players but do not shed light on the importance of observed practices and experiences for communities of players.

Aim of this thesis is to understand how perceived experiences of players reflect the price changes of cosmetic items on the secondary real trade game market. For that purpose, I analyze discussions of cosmetic items on Reddit.com on subreddit r/Dota2 which is the largest community hub in English language. In addition to discussions, I analyze market data of items being discussed and match the data to find out how subjective perception of consumption experiences reflects the price of virtual items.

I conduct this study using Structural Topic Modeling which is a framework of quantitative text analysis technique called topic modeling. With Structural Topic Modeling I analyze 4766 comments about 1088 virtual items in an attempt to answer four research questions which in general help to understand the process of cosmetic items evaluation:

- RQ 1a. What dimensions of players’ experience occur in the discussions of virtual items?
- RQ 1b. What is the relative prevalence of the extracted dimensions in the discussions?
- RQ 2. What is the relationship between items’ mentions frequency and their market properties?
- RQ 3. How do the dimensions of players’ experience interact with price change of virtual items?
- RQ 4. How do dimensions of players’ experience interact to each other?
Chapter 2. Game Description

Dota 2 is a game in a genre of Multiplayer Online Battle Arena (MOBA) that was released by Valve in 2013. Dota 2’s gameplay consists of short sessions (0.5 - 1 hour) with two teams of five players fighting against each other in an attempt to destroy the enemies’ base. Each player operates a virtual avatar called a “hero” with its unique abilities. During a session players earn levels and equipment for their heroes to become more powerful than the opponents. Earned equipment and levels do not transfer between sessions so that players are free to decide if they want to play the same hero or try another option.

Dota 2 is a free-to-play (F2P) game which makes it a game available for free with microtransactions for real money. In this monetization model the microtransactions are the main source of income for the game developer. In the July of 2012, Valve launched an in-game store Dota2 Store2 which let the players purchase the assets. The assets are virtual cosmetic items which do not affect the gameplay (i.e. by making the avatar more powerful).

Three basic types of virtual cosmetic goods are an item, an item set, and a treasure chest. The item is an individual object that takes one inventory slot (e.g. head or hand) and changes piece of a visual model. The items united into a set usually have common theme and color-scheme (See Fig. 1). The players can combine items from different sets and obtain separate items without acquiring a whole set.

![Figure 1. Example of equipment slots](http://www.dota2.com/store)

Some items have additional visual effects that change animations of heroes actions and magic abilities. Two characteristics of visual effects are Rarity and kinetic gems. Each item has a class

---

2 [http://www.dota2.com/store]
called “Rarity” which despite the name describes an intensity of visual effects. Rarity has eight classes: Common, Uncommon, Rare, Mythical, Legendary, Immortal, Ancient, Arcana. While common rarity describes the items with no visual effects, Arcana items can change the whole model of a hero, ability icons and visual effects of magic abilities.

Another source of effects are kinetic gems that are objects of specific type that can be extracted from the item and moved to another one. Since, some visual effects are assigned to items which created a confusion about visual effects and items.

The treasure chests, in turn, include several sets or separate items of the same Rarity. Once bought, the treasure chest gives a player one item (or set) chosen randomly, and then disappears. Despite the same Rarity some objects have different chances to be given away (drop chance) and those items are significantly more difficult to get in comparison to the items with normal chance. It creates additional inequality of distribution among items and increases the scarcity of particular items.

In Dota 2 there are several most common ways to acquire cosmetic items: Dota2 Store, Steam community market⁴, and in-game activities. The Dota2 Store is the primary source of virtual goods for players. It includes most of the sets or the items and treasure chests which players purchase the for the fixed price. Dota2 Store is also a source of in-game objects which reward the players with items and treasure chests for accomplishing various in-game activities such as tournament betting, predictions of match outcomes, quests, etc.

---

⁴ https://steamcommunity.com/market/search?appid=570

---

Figure 2. Items of the same rarity have different drop chances as well³
Another source of items besides Dota 2 Store and in-game activities is Steam Community Market. Steam Community Market is a secondary market for the players who are willing to sell the items. Market uses real money transferred into steam wallets, and though the players use real money to trade, they cannot withdraw money from Steam. On Steam Community Market, the players set the price themselves and the market is mostly unregulated by developers who do not interfere with the price formation process. However, developers take a commission for each trade deal and also make some items unavailable to trade on market.
Chapter 3. Related works and Theoretical Framework

Free-to-play model: market as a monetization tool

The first games based on Free-to-play (F2P) model were introduced to the world in 2000. The main difference of F2P games in comparison to classical models is that they are distributed to players for free but pieces of the content can be purchased for real money and a game usually has online- or social components. Often F2P games include two types of the in-game currency: soft currency and hard currency. While soft currency is virtual money obtained by player in the gameplay as a reward, the hard currency is virtual money transferred by players in exchange for real money. Currency type defines what virtual or digital goods players can purchase and usually hard currency can be transferred into soft one but not vice versa. According to statistics of 2013 year (Dredge 2013), 92% of downloaded applications on the AppStore and 98% of downloaded applications on Google Play are Free-to-play.

Overall, F2P raises controversial attitude in the game industry. Even though this is a generally positive opinion among practitioners about F2P model (Alha et al. 2014), the ethics of F2P methods questioned as game developers tend to use dark design patterns to reinforce the sales (Alha et al. 2014). Moreover, F2P games usage is associated with games addiction(Dreier et al. 2017) and money spendings which make the monetization model more controversial.

There are three most popular types of F2P games: pay-to-win, pay-to-pass boring parts, pay-for-visual (Heimo et al. 2018). The first type is Pay-to-win model which makes a revenue by selling the content that increases the relative power of the owner. Usually, pay-to-win games offer the players the powerful weapon or game avatar that will put the owner in an advantageous position in comparison to non-buyers.

The second type is Pay-to-pass-boring parts games. Games with this model put the players in the situations of waiting for the game to allow next actions. Usually, game developer includes energy points which the player uses to make actions and when all the points are spent the players must wait to fill their energy up and be able to act again. However, the in-game shop provides the players with the ability to purchase energy points for real money.

The last type of monetization models is Pay-for-visual games (Gyuhwan and Taiyoung 2007) which provides players with additional decorative content such as alternative appearance of avatars or interface elements. This model is becoming more popular in the industry as it does not place players in the unfair position as Pay-to-win system does and it does not discourage players from staying in the game as Pay-to-pass-boring does. According to this classification, Dota 2 is pay-for-visuals game that provides the players only with cosmetic goods that do not affect the gameplay. Next subsections will answer the question of why virtual goods are valuable to players in the first place.
Sociomateriality: why virtual items are valuable

One of the possible explanations of why virtual items become valuable to players lies in the concept of “magic circles” (Huizinga 1950) that create endogenous meaning. A game puts the player into specific context with its own rules, achievable goals, and meanings. Context of the game is what makes actions and items that do not seem to have any value valuable for players. For example, though a player cannot use Monopoly money in the grocery store, the players are encouraged to use it to achieve in-game goals by the rules of the game.

In-game meanings are supported by social reality (Searle 1995) shared by players (Montola 2005). Rules and meanings within the game are socially constructed unless they are enforced by outside forces. For example, rules of physical games do not exist unless players create them and do not work unless all the players follow them. At the same time, digital games include the rules based on the design of the game and exist regardless of the players. However, it does not make digital games limit the players’ decisions entirely.

Digital games still provide players with an opportunity to set the rules of the playing. Moreover, being a monopolist of endogenous meaning, digital games also affect player’s judgement of success and failure in the game. By setting the goal to a player or giving him an opportunity to create their own digital games encourage players to use digital objects to achieve these goals (e.g. players can create their own collections of items based on an affinity that is not provided by the game (Toups et al. 2016)).

As a result, players can judge their success based exclusively on game context. However the players somehow choose what items are more valuable and, more importantly, they decide if some items are worth of real money cost. Next subsection covers studies focused on this topic.

Three approaches to study virtual items consumption

The field of virtual consumption can be separated into three groups of studies. First group of studies is focused on psychological factors of virtual goods purchases (Hamari and Keronen 2016; Bleize and Antheunis 2019; Hamari and Keronen 2017). Second group is based on sociological understanding of virtual consumption (Lehdonvirta 2009; Lehdonvirta, Wilska, and Johnson 2009; Marder et al. 2019). In particular, the research is focused on virtual goods attributes that drive the players to purchase items. The last group of studies is led by HCI research and is focused on experiences virtual items create and practices that make players interested in using the items (Toups et al. 2016; Musabirov 2016; Musabirov et al. 2017; Bowser et al. 2015). Further, I provide an overview of all the aforementioned groups by discussing the most valuable works in their relation to each other.

Psychological factors of virtual consumption

Juho Hamari and Lauri Keronen reviewed (Hamari and Keronen 2016) the main works on consumption in online games. They wanted to understand what factors are the most important
In decision making about the purchase of goods. They defined two factors which are consumer behavior and consumer intentions as the most often mentioned in the field.

By analysis of behavior researchers are trying to understand what drives players to buy a product. They consider a variety of psychological factors, such as subjective norms, habits, and purchase intentions. The last two factors play a major role in consumption as the majority of studies show (Hamari and Keronen 2017, 2016). However, habits, unlike intentions, do not explain why users choose particular type of the product as this factor does not explain why customers make a decision to purchase something in the first place.

Intention, being a psychological characteristic of consumers, on the contrary, makes it possible to find out what drives people to purchase virtual goods. Meta-analysis of virtual consumption studies (Hamari and Keronen 2017) shows that such psychological factors as attitude has the largest correlation with purchase intention (corr = 0.7 among studies). Other well-correlated factors are flow, network size, self-presentation and subjective norms that are correlated with purchase intention on the level of 0.4. Presented factors demonstrate the variety of individualistic and social reasons for players to purchase virtual goods.

It has been several years since the publication of the literature review but it cannot be said the field has changed dramatically. The only major difference observed is the growth of works related to virtual consumption in the context of mobile gaming (e.g. Balakrishnan and Griffiths 2018) and increase of interest to gambling in online games (Macey and Hamari 2019).

In addition to literature review conducted by Hamari and Keronen, there was another literature review that also attempted to summarize the factors of virtual purchase (Bleize and Antheunis 2019). Based on the smaller number of papers the authors came to different conclusions about the most important factors of purchase. The authors highlighted four most important factors: entertainment, social influence, customization, and ease of use.

Nonetheless, it cannot be said that the results of both reviews contradict to each other. Rather they analyze the consumption from different perspectives. While Hamari and Keronen describe the factors on more abstract psychological level, Bleize and Antheunis are more focused on purchases as activities that help the users to accomplish some goal. In some way, literature review is closer to HCI field as it highlights the importance of player’s experience in purchases.

Virtual purchases as experience

Researchers of this field focus rather on game design and player experiences than psychological models of decision making. For example, Toups et al. (Toups et al. 2016) describes items from the point of view of collectible practices. They view the game as a system of rules devised by the developer. The rules define the goals and objectives of the players associated with both common and personal achievements. A game may encourage players to make a collection of figures to unlock game achievement, or players come up with their own collections, not provided by the game. Thus, a separate item is not as valuable to players as practices associated with its possession and use.

Researchers grounded their study ten types of value, developed by Livingstone et al (2014). In their study, Livingstone et al. found relationship between identity or self-expression and
preferences in characters. Rational reasoning based on utility and personal investment values also was found among motivations. In addition, they found sentimental value, which is expressed when characters store memories about experience. Looking for new experience is close to the aforementioned category and works in a similar way but towards the future activity. Players also described such concepts as enjoyment, aesthetic and creative values, which are closely related to the concept of hedonic value, i.e. emotional value (Aspers and Beckert 2011). Besides emotional value, there are also social factors, influencing choice of character, such as sociability and social communication value. While the first one is expressed in the ability of characters to find new or support old relationships, second one, in turn, means highlighting owner social status and belonging to specific groups or community. Thus, players assess characters based on how they spend time in the game, which reflects playing experience (Livingston et al., 2014).

The change of focus from objects to accompanying processes is well described in the study of consumption of vintage goods (Bowser et al. 2015). In their work, researchers analyzed the practice of buying second-hand goods. They determined that the process of choosing things is an important part of the purchase.

Consumers felt pleasant sensations from searching and choosing things. Choosing the right thing is like solving a puzzle, where the final picture will be your own unique appearance. Thus, it is not the purchase of a thing that is the ultimate goal of shopping, but the process of searching for that very thing that complements the look.

Lastly, there were two studies with analysis of communication on Reddit about virtual goods. First work is dedicated to study the secondary market of Dota2 professional players merchandise and fans’ discussions of players’ personalities on r/Dota2 (Musabirov 2016). The author presents a model of price formation on the digital autographs of Dota 2 players on the market where the prices are set by the customers. Using the price as a proxy to player’s value in the community, the author describes what makes the personal brands of players more valuable.

To serve this purpose, the model includes information about players’ game statistics, media coverage, tournamental achievements and players’ personal information such as nationality and position in the team. In this way, Reddit discussions are used to verify the interpretation of the model and to demonstrate other non quantifiable factors that contribute professional player’s brands such as loyalty to the organisation, personal style, etc.

The aim of the second work (Musabirov et al. 2017) is to analyze the discussions of virtual items to find out what logics players use during evaluation and what activities make items valuable to them. This work used topic modeling as well and range of logics and activities was extracted from communication. For example, the practice of collecting and combining into unique look were found. The players discussed the rarity and aesthetic quality of items, and judged if the item corresponded to lore and heroes’ background.
Sociological perspective of items’ purchase drivers

The last approach in studying virtual purchases is focused on general understanding of evaluation of virtual goods. The main classification of types of items’ values is the division into emotional, social and functional. Although researchers operationalize them differently and include different characteristics of goods / consumers (Guo and Barnes 2011; Kim, Gupta, and Koh 2011; Lehdonvirta 2009), this type of separation is the most common in the field.

Functional attributes express the ability of an item “to be used as instruments towards fulfilling some higher objective, usually a tangible material objective that is seen as related to some fundamental human need”. In other words, when purchasing an item, a player expects the item to be helpful in achieving the particular goal. The goal can be related to accomplishment of the game quest, getting an achievement or even a desire to look better.

Author connects this attribute to the Marxist theory, calling the functional attribute a *use-value* (Miller 1987). The use-value is defined by two aspects taken from marketing: *performance* and *features*. The *performance* of the items defines how powerful their owner will be in comparison to others (e.g. players compare which weapons give them more attack points) and in this way the item’s performance exists in relation to other items because “if everyone has high performance, no one has high performance” (Lehdonvirta 2009, p. 105). The *features* of the item, in return, define the actions this item allows a player to do. For example, in some online games players are allowed to fish but they need to acquire a fishing-rod beforehand.

Lehdonvirta states the *performance* is the most-common driver of virtual consumption in online games such as World of Warcraft. Additionally, the author also highlights the importance of the new actions defined by *features* of the items. These two aspects play the vital role in players’ decisions on virtual purchases. However, he also acknowledges the presence of the items that lack both the performance and features and for those types of items, as he suggests, have hedonic or social attributes.

Lehdonvirta defines hedonic attributes as the properties that evoke the visual or aesthetic pleasure and pleasant emotions in their owners. The most important aspects of hedonic attribute are appearances of a virtual item and visual effects it creates when used. According to Lehdonvirta (Lehdonvirta 2009,p. 102), hedonic value is a mix of pleasure and aesthetics expressed in visual and sound representation.

Another dimension of items evaluation is based on the ability of virtual goods to highlight the owner’s status and belongingship to a specific group. The value that expresses social position and self-identity of the owner is called social value and has its roots in the early sociological theory of the consumption proposed by Veblen (Veblen 2017). In his work, Veblen describes elite class in attempt to explain excessive expenses and behavior and reveals the need of higher social classes to emphasize the difference with lower classes which is supported through consumption of particular goods. In this way, consumed goods and conspicuous behavior work as a social marker and help to highlight the social distance between classes. Lehdonvirta applies this theory to online games in an attempt to describe the social value of virtual non-functional goods.
It is worth to note that all three values are interconnected in the same item. In this way, each item is a mix of functional, hedonic, and social values and depending on the proportion one of the values is dominating. In case of items with no functional value the item can be called non-functional, cosmetic, or decorative item.

This approach has not been changed significantly throughout the years. The classification of values is tri-dimensional in most of the papers and separates functional (utilitarian), hedonic (emotional) and social values. Most of the works focus on the functional aspect of the items and find it an important driver to purchase the items (Hamari and Keronen 2016).

The recent work on this topic (Marder et al. 2019) in some way supports the findings of Lehdonvirta’s work and extends the analysis of emotional and social value in virtual consumption. Analyzing interviews of League of Legends players the authors extracted nine key themes and found that emotional (hedonic) motivation has five aspects important for players: novelty, aesthetics, reciprocity, self-gratification and character dedication. Social motivation, in return, consists of four components: gifting, social distinction, showing reciprocity, visual authority.

Values as singularities of meaning

The virtual goods analyzed by Lehdonvirta are a bright example of what Lucien Karpik calls a “singularity of meaning” (Karpik and Scott 2010; Healy et al. 2011). According to his framework of “economics of singularities”, the evaluation of items with no visible utility (Karpik analyzes aesthetic goods) is a complex social process. While functional goods have visible scale of quality, aesthetical goods have no such attribute that would let customers compare two objects. Moreover, the aesthetical goods are a mix of different social and symbolic aspects which, in their interconnection, create “singularity of meaning”.

Though undoubtedly aesthetic goods have different value for customers varying in price between several dollars and millions of dollars, this interconnection of social and symbolic meanings makes it impossible to define the scale of value for aesthetic goods that would vary from “bad” to “good” (Karpik and Scott 2010). However, it is possible to reveal the most important attributes of evaluation and to analyze their relationship with price.

In the study of French wine market Beckert et al (2016) described wine price formation mechanisms and like Lehdonvirta did in his study of virtual items, Beckert et al. deconstructed wine into several aspects that are believed to play a role in price formation: wine age, year and place of origin, etc. Using hedonic regressions approach researchers revealed the relationship between those aspects and price of wine. In this way, hedonic regressions let researchers and producers predict and explain the process of price formation. Lehdonvirta, in return, did not try to reveal statistically credible relationship between item’s attributes and rather focused on the definition of the attributes that play a role in price formation.

However, in real life setting it is very difficult for customers to make calculations on the value of the goods or services they purchase (Karpik and Scott 2010). Nevertheless, they make a judgement about goods and what object to choose. Karpik suggests that people use judgement devices (ibid.) that help them to choose the right option and describes five types of judgement
devices: rankings, personal and not personal social networks, brands, ciceronis (experts) and guides, and marketing (ibid.). Kornberger (Kornberger et al. 2015) highlights that not only customers use particular judgement device to make a decision but those devices are used in their interconnection. Information about wine can be a judgement device that helps the customers to evaluate the wine. For example, the year of wine release defines its price if wine was associated with remarkable moments of French history such as war victory (Beckert, Rössel, and Schenk 2016). In this way, customers can learn about wines and their association with history and evaluate them higher than others.

Theoretical framework

Aforementioned approaches of studying virtual consumption cover the same process of virtual goods purchases but they do it from three different perspectives. While marketing and psychological research is focused on psychological prerequisites to virtual purchases, sociological and HCI research tries to shed light on what makes virtual items attractive to players. In the case of sociological approach to this question, researchers attempt to understand the processes of non-functional goods evaluation in general. This research mostly provides the field with models that use hedonic regressions and theories that explain how the market of non-functional goods works (Aspers and Beckert 2011; Beckert, Rössel, and Schenk 2016; Karpik and Scott 2010). HCI research, in return, makes the major accent on what experiences virtual consumption provides the players with. In this case, the motivations to purchase items are very contextual as games have sets of constraints that differ among games and even real life (Toups et al. 2016; Bowser et al. 2015). The players can have experiences that are entangled into design of games which is a focus of HCI studies that try to uncover relationship between those experiences and design elements.

This work is based on a combination of sociological and HCI perspectives. On the one hand, it is dedicated to understanding of the relationship between experiences virtual items grant to players and items price. In the same manner as previous studies (Bowser et al. 2015; Livingston et al. 2014; Toups et al. 2016), this work is trying to uncover the practices of usage and evaluation of items and players’ reflection on their experiences towards virtual cosmetic goods. On the other hand, it describes the mechanisms of price formation which have a social nature as they are based mostly on social and symbolic interpretation of what is good (Lehdonvirta 2009; Lehdonvirta, Wilska, and Johnson 2009; Beckert, Rössel, and Schenk 2016; Beckert and Rössel 2013) and what is not. Moreover, as previous study showed the important role of judgement devices in the evaluation of virtual goods (Musabirov 2016; Musabirov et al. 2017), this work is aimed to extend the discussion in this direction by including the items’ price into consideration.
Chapter 4. Research questions

The studies of virtual consumption cover a wide range of topics and problems. There are studies that focus on the psychology of consumption, studies that deconstruct the value of items into motivations to purchase items and studies that focus on consumption and usage as continuous experience worth studying per se. However, most of the works are mostly explorative and done in a qualitative way. It is a suitable approach if the aim is to learn about practices and activities related to evaluation in the game; however existing interaction among those practices and, what is especially important, interaction between those practices and price remains unclear.

For example, though previous study uses a large corpora of data, it does not reflect on the importance of each revealed theme for the community. Moreover, though the combinations of topics reveal interesting examples of activities, the size of connection between topics is not evaluated. In this way, this study is aimed to fill several gaps in the existing studies of virtual consumption.

Beforehands, the paper will update and specify what dimensions of player’s experience exists in the discussions. It is expected to extract the same dimensions as were extracted in the previous study and to measure the prevalence of each dimension in the discussions. It will allow to compare the dimensions and reflect on their relative importance for players:

RQ 1a. What dimensions of players’ experience occur in the discussions of virtual items?

RQ 1b. What is the relative prevalence of the extracted dimensions in the discussions?

Furthermore, the paper will uncover the relationship between discussions of virtual items and their market value. This relationship will be analyzed on two levels. First, items’ popularity in social media and its position on the market will be compared. Using frequencies of mentions in the discussions and items’ aggregated price and number of sold copies I am going to understand if items’ price and scarcity are related to the amount of attention they get on social media:

RQ 2. What is the relationship between items’ mentions frequency and their market properties?

Secondly, the relationship between price and dimensions of experience will be described. The study will show what dimensions are positively or negatively associated with price and it will allow me to describe what players perceive to play a role in items’ value:

RQ 3. How dimensions of players’ experience interact with price change of virtual items?

Lastly, as previous study showed, combinations of dimensions reflect practices as well as individual dimensions. In this study I am going to measure the strength of connection between
the dimensions and to analyze the pairs with high co-presence of dimensions. It will allow me to exclude the combinations interesting for studies yet not large enough to be a reliable finding:

RQ 4. How dimensions of players’ experience interact to each other?

Chapter 5. Methodology, data, analysis

Methodology: Netnography and digital studies

The aim of this study is to understand the community of Dota2 players who talk to each other about virtual goods in social media. For that purpose, the study uses a mixed methods approach known as Netnography (Kozinets 2015) that is primarily famous as ethnographic method of understanding the social interactions in Web- and virtual communities. Netnography was invented to fill the gap in classical methods which did not fit in the context of studying virtual communities.

The quantitative methods and data analytics approaches need correct interpretation of revealed patterns which is not a problem in traditional research quite often as researchers are usually a part of studied offline community. However, with online connection it is possible to access to an online community without getting to know its members’. It can lead to misinterpretation of observed practices and behavior as their true meanings are hidden from the outsider of the community.

Traditional qualitative methods (and most importantly ethnographic studies), in return, were created in the context of offline interaction but digital communication is different form of interaction with its own artefacts to observe and processes to study. In contrast to offline communication where a researcher observes the speech and body movements, the digital communication is mostly represented with a text or visual information which creates a different kind of data that should be studied in different ways.

First of all, written text differs from transcribed speech as the participants conduct different sets of actions in both processes. Moreover, the platform for digital communication per se defines what and how the participants write the texts. Type of the communication platform affects what participants can do to share their thoughts. For example, conversation in the chat room is framed by scrolling feed of messages and it is limited by human’s capacity to read finite amount of messages in a short period of time. In contrast, conversation on the online forum gives the participants time to think about the content because other texts remain visible and speed of sending the messages becomes less important.

Furthermore, in traditional ethnographic methods a researcher’s ability to gather the data is limited by geographical and time constraints. A researcher can talk with a limited number of participants in the particular geographical area and it is a time consuming process. Digital communication, in turn, provides a researcher with plenty of data which is already stored in a form of action logs, forum posts, messages in the chat, information in users’ profiles, etc. The new challenge here is not a shortage of data but its overwhelming amount. It becomes a problem to decide what texts worth the analysis and what texts describe community the most.
Netnography is an approach that helps to overcome those problems. The focus of this approach is to use quantitative methods to support qualitative analysis. In this way, quantitative study will show the most important patterns in communication of the target community, will highlight the most important texts that describe the community and will help to interpret the revealed patterns.

The aim of this study is to explore the dimensions of virtual items' evaluation in the discussions of Dota2 players on Reddit.com. The exploration includes identification of the key evaluation dimensions, the estimation of their relative importance in the discussions and finally the measurement of relationship between dimensions discussions and price change of virtual items on community market.

In the spirit of Netnography, I am analyzing texts gathered on subreddit r/Dota2 of Reddit.com as the largest platform for community discussions with 559 000 members. Using Structural Topic Modelling I process the discussions of virtual items, extract themes of discussions, estimate the relationship between discussed themes and price change and identify the themes that often emerge together in the texts.

**Ethics of Nethnography**

When talking about ethics of chosen approach, several concerns come to mind. In particular, the observed subjects are usually unaware of the observation (that's what makes the approach valuable due to less biased behavior of subjects) which leads to two problems. First of all, the privacy and confidentiality of participants can be violated with observation. As participants do not know they are being recorded, they can share personal information that makes them exposed to undesired reaction of third parties. Moreover, as participants did not consent to be studied, the participation is not voluntary. Some communities are not willing to invite strangers even the researchers which raise an issue of ethical conformity.

Aforementioned concerns are mostly related to cases when private communities are in the focus of research. When the researcher tries to enter closed community, she needs to have a consent form to make sure community participants agree to share their privacy with researcher. In this case, the confidentiality and privacy can be violated and agreement of participants is important as they demonstrate the intention to hide communication within a small group of people.

However, in the context of Reddit.com discussions which is the primary source of data those concerns are arguably an issue. Reddit users expect (even willing) to be seen by a group of strangers in the Internet, as the users’ attention is the main source of reputation on website. The discussions used in this study are publicly available and anonymized. It means that the users are expected to agree to be read by others and their privacy is not violated as their identities are not represented in the quotes and analysis of data.

Nevertheless, there is another problem relevant for both private and public communities. The users have a total control over information about them which creates second identity in the Internet known as digital double. In this case the researcher is not fully aware of the subjects’ background, hence, observed behavior can be misinterpreted and can be understood only
The psychology of virtual goods and experiences research is of less relevance to this study.

Methods

Topic modeling
To extract experiences of players reflected in the discussions a topic model was estimated. Topic modelling (Steyvers and Griffiths 2007) is a machine learning-based method that let researchers analyze large bodies of texts. Topic modelling algorithms treat texts as ‘bag-of-words’ objects which ignore words positions, their lexical meaning and punctuation and only count co-occurred words and their frequencies. Using information about words co-occurrence, topic model defines groups of words that tend to occur more often than others.

Each group of words is known as topic. In order to create a topic model a researcher defines the desired number of the topics. Regarding topic model output, each topic is probability distribution of all unique words in the text corpora. However, in each topic there will be different words with highest probability as different groups of words will co-occur in different texts. Usually topic has a long tail distribution meaning that several words accumulate 90% of probability while the vast majority of other words has a probability close to zero. As each topic is a mixture of words with different probabilities, each document is a mixture of topics with different proportions since documents consists of the same words.

Though model provides researchers with topics, topic model is not meaningful until topics are interpreted and labeled by researcher. In this task, researcher relies on the words most related to a topic and documents with the highest proportion of this topic.

Topic modelling was chosen primarily because this method let a researcher handle large textual data by clustering many disconnected texts into topics. Manual techniques such as thematic analysis are not suitable for the large text corpora analysis and Paul DiMaggio suggest three reasons for that (DiMaggio, Nag, and Blei 2013). First of all, manual analysis is difficult to conduct on a large body of texts which makes it time-consuming and impractical. Second, in the more complex analytical tasks it is harder “to achieve acceptable levels of intercoder reliability” (DiMaggio, Nag, and Blei 2013; p. 577) as they demand more intersubjectivity. Lastly, a researcher usually presumes beforehand what is worth finding which makes exploratory part of research flawed.

Alternatively, dictionaries on various themes could be used (and actually were used in the first iteration) to calculate the frequencies of each theme which solves the first problem. However dictionaries are conducted manually which reduces exploratory effect of text analysis, as not all word usages can be considered by researcher beforehand.

Topic modelling, in return, creates clusters of words in the topics which makes the research reproducible as model results can be conducted again and it shows a researcher the points of interest by giving a ‘map’ of discussions in text corpus as each document has prevalent topics.
Topic modeling application in social sciences

Topic modelling becomes popular tool for studying large bodies of texts in virtual studies and social sciences. One of the first researchers who applied topic models in the context of social science was a group of researchers (DiMaggio, Nag, and Blei 2013) who studied texts of political news using topic modelling algorithm Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003). Using LDA Dimaggio coded texts documents and interpreted how media covered news related to art. Dimaggio interpreted topics as frames which are “semantic contexts that prime [...] interpretations of the phenomenon in a reader” (DiMaggio, Nag, and Blei 2013, p. 578).

It is important to keep in mind that topic models describe interpretations of the authors about phenomenon rather than phenomenon per se. This work presents analysis of discussions about items with relation to price change. It does not uncover real player’s motivations to purchase items but uncovers their understanding of important aspects of item evaluation.

Another example of topic modeling is presented in the study of fan communities of the NBA teams (Zhang, Tan, and Lv 2018). Researchers analyzed how teams performances on the tournaments affect fans’ activity in the Reddit discussions about NBA and their teams. First of all, they found that the loss of top teams and win of bottom teams increases discussion activity. Moreover, younger team get more loyal fans in terms of the user retention in the team subreddit.

Finally, after building topic model, researchers conducted the analysis of the relationship between team performance statistics and topics of ‘season prospects’ and ‘future’. They found that while top teams are more associated with the topic of ‘season prospects’ and less associated with the topic of ‘future, the bottom teams show the opposite relationship. In summary, the paper is an example of how textual analysis and topic modeling in particular can shed light on the relationship between team performance and audience perception and interpretation of observed events.

However, the topics can be treated in a more specific way than frames of interpretation. For example, Guo, Barnes, and Jia (2017) analyzed reviews on TripAdvisor and described topics as the dimensions of guests experience in the hotels. By conducting LDA authors revealed 30 dimensions of guests experience with 9 dimensions that had not been covered before. Moreover, the authors analyzed the relative importance of experience dimensions in their connection to demographic information of guests and hotels classification.

There was an attempt to analyze dimensions of users experience in the case of items’ purchase and usage before (Musabirov et al. 2017). With use of LDA in the same manner as in the aforementioned work key topics were extracted from discussions on Reddit.com. However, not the topics were in the focus of the work but their combinations which revealed the types of logic and important dimensions of players’ experience that describe cosmetic items’ evaluation.

Structural topic modeling

This work, however, has a different set of research questions that cannot be answered with Latent Dirichlet Allocation, one of the most popular topic modeling algorithms in social sciences. This work is based on framework for topic modeling known as Structural Topic Modeling (STM)
STM is an unsupervised machine learning method that created clusters of words using metadata covariates that make topics composition and model interference more precise. The implementation of covariates in topic modeling allows researchers to conduct regression models between metadata variables and calculated topics.

Though STM is a newer method in comparison to LDA, recently it is actively used in social sciences and research related to experiences (Lynam 2016; Tvinneireim et al. 2017; Grajzl and Irby 2018; Chow et al. 2017). For example, Grajzl and Irby (2018) extracted themes of experiences for students studying abroad with the help of STM and the choice of method was motivated by using metadata that makes topic modeling more precise. Researchers found themes as relating to context of study (e.g. duration and location) as different dimensions of experiences such as immersion in a new culture, history & art, personal growth, etc. Moreover, authors analyzed how students’ demographic characteristics such as gender, age, academic performance, etc. are related to extracted themes. For example, while males shared more reflection on immersion in a new culture and relating to people, the females tended to share more about food culture and social habits.

STM model was conducted with help if R language package stm (Roberts, Stewart, and Tingley 2018). Not only stm allows to estimate topic model of the texts (will help to answer RQ 1), it also allows to estimate covariate effects on topics (will help to answer RQ 3) and calculate correlation between topics that co-occur more or less often than others (will help to answer RQ 4).

Data

Items dataset

Items dataset consists of 1088 unique items and dataset is constructed in accordance to several conditions: 1) items represent each rarity available in the game (see Table 1) 2) items have diverse release year 3) items are old enough to have discussions and to released in market as new items are usually forbidden to trade for particular period of time (see Table 2).

<table>
<thead>
<tr>
<th>Common</th>
<th>Uncommon</th>
<th>Rare</th>
<th>Mythical</th>
<th>Legendary</th>
<th>Immortal</th>
<th>Ancient</th>
<th>Arcana</th>
</tr>
</thead>
<tbody>
<tr>
<td>232</td>
<td>225</td>
<td>343</td>
<td>146</td>
<td>14</td>
<td>111</td>
<td>11</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1. Distribution of rarities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>148</td>
<td>400</td>
<td>347</td>
<td>129</td>
<td>64</td>
</tr>
</tbody>
</table>

Table 2. Distribution of release years.
Based on the distribution of these variables it has been decided to proceed with this data.

Text corpus

The first step was to extract threads in which specific items were discussed. The Reddit API offers a search engine tool that can be used for this purpose. The Reddit API was used to obtain the list of threads that mention an item from the items’ dataset. One search query included one item name, and as a result, 1088 search queries (one query per name) were done by automatic script. For each query, no more than 100 threads were gathered and threads were sorted by a number of comments since even the most popular items usually were not mentioned in a larger number of threads. Due to several items being mentioned in the one thread, some thread URLs were in the list several times. After all the duplicated URLs were removed there has been 2213 unique URL.

The next step after getting the list of thread URLs is collecting the comments in each thread. Package RedditExtractor for statistical language R\(^5\) provides such a tool that takes list of the URLs and collects the comments on the given links. The package collects up to 500 most upvoted comments and ignores the rest of the comments which is not a problem because the package collects the most popular comments -- the comments supported by the larger number of community members. Moreover, not every thread contains that number of the comments.

As a result, 103504 comments were collected and the comments with at least one item name were labeled by the script that detected the names in the text (see Table 3). In total 1821 comments out of 103504 include at least one name of the item.

<table>
<thead>
<tr>
<th>Length of the item name (in words)</th>
<th>Number of mentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 word</td>
<td>717</td>
</tr>
<tr>
<td>2 words</td>
<td>880</td>
</tr>
<tr>
<td>3 words</td>
<td>257</td>
</tr>
<tr>
<td>4 words</td>
<td>103</td>
</tr>
<tr>
<td>5 words</td>
<td>114</td>
</tr>
<tr>
<td>6 words</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total # of unique comments</strong></td>
<td><strong>1821</strong></td>
</tr>
</tbody>
</table>

Table 3. Distribution of names in the corpora

The comments on Reddit have a tree-like structure where one comment is one branch of the tree with plenty of smaller branches. One comment has replies which are treated as children nodes and replies to replies are also children nodes of this comment. In this way, the topic raised (e.g. name of the item) in the parental comment is related to replies but detection of the item names does not consider replies as related texts.

\(^5\) [https://cran.r-project.org/web/packages/RedditExtractoR/index.html](https://cran.r-project.org/web/packages/RedditExtractoR/index.html)
For that purpose, all the comments that included the item names were united with the comments subsequent in the structure so that replies to a comment, replies to these replies, etc. that potentially can be related to an item are also included in the analysis. The extended dataset includes 4766 comments.

Market price dataset

Information about the price dynamics of the items was gathered from steamcommunity.com web API. One query collects the price dynamics data on the specific item and gives back a list of dates for a particular item with a number of sold items and the average price of sold items for each date. In total, 1089 API queries were made and data of 999 items was gathered as some items were absent on the market at the moment and API could not process the given queries. Web API replicates information of price dynamics from item’s page on the item:

![Figure 4. Example of price dynamics on web-site](image)

Price per day was taken in order to detect how discussions are related to price change. In this sense, it was necessary to transform data as the price in absolute numbers did not represent the price change and the interpretation of the given variable could be wrong.

For that purpose, price change in comparison to previous day was calculated. For each item, the order of days was defined and then price change was found by subtraction of the price on day N out of price on day N-1. All the price changes that showed less than 10% change were equated to zero. The remaining value were categorized as Price Increase and Price Decrease. As a result, the final variable on price consisted of three categories: “Price Increase”, “No Change”, “Price Decrease”.

<table>
<thead>
<tr>
<th>Date</th>
<th>Item name</th>
<th>Text (truncated)</th>
<th>Price change category</th>
</tr>
</thead>
</table>

23
Analysis

Research Question 2. Pearson correlation

To analyze the relationship between item’s mentions frequency and average price I created a dictionary with item names, extracted item names from the texts and calculated the frequency of each item. Market data was aggregated by average per item name which allowed me to compare the frequency and market price. To estimate relationship, pearson correlation was calculated.

Research questions 1 / 3 / 4. Structural Topic model

Before conducting topic modeling it was necessary to prepare the model by deciding what covariates will be included in the model and by defining the optimal number of topics in the model.

The first step was to define covariates and Structural topic models use regression model formulas for that purpose. The regression formula was \( prevalence = \text{~} Price \ \text{change category} \) meaning that only the variable of price change was involved in topic modeling.

The second step was to define optimal number of topics. For that purpose there are statistics that help evaluate the model quality. In order to choose the optimal number of topics, several models with varying number of topics (between 10 and 45 with step of 5 topics) were calculated and diagnostic of models was conducted.

According to guidelines (Silge 2018; Grajzl and Irby 2018), it is necessary to find a trade off between high exclusivity (exclus), high semantic coherence (semcoh), high held out probability and low residuals value. The number of 35 topics was chosen because held out and semantic coherence were not lowest, exclusivity was almost the highest and residuals were lowest (see figure 6). As a result, topic model on 35 topics was conducted and the rest of the research questions could be answered.
The topic model identified 35 topics with different sets of the most associated words in each topic. Since, identification of topics is based on unsupervised machine learning algorithm, extracted topics are not affected by the biases of the researcher. However, the interpretation of topics based on the most probable words is a reflection of researcher subjective evaluation. To reduce the effect of subjectivity, it was decided to use two kinds of scores for words’ association with topics and the most associated texts for each topic.

For each topic two lists of words were presented. Firstly, there were 7 words with highest probability in the topic which is based on words’ frequency in the topic. However, in this case, the most frequent words in the whole text dataset would appear more often. For that purpose, there also 7 words with FREX (FRequency and EXclusivity) score which combines word’s probability and its exclusivity for particular topic. In this way, FREX tries to find the word both frequent and specific for particular topics which should make interpretation of topic more precise than it would be in case of analysis only bayesian probability.

To answer the RQ 1 (What dimensions of players’ experience occur in the discussions of virtual items?) topics were interpreted. Based on the most probable words and example messages it was possible to understand each topic. Table 6 shows an example of information necessary to interpret the topic. Based on the most probable words look, set, like, awesom, cool it could be concluded that topic unites the words that express evaluation of appearance and example text supported this interpretation. After topics were interpreted, distribution of their proportions in text corpus was calculated.

<table>
<thead>
<tr>
<th>Topic ID</th>
<th><strong>Highest prob.</strong>: look, set, like, realli, one, pretti, good</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td><strong>FREX</strong>: look, pretti, awesom, cool, good, gold, shadow</td>
</tr>
</tbody>
</table>

I mean some of the gold ones looked alright. Like the Gold Lina dress was nice, but moreso
for the spell effect, the gold riki blades from a while back look good. Golden Gravelmaw is ok, since it looks like a gold ingot for a hero of the earth. Gold Fortune's Tout was fine since those cats can often be golden to signify wealth or some shit. Gold Shadow Demon and AM from the trove look decent too.

Table 6. Example of topic

In order to answer RQ 3 (How dimensions of players’ experience interact with price change of virtual items?) the effects of covariate “Price change category” were calculated for each topic. Structural topic model calculated linear regression model for each topic, so there were 35 models and each model included coefficients of relationship between topic and each unique category: Price increase, No change, Price decrease. The primary interest of the study was how discussions are related to price change, that is why coefficients for “No change” category were withdrawn from analysis.

Lastly, to answer RQ 4 (How dimensions of players’ experience interact to each other?) correlations of topics were extracted from the topic model. Topic correlations show that topics co-occur together more or less often than other topics which would help to understand what combinations of topics are robust enough to be analyzed.
Chapter 6. Results

In this section I present the most important findings of the study and I do it in the following order: firstly, I will talk about RQ 2 as it has separated method of analysis which is frequency counting and pearson correlation, then I discuss RQs 1 / 3 / 4 as they are united by common analysis method which is Structural topic modeling, and lastly, I summarize the findings in a short form.

Research Question 2. Relationship between item price and discussions

Item’s mentions count per se is an important characteristic of the item’s value. Figure 6 shows the fifteen most mentioned items and demonstrates the unequal amount of attention to the items in the discussions. For example, the fifteenth item Yulsaria’s Glacier in the list gets six times fewer mentions that the most mentioned item Heat. What can possibly explain the difference in the popularity of items in the discussions?

Pearson’s correlation analysis shows a moderate positive relationship between an item’s average price and mentions count (cor = 0.4, p. value < 0.001). In other words, the more expensive items draw more attention to the discussions. However, the mentions count is not related to the item’s average number of sold copies (cor = -0.05, p. value = 0.4). Figure 8 shows the distribution of items by their mentions frequency and average price on the market.

Figure 7 demonstrates the presence of outliers that affect the correlation value. To overcome the problem, outliers were removed from dataset and correlations were recalculated. After removal of outliers (the items with average price more than 5 000 and mentions count higher than 200) correlation between average price and mentions count decreases to 0.2 but it is still statistically significant (p-value = 0.002778). The correlation between average supply remained insignificant. In this way, while scarcity of the item is not related to discussions on Reddit, the price of item certainly is. It means that more expensive items draw more attention or probably attention in the social media makes the items more valuable assets.
Figure 7. Distribution of items by price and mentions in discussions

Research questions 1a-1b. Dimensions of players experience

This section will describe the distribution of topics in the discussions and will include the description of the most important topics. I will start with the analysis of the distribution of topics and then present each topic worth knowing in a detailed manner. For each topic I present topic label formatted in bold text, most probable topic words that are formatted in italics. I also refer to text of Reddit.com posts and comments I found during my work to support my interpretation of topics:

"Pieces of discussions are put in double quotes and paragraphs have larger indent on both sides of the page in comparison to normal text."

If topics would be distributed equally, each topic would take 2.85% of discussions (100% divided by 35 topics). However, there is a group of topics much more frequent (see Fig. 8). These topics will be described in more detail as dominant topics of the discussions as well as the topics that stand out in others parts of the analysis despite having smaller share. The rest of the topics described with topic words and examples of texts in Appendix B.

Topics Cool look and Animations and effects each take about 7.5% of discussions which is dramatically larger than other topics. With regards to share, the closest topics are The item changes and Fix the bugs and they take 4.7% each. The next group are topics Expensive rare items, Market scam, Drop chances, Copying the items that take about 4.3% of discussions each. Topics Evaluation of items slots, Items with lore, Store sales, Comparison of Arcana items, in turn, take 3.5% of discussions each. Lastly, the topics Favorite cosmetics and Bugged Items occur in 3% of discussions each.
The topics **Confusion on rarities**, **Thanks fixing bugs**, **chances calculation**, and **Items with lore 2** do not take large share in discussions but they are important for the further analysis and will be described as well. The found topics can be analytically separated into three categories be relation to hedonic value, social value or expectation mismatch.

![Figure 8. Distribution of topics in the text corpus](image)

**Hedonic Value**

**Cool look**

The most widespread topic is named **Cool look** and the most probable words in this topic are *look, pretty, awesome, cool, good, gold*. The topic expresses the evaluation of items by the players and the evaluation is related mostly to item’s appearance or visual effects:

“I mean some of the gold ones looked alright. Like the Gold Lina dress was nice, but moreso for the spell effect, the gold riki blades from a while back look good. Golden Gravelmaw is ok, since it looks like a gold ingot for a hero of the earth. Gold Fortune's Tout was fine since those cats can often be golden to signify wealth or some shit. Gold Shadow Demon and AM from the trove look decent too.”

Though the positive characterization words are the very probable for this topic, it does not necessarily mean the positive evaluation of the item. In this way, the item’s evaluation can be based on the comparison with other “better” items that actually look “good”:

And as for the Crimson, especially for this year, that really comes down more to choice of what to apply the crimson too. Like there’s only really two crimson immortals that actually look good, and that's the spectre and void ones. The other three were ugly as they were already, but the red didn't help matters. Though they
kind of designed themselves in to a corner with that one since 8 of the immortals are already red to begin with so a crimson witness version would just be redundant.

Favorite cosmetics
In the topic **Favorite cosmetics** the players mention item names and characters that are their favorite and this preference is expressed in the most probable words of the topic: **sniper [hero], like, green, favorite[te]**. In this case the players either discuss the cosmetics they like the most:

“Death Prophet’s skirt is probably my favorite. Such beautifully made, and the custom effect is just awesome. Overall, I'm satisfied!”

Or their favorite characters that they would like to have a new cosmetics:

“Okay, I know you are probably thinking that I'm being desperate here guys, but hear me out. We NEED a Sniper Arcana. Sniper is by far my favorite hero and I'm surprised to see he made it as far as he did.”

Equipment slot evaluation
The topic **Equipment slot evaluation** consists of the equipment slots names **blade, item, mask** and characterization words (e.g. **nice**). The topic describes the practice of the players to relate to concrete equipment slots when discussing the items:

“Jugg [Juggernaut (hero)] relic sword, i like 100p jug which bobs when he runs. Also looks naked. Happy mask to make enemies rage at me when i wtf slash em”

The players discuss particular slots due to several reasons. First of all, the discussions are partly related to the visual effects and usually the one item of the whole set provides the given effect. Moreover, the items from different sets can provide different visual effects and it possible to combine the effects if two items with the effects occupy the different equipment slots.

Secondly, the players have a practice of combining the set pieces into a unique look. Dota 2 players even created a separated subreddit to exchange opinions on what items create the “coolest” appearance⁶. The practice was discussed in a detailed manner in the previous study.

As a result, the players can discuss and evaluate the items without connecting to an original set. In this way, players mention the equipment slot either because it is a part of a name (e.g. Reliq sword [of the Kuur-Ishiminarl]) or because it helps them to be more precise about the item (e.g. Bladeform Legacy⁷ name does not show that it is a mask).

Animations and Effects
The topic **Animations and Effects** expresses the discussions related to visual effects and animations given by items which include words **anim[ation], gem, effect, kinet[ic], item, chang[e]**. People discuss effects in the two contexts: when they evaluate items and when they learn how effects interact with each other and how effects are assigned to an item. While the first context is expressed in the topics **Cool look** and **Favorite cosmetics**, the topic **Animations and**

⁶ [https://www.reddit.com/r/dota2fashionadvice/](https://www.reddit.com/r/dota2fashionadvice/)
⁷ [https://dota2.gamepedia.com/Bladeform_Legacy](https://dota2.gamepedia.com/Bladeform_Legacy)
**Effects.** As it was said before, in the discussions with this topic, the players try to make sense of the system:

“[User 1] Let me paraphrase,
Item + [Kinetic] Gem = Animation
Item - [Kinetic] Gem = No Animation
Item + [Kinetic] Gem from other Item = Animation

[User 2] As long as the word "retain" means "to keep having something" I understood everything correctly. The item won't retain the animation after removal of the gem, which corresponds to "Item - [Kinetic] Gem = No Animation" line in your message.

[User 3] Well from my PoV I think he's saying "If I remove the [Kinetic] gem, will the [Kinetic] gem still have the custom animation?"

The presented discussion shows that players were not instructed by game developer about system of visual effects. By experimenting and using publicly available information the players learn the rules of modifying the items:

“A lot of people are, however, confused as to what the gems actually give and [here](http://dota2.gamepedia.com/Kinetic_Gem) is a good resource on what you can expect from transferring each particular gem."

Developer not only does not provide the players with an information, the designers change the rules of assigning the effects to the items:

”"This gem was originally used to grant custom animations to Muh Keen Gun, but the animations were later added to item itself due to clipping issues." As per the wikipedia."

**Items with lore**

The topic **Items with lore** reveals the discussions of items that corresponds to either game lore or game elements. Usually, the players discuss the decorative items that look like functional items. Since purchase of functional items does not affect the appearance of the hero, the players make lists of their cosmetic duplicates:

“Here is a list of the specific items that alter the icons. Note that for some of the icons, multiple items will change them (Like 3 of the Gyrocopter items will change Rocket Barrage's icon)"

The words related to the topic are usually the names of functional items or names of the heroes those cosmetic items can be purchased for: *lanc* [Hero name: Phantom Lancer], *diffus* [Functional item’s name: diffusional blade], *edg* [Functional item’s name: Demon Edge], *phantom* [Hero name: Phantom lancer], *brewmast* [Hero name: Brewmaster], *reaver* [Item name], *centaur* [Hero name: Centaur Warrunner].
Expected mismatch

Comparison of Arcana items

Another example of items’ comparison is the topic **Comparison of Arcana items** which consists of the words *arcana, lina, catalyst, cap, hair, legion (names of heroes and items)* and describes the discussions of items with Arcana rarity:

“I have a few Arcana and am of a split mind. On the one hand, you're looking at specific ones and not thinking the example through on everything. Legion Commander and PA both completely remake the model, and add duel wield, and the particles are basically just weapon or general not mattering what items you have on, so arguably any weapon could be added to the arcana and map the particles on the same way all doom swords basically use the same particles at different lengths for the blade, sure.

But look at Crystal Maiden, whose arcana is a cape, where almost all the other cape pieces are all rigid, so would need to be altered a lot to fit, and many are wider (like the Crystalline Regalia) or start from different places (Tundra Warden, Icebound Floret) that would look awful trying to fit them onto the arcana model with or without attempting to animate them. It's not to say some should work, but it'd really need to be an all or nothing deal.”

People actively compare arcanas because the rarity grants extremely significant effects yet it does not define the visual effects of each particular item. In other words, items of the same arcana rarity can have different types of modifications. For example, quote above describes an Arcana item that changes the whole hero model (“Legion Commander [hero] and PA[hero] [arcanas] both completely remake the model” ) and an item that changes only the slot of back (“look at Crystal Maiden [hero], whose arcana is a cape”).

It creates confusion among the players about the properties of Arcana items because it is not clear what defines Arcana rarity. Moreover, the lack of consistency leads to breakage of other cosmetic items (“and many are wider [...] or start from different places [...] that would look awful trying to fit them onto the arcana model”).

Confusion on rarities

Another topic describing the same problem is **Confusion on rarities** however it covers the discussions of different rarities and focuses on the item properties instead of names: *rarity*, *scheme*, *color*, *ursa (hero name)*, *set, item, like, immort, one, look* such as color scheme and item slots. The players are trying to make sense of the rarities because of the inconsistent effects given to an item:

“Not even immortals [rarity]. There are legendaries [lower rarity] that do more than immortal things. They REALLY fucked up item quality with last year's International.”

Moreover, the same rarity was assigned to particular items in different circumstances:
“Immortal should have been reserved for items with special circumstances, like DC hook or Timebreaker.”

In the last quote the player describes the change of Timebreaker which rarity was changed from Uncommon to Immortal. The reason for change was that the original version of Timebreaker was a replica of the item from the different game. To avoid the scandal, the developers changed the item’s appearance, removed it from the trade system and assigned a different rarity to the copies remaining the players’ inventories.

At the moment developers, release tens of Immortal items annually. They are not as rare as Timebreaker and they have more intense visual effects which creates confusion about what it is for item to have Immortal rarity.

Fix the bugs

The topic Fix the bugs is related to words fix, heat, bug, claw, just, cosmef[ics] which partially express the desire of players to fix the bugs in their cosmetics (e.g. Viridus Claw):

“Void’s portrait while using Viridus Claw has been fixed, but using Tentacular Timelord head without Viridus Claw makes tentacles clip with default model's chest because of the stance. I use both items, the tentacles get really bad. I just played without it :( If they fixed Viridus in a day, it is possible more small fixes are coming.”

In texts related to this topic the players usually share their frustration on mismatch between expected and observed result of the cosmetics. However, the topic does not limit with mention of bugs. In Dota 2, some heroes undergo the model rework which changes the default appearance of the hero and usage of the old cosmetics on the new models can be broken:

“There is one more thing that bugs me. Using any Head cosmetic gives us default jaw/mouth/teeth, the ones old model used. And new one is so much better, I hoped thy would merge cosmetics with new jaw, but they haven’t. I also noticed this, all the cosmetics still have the old mouth. I dont think they will fix it, and if they do, it will take a long time imo I wouldn't even call it a fix.”

Bugged items

The topic Bugged items is similar to the topics Fix the bugs and consists of the words bug, axe, loadout, blood, use, also. Though the topic reveals the theme of bugs too, the content of the most probable texts is slightly different. In the discussions prevalent in this topic the players rather share the lists of the items with bugs than express the annoyance:

“I've reported it a couple weeks ago on the devs forum (after the update that changed the loadout) but it's still not fixed.

Also, not related to cosmetics, but when you click the icon ingame to check Rubick’s equipment, [he looks too big](http://i.imgur.com/iF0mgll.jpg) and does not fit the screen properly. skyhigh warship bomb does not show up cloak of inscrutable zeal clips with all rubick heads other than the one the set is for. Yulsaria's crown (cm immortal) sometimes bugs out on death, leaving CM bald enamelled shield of sir davion clips with his leg in loadout stance
Emerging dragon missile does not show in game when combined with the sky high warship set*.

Thanks for fixing bugs

In contrast to the topics Bugged items and Fix the bugs the topic Thanks for fixing bugs expresses messages addressed to game developers with the words of gratitude:

“Thank you guys for all your hard work.

The Terrorblade arcana still has a bug, when you swap with someone who doesn't have the arcana and you die and respawn with the arcana, the model doesn't have a back and the icon doesn't change either on the top.*

However, as it can be seen from the example, the gratitude is entailed by the complaints on other bugs. The messages of this topic does not seem to express frustration though. On the contrary, players try to help the developer to fix the bug:

“> All of the Snowball Stingers have the correct gem now, if they had room free for a gem to be placed in them. 478 Snowball Stingers had all 5 socket slots taken, so their owners were granted a new Snowball Stinger for each they owned that was full [developer’s message]

I had a stinger with 5 gem slots but only 4 filled but I was given another one. Perhaps there was some glitch because I had renamed it or had the open gem slot as being the 3rd one, that's all I can think of. Still thank you Valve people for fixing it.*

Market scam

Topic Market scam express the discussions of situations when the players became or felt like the target of scam. The words trade, scam, price, worth, steam related to this topic can characterize every message about these situations:

“I also took some extra SS [screenshot] to illustrate this. The same seller is listing two BladeBiters for the same price, but one doesn't have the Kinetic gem ([url]), while the other has ([url]). Another example: one seller is listing a BladeBiter with gem for a determined price ([url]), while another seller is listing a Bladebiter without a gem for a higher price ([url]). Now this is where it gets iffy. In the case of the first two screenshots, it's obvious there's a difference yet the price doesn't reflect that and in that case it *does* seem like the seller is trying to profit from people's naivet. Actually a scam.”

As previous topics showed, not every player understands how system of visual modifications works as rules for particular items are inconsistent and here is no instructions of developer. In most cases, kinetic gem defines the presence of visual effects and kinetic gem can be replaced into another item. It gives the sellers an opportunity to raise the price for items that “should” have visual effects while they actually don’t since the kinetic gem is removed. Players give that responsibility to the developer who implicitly supports market scam by keeping flawed interface of community market:
“Although I don't believe this is a scam per se, they are clearly taking advantage of Steam’s poor interface and people’s ignorance (naiveté). We have to remember that plenty of people use the Community Market in order to avoid sharks and scammers from the usual trade channels. Supposedly the Community Market provides a safe way for you to get your item and one that doesn't require you to be on full alert to not be scammed.”

Copying the Items

Another topic related to market scam is also related to kinetic gem[s] and reflects the situation when software bug let the players make copies of items. The situation with Copying the items dropped the prices of particular expensive items and forced the developer to ban every player somehow related to copying. The players discussed mostly the procedure of items copying (duping) and consequences for other players:

“I didn't do it myself (bought 2 sigs) but I'm all for having equal opportunity. I guess it's fair. But its shitty for people who had legit items. I didn't personally have some, but I would be mad if my Dendi signature went from 150 to 2$ overnight because of a dupe abuse.”

While the players felt sorry for the owners of original items who could spend hundreds of dollars on purchase, they still bought the items because the price rapidly dropped and purchase became feasible. In the example the player shares his thoughts about digital signature of professional Dota2 player Dendi whose autographs cheapened by 75 times during one night (see Fig. 9).

This topic can extend the discussions of market scam as the players sold copied items that were perceived as “fake” ones. Moreover, the copied items differ from the original ones in visual representation. Hence, the players shared their opinions about purchases of copies:

“Player X: i bought some REAL autographs for 2C and not the kinetic fake things.[…]”

Player Y: i hope that the fake kinetic things get deleted some how and that the real autographs rise in price again...lets see:)” (Musabirov et al. 2017)
Figure 9. Example of item’s price fall after duping incident

Social value

Expensive rare items

The topic **Expensive rare items** consists of names and the names are related to scarce items that difficult or impossible to get in any way other than on the market: *dragonclaw [hook]*, [dragonclaw] *hook*, *courier*, *unusu*[al], *golden*. There three types of items mentioned in the most probable words. Items such as Dragonclaw hook that were prohibited from being obtained in anyway; Unusual couriers, in return, are couriers that were modified by players to have a glow around the courier’s model until the functionality of modifications was stopped by developers. Lastly, golden couriers are golden versions of some couriers that either have dramatically smaller drop chance in comparison to standard ones or need players to accomplish very difficult in-game activities. Mostly players discuss the price of such items and their scarcity:

*Alpine Ursa Set - Worth ~$1k+, roughly 250 in circulation
Dragonclaw Hook - Worth ~$200, roughly 10k in circulation*
Golden/Platinum Baby Roshan, rewards for 2012/2013 Diretide - Worth Upwards of $1k, roughly 50 Golden Roshans and 100 Platinum Roshans in circulation

With regards to items' high price the players discuss their quality as well:

"Not really, price has nothing to with how good an item is. Does not mean I don't like them but I'd rank Timebreaker and Dragonclaw a lot lower than most immortals [...] If Dragonclaw was released for the first time today or even 2 years ago, it would not be an immortal. For sure yes."

The discussions on quality and price happen because the players perceive such items as less aesthetically appealing because such items are usually very old (two or three years prior to discussions) and their design is not compatible with the quality level of the newer assets. It creates the collision between the scarcity and beauty of the item. Players understand this collision yet some are ready to purchase the items.

Moreover, due to their scarcity and stable price these items are discussed as a replacement for virtual currency:

"The only reason these items are worth huge prices now is because of their immortal status. As far as I can tell its not really the cosmetic appeal, but used more like currency. Since there is a limited number of those items available, the price can be reasonably stable."

Chances calculation

To some extent, the topic Chances calculation is similar to the topic Expensive rare items as it reflects the discussions of scarcity and drop chances. However, this topic is more of reflection of players calculating the chances to obtain the item. It mostly consists of the words chang[e], lucki, open, roll, row, chest which describe the discussions of opening the treasure chests. There are long threads with calculating the chance to get the ultra rare item and making sense of the drop system as a whole:

"We assume there are 10 different items.
Probability of getting the same item 10 times is: (1/10)^10
Probability of getting 10 different item is: (1 * 9/10 * 8/10 * ... 1/10)
Probability of getting Item A, Item B, Item C in this particular order is: (1/10)^10 your system assumes that the choosen items are no longer in the mix, but indeed they are each chest has a fresh set of all 10 items and therefor the probability is 1/10^10"

Among calculations made there are messages that are aimed to contest the statistics as some players face getting the same item several times in a row:

"Opened 5 IT chests in august and got 5x craniomancer in a row but if you tell people the system is rigged they all go "bullshit it's totally random!"

Moreover, the players compare the drop chance with item’s value which perceived much higher than it should be:"
“Buying a golden empyrean for 300$ which originated from a 15$ luck chance doesn't make any sense.”

Drop chances

Topic **Drop chances** extends theme of chances and treasure chests. However, while the topic **Chances calculation** is mostly focused on the sense-making of the drop system as a whole, topic **Drop chances** is focused on discussions of Treasure chests with items of Immortal rarity, the value of chests and value of the items in those chests:

“The immortals will probably end up cheap. If you just wanna get them now, go ahead. If you want to get it for trade, probably best you don't. The only immortals that might end up &gt;1$ will be the rare ones. For example, I have the enigma eidolons and they go for like 3.50 euro, the warlock immortal goes for like a euro. However, the most expensive TI3 immortal is actually the Spectre Soul Diffuser and that wasn't rare and it's like 7.62 euro.”

Other

Store discount sales

Besides scamming and copying items there are other ways to decrease the price of items (e.g **Store discount sales** which extracted to separate topic). The most usable words in this topic are set, sale, lucentyr, luna. First two words relate to discussion of purchases:

“That's it folks, **sale is over**!
If you didn't get the set you wanted in the short time window, check Steam Market!
Maybe you can get it there by a friendly price!”

Words lucentyr and Luna, in return, relate to particular set “Blessings of Lucentyr Set” for particular hero named Luna. Apparently, this set was one of the most popular assets on the sale. However, another possible explanation is the announcement showed some of the sets before the sales and the set was one of them:

“Two invoker sets and luna one will be in it 100% cause they listed as examples for what will be sold”

While sales in Dota2 Store organized by developer are not surprising, it is interesting to see particular topic describing it. Having one separated topic for sales means the discussions of this kind are not spread in the comments but rather focused in particular threads and related to specific items. In contrast, if the sales were discussed in relation to most of the items, it would be present in most of the documents and had not been extracted in a separate topic.

The item changes

The topic **The item changes** describes communication about one particular item Timebreak[er] which changes weapon slot of the hero named Faceless Void. The discussions of this item were united under one topic because the item’s design was changed as the original design was a
copy of another item in another game. To avoid the scandal the developer removed the item from the trade and changed the design of purchased assets:

"[User A] Can someone explain the timebreaker situation?
[User B] It was an item for Faceless Void that turned out to just be a direct port of an item for another game which made it onto the store. Valve had to remove the item, but also had to make a different replacement version to give to people who had already bought it."

Event venue

The topic **Event venue** describes the discussions of tournaments during those years. In particular, the players talk about The International which is the biggest tournament in Dota 2 organized by game developer and accompanied by release of many items, treasure chests and in-game activities that let the players obtain items. However, the topic reflects the tournament itself, the venue, and esports side of the game: professional **players**, their **teams** and organizations, and **sponsorships**:

"Interesting read. I'm on the other side of things, gathering sponsors and making sure they get value for sponsoring events and such. The sponsors I typically get are those trying to sell a product or brand, much like the groups that sponsor tournaments. I went through your post and here are my thoughts on a few things:...

**Relationship between topics**

However, players do not necessarily discuss one particular topic in the thread. Sometimes the discussion can be a mix of several topics. When the topics was interpreted the next step was to analyze the combinations of topics that occur together more often than others. Structural topic model calculates the correlations among topics analyzing what topics occur together more often and less often than others. The correlations were used to build the networks of topics and find the pairs of topics that are often combined together. The values of correlations varied between -0.11 and 0.125 which shows that overall the discussions are mostly focused around one topic and there are weak relations between topics.

An algorithm cuts off the correlations that lower or equal zero which leaves even the correlations of 0.01 in the network. As there were no recommendation on what correlation is acceptable for structural topic model, the pairs of topics with correlation higher than 0.04 were analyzed thoroughly.

Since each topic assigned to each text with a particular share, the most correlated topics should have an equal proportion in the texts. To verify whether the correlation is meaningful or random, distribution of texts for each pair of topics was visualized. For each pair of topics the texts were selected to have equal and as high as possible share of both topics. Figure 11 demonstrates the visualization of topics **Market scam** and **Expensive rare items** and shows the distribution of texts by proportion in both topics. For example, text 1324 has a proportion of 0.25 for topic **Expensive rare items** and close a proportion of 0.5 for topic **Market scam** meaning the text is characterizing topics almost equally.
The correlation was called meaningful if the text examples could help to explain the relationship between two topics. Based on analysis of the texts, the pairs with correlations lower than 0.05 could not get meaningful interpretation, so the network of the topics was cut off at this level. The remaining pairs formed the network presented in Figure 10.

Figure 10. Network of correlated topics

The resulting network consists of four topics: **Expensive rare items** which related to **Market trade scam** which related to **Tournament venue** which related to **Store sales**. For each pair I present the distribution of documents by topics’ share in the document. Furthermore, I present examples of texts that have equally high share of both topics and describe the interpretation of the given pairs.
The pair of topics **Market trade scam** and **Expensive rare items** (Fig. 11) has been almost equally present in the texts 262 and 1324. Below are quotes from both texts:

"Like valve caress about pricing, look what they did to dragonclaw hook [expensive item] at one point. What did they do? I see the price dropped from $260 to $125 over ~10 months after steadily increasing. Was that due to Valve's action?\nThey changed it to mythical and added to drops list for some time." [262]

"My Pudge's inscribed whalehook from ti3 I'd say. Not too attached to most of my stuff. I like to sell all my stuff to buy more games. Just sold an unusual courier for $13 and bought hotline Miami. Thank you dota" [1324]

The examples demonstrate that players discuss trade of expensive items. In general players discuss what and why they trade. For example, the text 1324 depicts the player's way of purchasing new games through selling the items he earned in the Dota 2.

Besides general discussions of trade it is important for particular players to reflect on game changes and how it affects the trade. For example, in the text 262 the player complains about decreased value of the item and tries to understand why it happens (presumably due to change of rarity and adding into drop lists).

As a result, the pair of topics is based on discussions of expensive items and reflections on their monetary value. These items can have inadequate price for traders (as the example shows, because it is lower of expected price) and this inadequacy evokes such discussions among players.

Another pair of topics is based on topics **Market trade scam** and **Tournament venue** (Fig. 12) and reflects the discussions on tournament quality. The players share their feelings about
attending the tournament and judge if it was worth it or not. The same kind of communication is related not only to a tournament as a whole but to its particular aspects. Text 610 (text 611 is a quoted version of this text) is an example of such evaluation:

“10) FOOD TRUCKS ARE TRASH. Food is trash, they are scams imo, everything is 6-15$ for 5$ worth of food, standing in line in heat is a nightmare, don't waste ur time. Walk to a safeway or fast food restaurant if you really want to, and buy something made in a room that isn't 300 degrees Fahrenheit.”

In the example the user complains about service quality and nothing particular about the items in the messages. The next pair of topics, however, uncovers an interesting aspect of Dota 2 cosmetic items system.

```
Figure 13. Distribution of text documents between topics Tournament venue and Store sales
```

The last pair of topics consists of discussions of **Store sales** and **Tournament venue** (Fig. 13). Example texts demonstrate the connection between topics on the level of item names. The topics have common names frequently mentioned in texts. However, some of those items are associated with professional organizations and players. The texts 459 and 712 demonstrate this connection:

```
“Spent 10 \xa4 no more money U_U
Bought Chains of the Black Death [Associated with DotaCinema]
Dread of the Gleaming Seal [Associated with Alliance]
Bindings of Deep Magma [Associated with Evil Genius]
Blessings of Lucentyr
Pauldron Perch with kinetic [Associated with Dendi]
I hope the market prices will drop for the earthshaker and luna set. People stockpiled a lot of the Luna set to sell, that one might drop. Not sure about Earthshaker. Why? maybe so that he can afford one?”
```
"Is lucentyr set next? Based on Chinese websites, Lucentyr would be one of next ones. I hope so, managed to get pudge’s and shaker sets’ would be great to get 3 for 1 Yeah, from what I saw today, Black Death [Associated with DotaCinema], Gleaming Seal [Associated with Alliance]; Deep Magma [Associated with Evil Genius] sets were such a good bargain. Lucentyr would be too. When exactly is the next one starting? I really don’t want to miss the luna one. It’s up. Yeah Thanx mate. It’s up. It's Lucentyr now. :D" [459]

Apparently, when players mention the items, they mention team and player names as well, and probably they have the context of the tournament and teams participating in it in their minds. While it is difficult to say if the players discuss and evaluate the players when mentioning the items but they certainly use names when explaining what item they talk about. For example, "I'm trying to the the [A] PA set" [piece of 457 text, team alliance set].

**Relationship between topic and price dynamic**

![Graph showing the relationship between topic and price dynamic](image)

**Figure 16. Regression coefficients of topics with significant coefficients.**

Full version is in Appendix A

When the regression model on topics and price dynamics was done, there were four topics (see Appendix A for visualization of coefficient for other topics) that had statistically significant relationship with price dynamics category: **Animations and Effects, Thanks for fixing bugs, Expensive rare items, Confusion on rarities**. Though presented effects are not larger than 0.07, they describe change of topic proportion in texts which on average is 0.0285. Despite small number, shown effects can express rapid change of topic proportion in comparison to its proportion in the whole corpus.

The topic **Animations and Effects** has a strong positive relationship with price increase (coef = 0.047, p.value = 0.002). Based on the given coefficient it can be suggested that the more players discuss item’s visual modifications the higher its price or vice versa since the causal relationship is unknown. Perhaps, such discussions encourage players to buy more which
increase the price or increase in price make people discuss why monetary value of items has increased.

Another topic that is positively associated with price increase is Expensive rare items (coef = 0.037, p.value = 0.002). Like in the previous case, the more discussions on this topic occur on Reddit, the more expensive it becomes. Though, like in previous case, it is impossible to say the exact size of the effect, but regression shows that discussions on this topic are related to price increase. The interpretation is the same as in the previous example.

Topic Confusion on rarities, in return, have a strong negative relationship with price increase (coef = -0.022, p.value = 0.002). It means the discussions on this topic occur less often with regards to items which price increase. It does not mean the discussions are related to price drop but it shows that discussion of the confusion with item’s rarity certainly does not increase its price. Alternatively, when player discuss why price has increased, they look for other reasons rather than inconsistent rarities.

The only topic that have positive relation with price decrease is topic Thanks for fixing bugs (coef 0.032, p.value = 0.001) which makes this topic be twice more probable in texts related to price decrease. As analysis of topic shows the gratitude is accompanied with complaints on unsolved bugs and probably complaints are prevalent in the discussions on this topic. As a result, the growth of the topic can reflect the higher number of people who suffered the bug or price decrease could evoke discussions on the possible reason for price drop.

Results summary

RQ 1a. What dimensions of players’ experience occur in the discussions of virtual items?

RQ 1b. What is the relative prevalence of the extracted dimensions in the discussions?

During analysis 35 topics were extracted and labeled. The topics reflect different dimensions of experiences, and those dimensions have different importance for players as their expected proportion varies between 0.5% and 7.5% of discussions. The most important topics Cool look, Animations and effects, The item changes, Fix the bugs, Expensive rare items, Market scam, Drop chances, Copying the items, Evaluation of items slots, Items with lore, Store sales, Comparison of Arcana items, Favorite cosmetics, Bugged Items, Confusion on rarities, Thanks fixing bugs, Chances calculation were analyzed in a great detail using the most probable words and the most probable text documents.

Analysis showed three groups of dimensions that can be connected to three theoretical concepts: hedonic value, social value, expectation mismatch. Topics Cool look, Animations and effects, Evaluation of items slots, Items with lore, Favorite cosmetics reveal hedonic value (Lehdonvirta 2009) of items as they appeal to aesthetic quality or emotional reaction of the players. Most of the topics related to aesthetic quality but there are also topics related to lore which triggers positive emotional response (Lehdonvirta 2009). Moreover, there are two types of
aesthetic judgement: players either judge whole appearance of item (Cool look and Favorite Cosmetics) or discuss particular mechanics of making the item beautiful (Animations and effects and Evaluation of items slots).

Topics Expensive rare items, Drop chances, Chances calculation express social value (Lehdonvirta 2009) of items because they reflect discussions of items scarcity and its relation to items value. In this context people discuss two things in particular: how scarce items make owners more visible in the community and how the system of getting the items works.

Topics Fix the bugs, Market scam, Copying the items, Comparison of Arcana items, Confusion on rarities, Thanks fixing bugs, in return, reflect discussions of expectation mismatch (Musabirov et al. 2017) when the players do not get what they expected. The players can express their confusion about items’ perceived value and their price (Comparison of Arcana items, Confusion on rarities) or about failed design that breaks trade (Market scam, Copying the items) but mostly players complain on the visual bugs and explain why they happened (Fix the bugs, Thanks fixing bugs).

RQ 2. What is the relationship between items’ mentions frequency and their market properties

Analysis also showed that on discussions have different patterns on the level of individual items. First of all, analysis showed that price of item is related to its popularity in discussions. Aggregated price of item is positively associated with number of mentions in the discussions and expressed in statistically significant correlation estimate equal 0.2. It means, the more discussed items are more expensive on the market. However, the popularity of item is not related with price scarcity (number of sold copies), meaning that players tend to discuss expensive items rather than the scarce ones.

RQ 3. How do dimensions of players’ experience interact with price change of virtual items?

In an attempt to find alternative mechanisms of price formation, extracted topics of discussions were analyzed in connection to price change of item in comparison to a previous discussion of item. In summary, topics reflecting hedonic (Animations and effects) and social value (Expensive rare items) are positively related to price increase. It means that if the price of item has increased in comparison to previous discussion, the proportion of those topics is higher. Expectation mismatch, in return, has either negative relationship with price increase (Confusion on rarities) or positive relationship with price decrease (Thanks for fixing bugs).

RQ 4. How do dimensions of players’ experience interact to each other?

In total, when correlations between topics were calculated three interactions were found in the model but only two of them were of interest in this study: pair Market scam - Expensive rare items and pair Tournament venue and Store sales. First pair reveals the discussions of expensive rare items in the context of market trade. Especially players focus on how monetary value of items depreciated due to developer’s actions. Second pair, in return, reveals the
esports nature of Dota 2 by connecting the topics of items sold on sale and items related to esports brands such as professional players or organizations.
Chapter 7. Discussion

Based on the analysis of Reddit.com discussions of virtual goods in their connection with price dynamics, this study brings new insights into players experiences connected with virtual goods evaluation, extending previous studies in this area.

First of all, it was found that items’ popularity in the discussions correlates positively with their price: the more frequently the item is mentioned in the comments the higher its average price on the market is. In contrast, the number of copies sold does not significantly correlate with discussion popularity, i.e. items that are bought often does not necessarily spur discussions, while expensive items do.

This focus of discussions on expensive items is reflected in the results of topics modeling which captures a particular discourse of discussing these goods (e.g. Expensive rare items and Store sales) and is connected with status dimension of social value or emotional dimension of hedonic value (Lehdonvirta 2009; Marder et al. 2019).

Aesthetic value in appearance and visual effects

Topic model reveals the evaluation of aesthetical value on several levels and describes the experiences of different kinds. Topics Cool look and Favorite cosmetics describe the general evaluation of the items where players judge what they find cool (Raptis, Kjeldskov, and Skov 2013) and what they do not like. Topics Animations and effects and Evaluation of item slots, in return, are focused mostly on more specific kind of evaluation that is connected to a specific set of experiences. First of all, the visual effects have bugs which can ruin the whole experience while item’s appearance cannot be broken. Moreover, the topic Animations and effects is connected with positive price change which demonstrates that visual effects play a major role in the experience of players.

The vast prevalence of such topics supports the previous theoretical and empirical findings (Lehdonvirta 2009; Marder et al. 2019) who found the visual representation plays an important role in player’s decision to purchase non-functional items. Not only does the analysis conducted in this work demonstrate that aesthetic value plays the vital role in evaluation of non-functional items (range and number of topics associated with aesthetics) but also it shows how the discussion of visual effects reflects the increase in the value of the assets. Moreover, analysis shows that aesthetic value can be decomposed into different dimensions as players describe different experiences when talking about Cool appearance (Raptis, Kjeldskov, and Skov 2013) and Animations and effects.

Another dimension of hedonic value that occurred in discussions is Lore of items (topics Items with lore) which was suggested to be an important part of emotional value because it includes “related background fiction or narrative presented to the user” (2009, p. 106). Though lore is not as important as aesthetic value because shares topic related to lore are much smaller (2.7% - 3.5%) than shares of topics related to visual effects (7.5% for Cool look and 7.5% for visual Effects and animations), it is surprising to see the presence of this dimension in Dota 2, the game with no story or thoroughly described universe. Probably the game without a rich story
does not depreciate the lore element. On the contrary, shortage of lore makes it a tool of judgement about items quality. However, further research needed as the relation between lore and item’s price remains unclear.

**Social value and scarcity**

The presence of such topics as **Expensive rare items, Drop chances, Chances calculation** reflect the interaction between item’s availability and its price on the market. This interaction is what defines social value because scarce items have more social value as they are better at highlighting the status of players. However, these topics focus on different aspects of this experience. For example, **Expensive rare items** reveals status dimension of consumption as it consists of discussions of expensive items that are valuable due to their scarcity and its presence have a strong positive relationship with price increase (coef = 0.03, p.value = 0.0024).

Topics **Drop chances** and **Chances calculation**, in return, are focused on mechanisms of artificial scarcity which are being decomposed by players during discussions.

The theme of status dimension is supported by the studies that identified the scarcity as a primary factor of price formation (Yamamoto and McArthur 2015; Lehdonvirta 2009; Marder et al. 2019). The scarcity in this sense is discussed mostly as a tool to highlight the status of the owner (Lehdonvirta 2009; Marder et al. 2019).

Moreover, analysis of topic **Expensive rare items** revealed the discussions of expensive items as replacement for currency in trade. According to the text, the expensive items are a great currency due to their limited availability and stable price on market. Though usage of assets as money in barter-like trade has been mentioned in previous studies on virtual consumption (Yamamoto and McArthur 2015; Lehdonvirta, Wilska, and Johnson 2009), there is a difference that makes that example stand out.

In previous studies, the candidates on being money were widespread items that were so cheap and large in number that their value was stable regardless game updates and other factors. For example, Yamamoto and McArthur (2015) described keys in “Counter-Strike: Global Offensive” that opened cases with virtual items as a currency that became a unified currency in unofficial trade platforms. Lehdonvirta (Lehdonvirta, Wilska, and Johnson 2009), in return, explained that players “denominated prices in plastic chairs” (p. 1072) because the game itself did not provide the players with the currency for trading.

In contrast, in Dota2, the expensive items become this replacement for currency. It can be interpreted as a sign that Steam community market is not a place for “premium” trading with unique items. Such trade deals, possibly, get done on different platforms: the most famous case of EF Pink War Dog being sold for $38 000⁸ was done in auction on special subreddit r/Dota2Trade⁹ which took 2 minutes the original owner to find a buyer.

---


⁹ [https://www.reddit.com/r/Dota2Trade/comments/1q0kxp/auction_ef_pink_war_dog_191_78_123_with_bo/](https://www.reddit.com/r/Dota2Trade/comments/1q0kxp/auction_ef_pink_war_dog_191_78_123_with_bo/)
In this way, Steam community market can be an inappropriate place to trade expensive items due to the tax charge that service takes from each transaction and impossibility to withdraw money from the Steam. In this situation, the players set barter-like trade where expensive items become a unified currency for the serious traders.

**Expectation mismatch**

topics **Market scam** and **Copying the items** reflect the mismatch between players’ evaluation of items and their market price which is perceived as unfair and frustrating experience. Besides those topics, there are several groups of dimensions related to expectations mismatch. In particular, analysis showed the presence of bugs and inconsistency of rules in the system of cosmetic items.

Topics **Fix the bugs** (4.7% of discussions), **Thanks for fixing the bugs** (2.4% of discussions), and **bugged items** (2.9% of discussions) reveal the discussions of visual bugs that players face using cosmetic items. As having the real product broken makes the users feel frustration, bugs have the same effect on players in case of virtual world.

The difference is the real life trade has legal regulations that insure the customers in the case of broken products. In Dota 2, it is difficult to get compensated for bugged item. As a result, the players tend to value the bugged items lower. This conclusion is supported by topic **Thanks for fixing the bugs** which includes both gratitude and new complaints and which is positively associated with price decrease (0.029, p.value = 0.0017).

Having the bugs can be an example of expectation mismatch (Musabirov et al. 2017) when the real value of the asset does not comply with expected one. The users do not get what they expect to get and though developers try to avoid the bugs, it cannot be always under control. When bugs occur, the developer must choose what is the company’s priority. Either developer can focus on fixing the bugs of the items that already sold and spend resources for no new revenue or developer focuses on creation of the new content that will bring revenue in the future.

At first glance, creating the content can be more profitable option as bugs occur at already sold assets but the collision between bugs fixing and creating the content is not easy to solve. As analysis of discussions shows, the discussion of item’s bugs is related to their price decrease. In this case, the developer gains less money from commissioning the trade and loses potential income because the company product will be less trustworthy. In this way, the choice between creation the new content and fixing the current problems costs money in both scenarios. That is why it is important to find the balance between bugs fixing and ignorance of the problem.

Reddit discussions provide us with another example of expectation mismatch based on inconsistency of rules introduced by developer. Such topics as **Comparison of Arcana items** and **Confusion on rarities** reveal the attempts of players to understand what rarity actually represents
The inconsistency of rarity properties makes the players to figure out the rules collectively in the discussion and appealing to wiki-like websites\textsuperscript{10}. They argue for how good item should be to have a particular rarity, what visual effects it must have and what bugs player expect to get when mix arcanas with other items. As a result, discussions of rarity inconsistency and \textbf{Confusion on rarities} are negatively associated with price increase (coef = 0.022, p.value = 0.0028).

\textbf{Judgement devices of items evaluation}

The discussions also express the presence of judgement devices (Karpik and Scott 2010) used by players in purchases of items. First of all, topic \textbf{Confusion on rarities} demonstrates that Rarity as class must give the players some understanding of item’s value; however, as it is seen in discussions, sometimes it does not work this way. Moreover, not only it does not help the players to decide on the item’s value, it can contradict with players’ opinion about what items are more valuable than others. This example clearly shows that judgement devices designed to help the players can fail in their purpose.

Second of all, analysis revealed the vast presence of lists created by players. Themes of lists are related to various aspects of game: lists of items with bugs, lists of items with higher lore value, lists of items with lowest chance to get (and calculations of chances). Lists per se can be a judgement device as they demonstrate the players whether the items are ‘cool’ or, on the contrary, bad.

Lastly, this thesis revealed the presence of brands in discussions of items because particular items officially associated with teams and players\textsuperscript{11} were often mentioned in the texts that combine topics \textbf{Tournament venue} and \textbf{Store sales} (corr = 0.12). It can be suggested that in discussions of \textbf{Tournament venue} the users discussed esports teams and professional players and mentioned the items related to them.

This thesis supports some findings of the prior work (Musabirov et al. 2017). Though previous paper analyzed different data and used a different approach, there are several themes that occurred in both works. First of all, as it was mentioned before, the study revealed the importance of hedonic and social values. Secondly, current work found the presence of esports nature in discussion of items as particular items officially associated with teams and players were often mentioned in the texts that combine topics \textbf{Tournament venue} and \textbf{Store sales} (corr = 0.12). It can be suggested that in discussions of Tournament venue the users also discussed the esport teams and professional players and mentioned the items related to them.

\textsuperscript{10} https://dota2.gamepedia.com/Rarity
\textsuperscript{11} https://dota2.gamepedia.com/Brand_Items
Chapter 8. Conclusion

In this thesis I present the results of the study of consumption experiences of players in Dota 2. Based on data of Steam community and discussions of virtual items on Reddit, I analyze how players’ communication reflects value of cosmetic items. Using Structural Topic Modeling I extract topics from the discussions which I perceive as experience dimensions of consumption. In result of analysis 35 of topics were extracted and the most prevalent topics were described.

Analysis showed the presence of three general categories of observations: dimensions of hedonic value, dimensions of social value and dimensions of expectancy mismatch. In hedonic value there were found three most common dimensions which cover general evaluation of appearance, evaluation of visual effects and lore. In social value, dimensions of status and chance were found, and also was found the practice of barter-like trade with expensive items. In category of expectancy mismatch were found themes of inconsistency in rules, bugs, and market system breach. While some dimensions of hedonic and social value have positive relationship with price increase, dimensions of expectancy mismatch are either positively related to price decrease or negatively related to price increase.

Theoretical and practical implications

The role of expectancy mismatch in items evaluation is the main theoretical contribution of the work. While the prior works mostly focused on factors that encourage the players to purchase virtual assets, this study demonstrate the factors that discourage the players from doing it.

Besides that, the study contributes to understanding of consumption experience in several other ways. First of all, it verifies prior works which was mostly based on ethnographic observation or interviews. Study supports original findings related to values of non-functional goods (Lehdonvirta and Ben). In particular, analysis showed the presence of hedonic and social value in the discussions. At the same time, the thesis extends current discussion by decomposing values into several dimensions. For example, the hedonic value can be split into item’s appearance and presence of visual effects both of which are of aesthetical value for players.

However, the work does not corresponds to prior work on items collections which have two possible explanations. Probably, the collections of items are discussed on other platforms (as practice of creating unique look has moved to separate subreddit). Alternatively, while the present among players, the practice is not present in the discussions. Possibly, it is necessary to reconsider the role of social sharing in creating collections (Toups et al. 2016).

In addition, the work sheds light on several dimensions that are not covered to a great detail in the field yet. The case of confusion on rarities can be an example of rules inconsistency which frustrates the players affects the assets pricing. Another case of experience dimension is lore which is only briefly mentioned by Lehdonvirta (Lehdonvirta 2009) despite playing a visible role in players’ experience.

The work also has practical implications that would be useful for practitioners of game design and community construction. Two cases in particular can be considered in the design of online
markets. Firstly, the study revealed the impact of bugs on the market which makes the problem not just disturbing to players’ experience but harming developer income. Secondly, the study demonstrates the importance of consistent rules in virtual items system. Collision between items’ characteristics and player’s perception of items’ quality can become a real problem for developer as confusion on rarities certainly does not reflect the increase in price. Moreover, the ability to detect such problems by analyzing texts gives developers an opportunity to use discussions related to their game as a form of feedback about game.
References


Appendix

Appendix A

After attempt of interpretation topics **Game updates, Bugged visuals** were withdrawn from the analysis. Those topics had word Heat as one of the most probable words which meant temperature heat and one of the abilities in the game\(^\text{12}\). They were withdrawn as they were mostly connected to price dynamics of item named Heat due to polysemy.

Appendix B

[https://docs.google.com/spreadsheets/d/1LoynS844zqhr6HfuTwTStr9h3EBzLZjgsdTXKR4hkhE/edit](https://docs.google.com/spreadsheets/d/1LoynS844zqhr6HfuTwTStr9h3EBzLZjgsdTXKR4hkhE/edit)

URL link grants access to the table with topic words and examples of texts for each topic.

---