Patches and Player Community Perceptions: Analysis of No Man's Sky Steam Reviews

Chien Lu, Xiaozhou Li, Timo Nummenmaa, Zheying Zhang, and Jaakko Peltonen
Tampere University
Kalevantie 4, 33100 Tampere, Finland
firstname.lastname@tuni.fi

ABSTRACT

Current game publishing typically involves an ongoing commitment to maintain and update games after initial release, and as a result the reception of games among players has the potential to evolve; it is then crucial to understand how players' concerns and perception of the game are affected by ongoing updates and by passage of time in general. We carry out a data-driven analysis of a prominent game release, No Man's Sky, using topic modeling based text mining of Steam reviews. Importantly, our approach treats player perception not as a single sentiment but identifies multiple topics of interest that evolve differently over time, and allows us to contrast patching of the game to evolution of the topics.

Keywords

No Man's Sky, Player Modeling, Topic Modeling

INTRODUCTION

Steam is one of the biggest digital game distribution platforms. In addition to being a platform for purchasing and playing, it is also a community where members have mutual discussions and game owners can write game reviews to share their opinions and game-play experiences. When writing a game review, the player can label the game as "recommended" or "not recommended". How much time (hours) the player has played the game when writing the review is also recorded. Steam reviews form an important view into a game's reception among players; an overall summary of the proportion of positive reviews is prominently reported on a game's Steam store page and is also often reported on third-party sites and in various news and social media discussion on a game. However, such an overall summary does not reveal the main concerns reviewers report on, or the change of their perceptions over time.

Steam allows for games to be easily updated once released. Developers release updates on their own schedule; on the players' side, available updates can be set to happen automatically, making it easy for users to keep their games up to date. The dynamic of such updates is very different compared to the earlier practice without a unified platform, where users needed to download patches from developer or publisher websites. Steam also supports the sale of downloadable content (DLC), allowing developers to esily add paid content to already published games. In addition to normal game releases, Steam has a programme called early access (*Steamworks Documentation: Early Access* Accessed 8 Dec 2019). It

Proceedings of DiGRA 2020

allows developers to already sell games that are still in development and not ready for a final release on the platform. The impact of such an environment, where game updates are commonplace, on player perception has not been explored in game research in a comprehensive quantitative way. In this work we do so in context of a particular game, No Man's Sky.

No Man's Sky is an action-adventure survival game first launched in August 2016. It received strongly critical responses from players due to lacking features that had been promised to be in the game. However, since launch the game has been supported by rapid updates; it has had 8 major updates so far, denoted by versions 1.00, 1.10, 1.20, 1.30, 1.50, 1.70, 1.75 and 2,00, released on 12 August 2016, 26 November 2016, 8 March 2017, 11 August 2017, 24 July 2018, 29 October 2018, 22 November 2018 and 14 August 2018, respectively[1].

The changes to No Man's Sky exemplify several typical types of change in digital games. Digital games, as other software products, evolve during their lifecycle. The evolution of games can be in the form of emerging change (designing a space for the players to mold their own game experiences), reactive change (changing the game by reacting to direct or indirect feedback from the players) or pre-planned change (content that is already designed, or in some cases already produced, before the launch of the game), and their evolution has similarities to how more utility focused software products evolve, but not all properties are equal (Nummenmaa et al. 2013). No Man's sky is a mixture of all three types of evolution. Emerging change is built into the system, even if lightly, as the game world is generated piece by piece when users access new worlds. Reactive change is prevalent, as new changes are implemented due to feedback in the form of patches. Pre-planned change is also present, as the developer has implemented features that have already been promised prior to release. As Newman (2012) has pointed out, due to two factors, ports (transferring to different operating systems or platforms) and patches (updating or adding new features), a game itself is an unstable object. Due to this nature of games, the player perceptions or experiences of a game are therefore dynamically changing over time.

Despite research on various aspects of game development, and other research on impact of game updating strategies, there is a lack of research coupling large-scale analytics of different aspects of players' reception of a game to aspects of the update strategy, and in this work we do that. Our research questions are: RQ1 - What are the main topics of discussion (e.g. themes of concern or appreciation) in players' reviews? RQ2 - How do the contents of reviews change over time, are some topics rising or falling over time, and at what rates? RQ3 - How does the presence of the topics differ in reviews recommending versus not recommending the game? RQ4 - How do updates carried out to the game coincide with changes in the topical content of the reviews?

To answer the research questions, in this research, a collection of more than 85 thousands user reviews across roughly 3 years of *No Man Sky* from steam were analyzed. We use a machine learning based text mining technique called topic modeling to analyze the collected data in a computational manner. The model we use both extracts topics as semantically meaningful themes in players' reviews, and also models the relationship between the presence of those themes, time of the review, players' playing hours, and players' attitude in the are distilled/extract their relationships between players' overall attitude towards the game in the sense of whether they recommend the game or not.

In the following we first discuss a selection of related work on challenges in game development, customer feedback and review analysis, update strategy planning, and analysis of updates and reviews on Steam. We then discuss the method from data collection to the text analysis. Next, we discuss the results first in terms of the extracted topics and in terms of their prevalence over time. Lastly, we provide discussion and conclusions.

RELATED WORK

Many studies have addressed the challenges and issues in computer game development practice from software engineering perspectives (Alves et al. 2007; Kanode and Haddad 2009). Kanode and Haddad (2009) list and specify several challenges in game development in terms of software engineering, including assets, scopes, process, publishing, management, team organization and third-party technology. Ampatzoglou and Stamelos (2010), by reviewing the literature, examine the use of software engineering theories, methods and tools in game development practice and find that game developers tend to fit traditional software engineering methods to game development with certain adjustments. The authors also indicate that the maintenance activities within game development are mainly corrective and maintenance and verification in game development are often neglected. However, the game products can be changed significantly due to the feedback from testing phase and the market (Kasurinen et al. 2014). Thus, enabled by the contemporary online distribution channels, updating games correctively and perfectively can improve them significantly towards enhanced customer satisfaction.

In order to improve products and services, Customer feedback is an important data source for companies to understand the market and the needs of their customers (Cho et al. 2002; Wu et al. 2010). In the software engineering domain, end user feedback is also critical for facilitating the evolution of software products and services (Burnett et al. 2004; Ko et al. 2011). The importance of end users as stakeholders is particularly enhanced for mobile applications, since they are commonly distributed through online platforms (Holzer and Ondrus 2011). The combination of collectable end user feedback and traceable software evolution allows further requirements analysis to be done effectively through statistical and data-driven methodologies, in order to plan future changes and to be aware of how the changes may impact user satisfaction (Palomba et al. 2015). Following advances in data mining, many studies have provided various approaches towards effective review analysis to uncover critical user needs (Fu et al. 2013; Chen et al. 2014; Guzman and Maalej 2014; Li et al. 2018). Hence, despite their differences from mobile app reviews, video game end user reviews can also provide valuable information that game developers can take into account in order to improve their game products (Lin et al. 2019).

The evolution of software products is considered important for maintaining their quality, when together with the widely applied incremental and agile development methods, users receive early releases of software products and are more likely to support their evolution with meaningful feedback from which requirements are elicited and prioritized effectively and continuously (Greer and Ruhe 2004). Thus, an effective planning for software release is highly required. Many studies have contributed to the practice of software release planning, in terms of the process, decision making, strategic models and tools (Ruhe and Saliu 2005; Saliu and Ruhe 2005; Svahnberg et al. 2010). In particular for mobile applications that are distributed at unified online platforms where updates are easily delivered and feedback are

instantly received, Nayebi et al. (2016) find that developers tend to follow predefined rational update strategies and mostly believe that frequent updates and the different release strategies shall affect the app quality and users' feedback. Hence, towards effective mobile application specific release planning strategies that ease the developers' efforts and respond swiftly to users' concerns and complaints, many studies have proposed approaches and tools facilitating mobile app release planning (Villarroel et al. 2016; Ciurumelea et al. 2017; Scalabrino et al. 2017). Comparatively, even though many players still choose to purchase physical copies, the video games online distribution platforms have grown rapidly with the advantages of easiness to find relevant games based on preferences, affordability, easiness of payment, and so on (Toivonen and Sotamaa 2010). The mechanisms of the platforms also enable developers to constantly listen to players' feedback and update accordingly, but also requires them to plan the updates properly, especially for "Early Release" games (Lin et al. 2018).

As one of the most popular digital game distribution platforms, Steam provides not only video game purchasing and downloading service but also online communities for the players to review games and for developers to respond. Due to the notable volume and dynamic of the data contained in Steam, it has been widely used for research purposes (Kang et al. 2017; Lin et al. 2019). Lim and Harrell examine the players' behaviors of profile and social network maintenance and analyze the differences in their player identities (Lim and Harrell 2014). Slivar et al. analyze the the impact of game types and video adaptation strategies on the quality of the experience (Slivar et al. 2015). To investigate player behaviors on Steam, Sifa et al. analyze the players' different playtime frequency distribution and investigate their engagement and cross-game behavior (Sifa et al. 2014; Sifa et al. 2015). Regarding game updates on Steam, Lin et al. conduct an empirical study of the urgent updates of the 50 most popular games and find that the choice of update strategy affects the proportion of compulsory urgent updates (Lin et al. 2017). Furthermore, regarding game reviews, Lin et al. perform an empirical study on the reviews of 6224 games on Steam and analyze the review content and the relation between players' play hours and their reviews (Lin et al. 2019).

METHOD

Data Collection

In this study, we use Scrapy[2], an open-source Python-written web crawling framework, to obtain the user review data from the Steam platform, specifically for the game No Man's Sky. Scrapy was first released in 2008 with its latest version 1.8.0 being compatible with Python 3.5 and later versions. Technically, in order to crawl structured review data from the community page of a particular game, we define a Spider with Scrapy and run it through the crawler engine. The crawling process starts with a request on the URL of the game community page and calls the default callback method, which loops through the elements (i.e., review items) with CSS selectors and yields a dictionary with the requested information. Notably, Scrapy is able to crawl the content loaded with Javascript via users' scrolling that cannot be obtained using BeautifulSoup[3], which is crucial for crawling Steam user review data.

As a result of the crawling, we obtain the 85805 unique user reviews on No Man's Sky from its release date, August 12th 2016, to October 5th 2019. The features of the data include

review publication date, review text, user ID, the recommend/not recommend flag for the review, the user's play hours, the number of products owned by the user, the number of people who rated the review as helpful and the number of people who rated the review as funny. Among the obtained reviews, over half (44335) were given within the first month of the game release, with 59.14% of the users not recommending the game. During the timespan of the data set the overall recommendation rate has increased to 53.03%.

Text Analysis

We employ a text analysis technique called Structural Topic Model (STM, Roberts et al. 2016) to analyze the collected review texts. STM has been applied to analyze gaming discussion on game development (Lu et al. 2019) and trophy hunting (Lu et al. 2020). Topic modeling represents document content as a mixture of underlying topics, each of which has a distribution of typical words; these underlying topics and their prevalences in each document in the collection are found by fitting the topic model to the data set. The resulting topics and their prevalences over the documents can then be analyzed. The resulting topics can describe subjects of discussion, but can also describe other elements such as tone of writing. Among a set of multiple topics, some may differ greatly from one another while others may be more similar, describing differences of emphasis within a common theme. Compared to methods such as Latent Dirichlet Allocation (LDA, Blei et al. 2003) and Dynamic Topic Model (DTM, Blei and Lafferty 2006), the STM technique that we use is a more advanced model which is able to take available document-level covariates into account when modeling the text. For Steam reviews, we take into account several covariates in the modeling: the user recommendation (recommend or not) indicates the general positive/negative evaluation of the game, thus it is taken as one of the covariates; we also take the posting time as a covariate in order to model the evolution of the review content over time; the user's playing hours which reflects their amount of experience with the game as a player is also included as a covariate in the model.

Before the model traing, stop words (e.g. 'is', 'this', 'etc') and rare words (words that only appeared once in the whole text corpus) were removed and all the words were then lemmatized. The lemmatization technique takes morphology into account and can find unified forms for more complicated cases, such as irregular verbs (e.g. 'drive', 'drove' and 'driven' are lemmatized to their common lemma 'drive').

The final model was decided based on the criterion called held-out likelihood. To compute the value of the criterion, a proportion (50%) of a small subset of the collected documents is considered unobserved ("held out") and is not used to build the topic distributions, and the STM models are evaluated by their likelihood on this held-out portion, representing the ability of the models to represent previously unseen text. In our large collection of reviews, it is possible to find a large number of underlying topics, and thus it we chose the number of topics by a careful search. We first searched among topic numbers $K = 10, 20, ...$ to 100, with an interval of 10. After finding that the model with $K = 50$ had the maximum value of the held-out likelihood criterion, the search was focused around $K = 50$, and a more detailed search for possible improved values with $K = 41$ to 49 and $K = 51$ to 59 was conducted. The model with $K = 55$ ultimately turned out the have the best value of the held-out likelihood after the search and was chosen for the model. Note that the criteria of held-out likelihood has received criticism by e.g. (Chang et al. 2009), however, it is still

a common practice to decide the number of topics with STM (Stamolampros et al. 2019).

After deciding the number of topics, the semantic coherence value (Mimno et al. 2011) was taken as a criterion to choose the best model from multiple runs with different initializations. The semantic coherence value measures how strongly the top words in each topic co-occur over documents, thus, it can be employed to evaluate the performance of topic models and to choose the best-performing model among several models. We built 10 models with 55 topics using the whole dataset, starting from different initializations, and the model with the best average semantic coherence value over topics was selected as the final model.

RESULT

We first discuss the themes found in the extracted topics, then we discuss the evolution of the topics' prevalence over time, and we further discuss the impact of the users' play hours on the prevalence of different topics in their reviews.

Extracted topics

The top 10 words of each topic are listed in Tables 1, 2, 3, and 4. In each table, we list for each topic its most common words, and also its overall proportion ("Pr (%)") representing how much of all review content arises from that topic. The topics are also given descriptive names by the authors by analyzing their top words as well as analyzing example reviews that prominently arise from the topic. Notably, the rich review content in Steam allowed us to extract a large variety of topics with clear semantic meaning; this both indicates that players have a rich variety of concerns relating to the game and its development, and shows the benefit of using text mining approaches for review analysis. We next discuss the found topics.

The topic **Evaluating Game-play** with terms such as *play, fun, get, ...* holds the highest prevalence. One example quote is

´´It's fun at first, but gets boring after a while and then weirdly, gets fun again It's a good game, though it has it's flaws 9/10, go buy the game goddamnit".

Followed by topics **Reaching Recommended Status** and **Appreciating Improvements**. They reflect players' positive perceptions after improvements of the game. Other similar topics include **Gradual Improvement** which emphasizes temporal aspects (e.g. with terms such as time, long and due); For example, one quote of the topic **Gradual Improvements** is

´´this has supprised me the last few times i have played. at first i was not overly happy with my purchase but as time has gone by i have become more happy. is it poerfect, no, but it is fun to play now."

A certain amount topics are related to updates, including **Updates and Added Content**, **Change of Game** and **Upgrades and Items**. One example quote from the topic **Updates and Added Content** is:

´´The Foundation Update has added a good amount of new features to the game, including

building bases and owning Freighters. This is just the foundation of future updates so you can see Hello Games are actively working to make the game better. Kudos to them."

Among other topics with higher prevalence, some of them are related to game purchase such as **Worth the Price**, **Pre-order** and **Refund**; some are directly related to disappointment to the promises, one example is the topic **Disappointment to Promise and Hype**, one example quote is :

´´I preordered that game thinking it would be magical after seeing the trailers. It was not what was promised then. I tried again after beyond because they said they made it a lot better. yes some things changed but this game is so empty and ugly compared to what they said it was. the only fun i had was building my base which is now broken after a patch. I don't even want to figure it out. Quests ? boring as hell and there is nothing rewarding in them. thank you hello games, i was naive, now i will never preorder anything again, I'll wait to see what kind of crap a game really is before jumping in...."

another similar topic is **Lies and Miss Promises**, one example quote is:

´´The trailers lie. Most of the stuff in the trailers don't exist in the game. They are selling this to us under false pretenses which is illegal."

Other topics that directly reflect to the disappointments include **Strong Dislike, waste of money**, **Feels Unfinished, Indie vs AAA (Quality Level)**, **Quick Disappointment**, **Lack of Content / Grind** and **Recurring Bad Game-play**. One example quote from the topic **Strong Dislike, waste of money**

´´♥♥♥♥ YOU ♥♥♥♥ THIS GAME. YOU SHOULD GO BANKRUPT . NO ♥♥♥♥ING WAY IM GONA BUY FROM YOUR ♥♥♥♥IN COMPANY GAME. ♥♥♥♥ ♥♥♥♥♥"

and example quotes from the topic **Feels Unfinished, Indie vs AAA (Quality Level)**

´´In it's current state it feels as if I purchased an early access game.", ´´AAA price. Indie gameplay. 11/10".

Despite the negative feelings, there are some topics related to the appriciations including **Appreciation**, **Enjoyment of Play Experience** and **Enjoyment Despite Flaws**. One example quote of the topic **Appreciation** is

´´Best game,a lot of oprtunities and no limits you always have something to do,they have redeamed theirselfs.".

There are also topics that are related to other reviews. **Reviews vs. Reality** with top words such as *review, read, see, okay, say...* is about players' reflections on other reviews or the comparison between the reviews the players have read their own experiences. One example quote of this topic is

´´I really do not understand why there is so uch negative comments and posts about this game. I totally enjoy it and..."

Another topic related to other players' reviews is **Not as Bad as People Say** with top words such as *people, say, review, give, think, everyone ....* Where are some example quotes of this topic that are trying to defend the game:

´´This game is amazing, I woud love to wright a proper review but I am speechles and I dont understand why the reviews are so negative, thats all I have to say, this game is amazing..."

and

´´Some people like it, some don't. Some people anticipated more, some didn't. Some people feel tricked, some don't. Some people love Trump, some don't. I personally think that if this game was named as "Early Acces Game"(which it technically is) then a lot of things would be different"

Technical issues are also discussed in the collected game reviews. Relevant topics include **Graphics Settings**, **Crashes** and **Bugs and Glitches**. For example, one quote from the topic **Crashes** is

´´i cant even start playing, the minute the game starts to load it crashes".

Some topics are related the to details of game-play, including **Moving and Looking**, **Base-building and Desired Content**, **Spaceship Travel and Combat**, **Repetitive Resource-collection Game-play**, **Exploration and Discovery** and **Material Collection**. One example quote from **Repetitive Resource-collection Game-play** is

´´Explore, collect resources so you can keep exploring, repeat..."

Besides, some topic are specifically associated to PCG content, e.g. **Procedurally Generated Universe and Exploration** and **Procedural Content Generation of Maps**.

One quote from the topic **Procedurally Generated Universe and Exploration** is

´´A very interesting game where you explore a procedurally generated universe. It is very chill and if you enjoy relaxedly exploring a universe it's a great time. If you're looking for a survival game or fps or flight sim its not really that, though it has elements of that. Good stuff if you like exploration though."

Topic Prevalence over Time

Figures 1, 2, and 3 show the temporal dynamics of the topic prevalence. The blue line presents the prevalence among players who recommend the game and the red line presents the prevalence among the players who do not recommend the game; dotted lines around them represent 95% confidence intervals of the topic prevalences. The date of updates are also marked with black and green bars; the black bars represent the starting date of each major update and green bars are smaller updates.

In the plots of topic prevalence over time, 20 out of 55 plots show that the prevalence of the topic is always high among players who do not recommend the game; note that Figures 1, 2, and 3 only show a selection of interesting plots; the 20 out of 55 ratio was verified for the

Table 1: Extracted topics, part 1

Topic	Pr (%)	Top Words
Evaluating Game-play	4.46	play, fun, get, buy, hour, pretty, first, bore, good, like, couple, look, day, soon, gameplay
Reaching Recommended Status	3.69	update, launch, still, next, since, improve, change, come, beyond, finally, recommend, original, long, recent, theyve
Appreciating Improvements	3.46	good, great, lot, amaze, awesome, nice, keep, ton, job, perfect, overall, game, work, friend, fun
Worth the Price	2.94	worth, price, sale, recommend, full, definitely, quite, pay, enjoyable, pick, say, tag, good, chill, like
Game Lifecycle, Work of the Developers	2.81	release, year, dev, game, work, late, continue, ago, free, developer, day, finally, week, dlc, month
Interaction with Other Players	2.73	want, try, play, can, give, friend, another, multiplayer, like, else, make, back, someone, think, time
Not as Bad as People Say	2.68	people, say, review, give, think, everyone, hate, like, negative, see, alot, hope, positive, good, personally
Graphics Settings	2.64	run, setting, gtx, graphic, ram, low, high, max, fine, spec, set, smooth, card, window, computer
Disapointment of Promise and Hype	2.62	promise, hype, wait, buy, deliver, disappoint, title, worth, live, hope, game, pay, trash, huge, preorder
Strong dislike, Waste of money	2.54	money, ♥♥♥♥, waste, suck, buy, crap, ♥♥♥♥ing, copy, garbage, paste, piece, back, scam, ever, ♥♥♥♥♥♥
Lies and Missed Promises	2.53	lie, feature, show, advertise, trailer, miss, promise, developer, video, false, product, many, multiplayer, sell, unfinished
Appreciation	2.30	love, time, play, start, hour, keep, cool, idea, absolutely, feel, always, beautiful, first, put, experience
Enjoyment of Play Experience	2.30	enjoy, explore, like, far, bite, thing, play, look, feel, may, relax, slow, find, although, kind
Feels Unfinished, Indie vs AAA (Quality Level)	2.20	feel, content, early, potential, access, current, amount, indie, aaa, simply, depth, extremely, high, shallow, market
Crashes	2.16	crash, load, start, screen, work, minute, play, try, computer, playable, unplayable, past, fix, open, min

Table 2: Extracted topics, part 2

Topic	Pr (%)	Top Words
Lack of Content / Grind	2.11	nothing, grind, anything, real, end, like, literally, reason, whole, thing, basically, way, empty, stay, youll
Comparing with Other Games	2.06	space, exploration, like, minecraft, world, survival, elite, dangerous, combat, deep, open, game, adventure, sandbox, simulator
Quick Disappointment	2.00	bad, spend, half, hell, like, hour, big, tech, look, demo, page, ever, forget, seriously, straight
Issues and Patches	1.98	issue, fix, patch, problem, support, performance, work, need, optimization, edit, technical, hopefully, state, experience, poor
Refunds	1.95	refund, steam, hour, ask, return, realize, alpha, buy, playtime, wish, attempt, hope, policy, game, store
Updates and Added Content	1.93	new, update, add, game, bring, content, community, future, major, forward, feature, stick, atla, foundation, improvement
Reviews vs Reality	1.92	review, read, see, okay, say, write, time, watch, decide, know, check, sure, think, post, need
Change of Game	1.80	stuff, good, make, need, like, decent, super, lot, big, easy, yet, little, still, take, slowly
Gradual Improvement	1.73	however, purchase, time, game, happy, long, due, developer, concept, recommend, effort, offer, regret, massive, point
Discussion of Game Versions	1.72	actually, ever, big, game, version, edit, make, call, good, disappointment, back, put, come, sorry, late
Reactions	1.70	know, please, thank, guy, hard, yes, damn, stop, kinda, wow, good, god, let, work, like
General Opinion Words	1.64	thing, like, game, want, think, enjoy, hear, know, type, say, anyone, believe, fantastic, person, follow
Pre-orders	1.58	never, everything, pre, order, every, almost, like, make, ever, first, sit, see, imagine, look, time
Moving and Looking	1.57	around, turn, take, away, look, see, like, move, head, walk, way, hit, one, figure, blow
Variation in Content	1.54	planet, different, look, animal, see, plant, every, rock, color, thing, similar, variation, type, like, generate

Table 3: Extracted topics, part 3

Topic	Pr (%)	Top Words
Performance	1.53	drop, run, terrible, rate, frame, bad, stutter, horrible, port, lag, optimize, poorly, constant, like, console
Basebuilding and Desired Content	1.53	build, base, player, multiplayer, add, story, make, single, character, good, vehicle, friend, new, still, able
Repetitive Resource-collection Game-play	1.50	resource, galaxy, planet, repetitive, center, hour, gather, find, collect, another, repeat, bore, become, first, reach
Procedurally Generated Universe and Exploration	1.45	universe, experience, story, generate, explore, action, procedurally, vast, discovery, infinite, unique, exploration, visual, wonder, truly
Bugs and Glitches	1.44	bug, save, break, progress, time, con, pro, fix, buggy, hour, glitches, play, start, many, file
Enjoyment Despite Flaws	1.43	many, despite, experience, mod, moment, see, along, yet, incredible, true, flaw, dream, consider, become, world
Hype and Expectation	1.41	expect, expectation, small, hype, game, exactly, team, train, review, think, sci, plenty, fan, people, gamers
Lack of Interest and Reward	1.40	feel, lack, make, interest, reward, simple, little, place, point, gameplay, thing, certain, kind, find, purpose
Minor Complaints	1.37	seem, like, use, right, thing, first, sure, able, otherwise, sad, complaint, minor, look, see, lucky
Developers and Studios	1.36	fact, gaming, game, matter, example, say, develop, value, total, never, history, compare, studio, trust, let
Acknowledging and Expecting Improvement	1.35	man, sky, game, become, wonderful, ever, say, truly, come, leap, good, upcoming, next, experience, freelancer
Exploration and Discovery	1.33	planet, find, alien, discover, learn, creature, system, name, new, explore, word, language, race, fauna, species
Control Difficulty	1.19	rather, less, time, mouse, step, hold, three, require, handle, appear, spore, play, impossible, avoid, button
Recurring Bad Gameplay	1.17	happen, still, make, time, constantly, mess, good, instead, box, thing, apparently, somewhere, need, manage, fill

Table 4: Extracted topics, part 4

Topic	Pr (%)	Top Words
Spaceship Travel and Combat	1.14	ship, mine, fly, space, land, planet, station, tool, fuel, sell, resource, attack, multi, weapon, combat
Death	1.13	lose, kill, back, die, leave, annoy, time, find, shoot, sometimes, spawn, try, inside, death, power
Material Collection	1.11	life, find, every, material, planet, farm, start, time, need, minute, thing, walk, take, tutorial, except
Falling Short of Expectations	1.10	tell, gameplay, graphic, short, fall, suggest, requirement, average, little, good, will, hard, level, discount, significant
Travel Between Star Systems	1.10	system, star, travel, jump, entire, end, light, leave, can, explain, set, life, make, find, take
Control Interfaces	1.08	control, option, menu, flight, hold, change, click, press, key, controller, texture, interface, hand, force, button
Upgrades and Items	1.06	ship, upgrade, inventory, item, slot, suit, fight, find, sentinel, trade, management, space, system, need, blueprint
Procedural Content Generation of Maps	0.98	map, design, generation, sound, procedural, limit, world, surface, element, variety, vary, effect, engine, system, encounter
Survival and Challenge	0.91	survival, mode, craft, mechanic, need, challenge, grind, easy, normal, progression, make, creative, use, satisfy, equipment
Interaction with Factions	0.84	large, planet, battle, see, interaction, faction, trade, giant, space, planetary, terrain, player, creature, war, close
Quests	0.83	quest, freighter, mission, pirate, ability, ship, fleet, space, system, base, use, trade, enemy, multiple, main

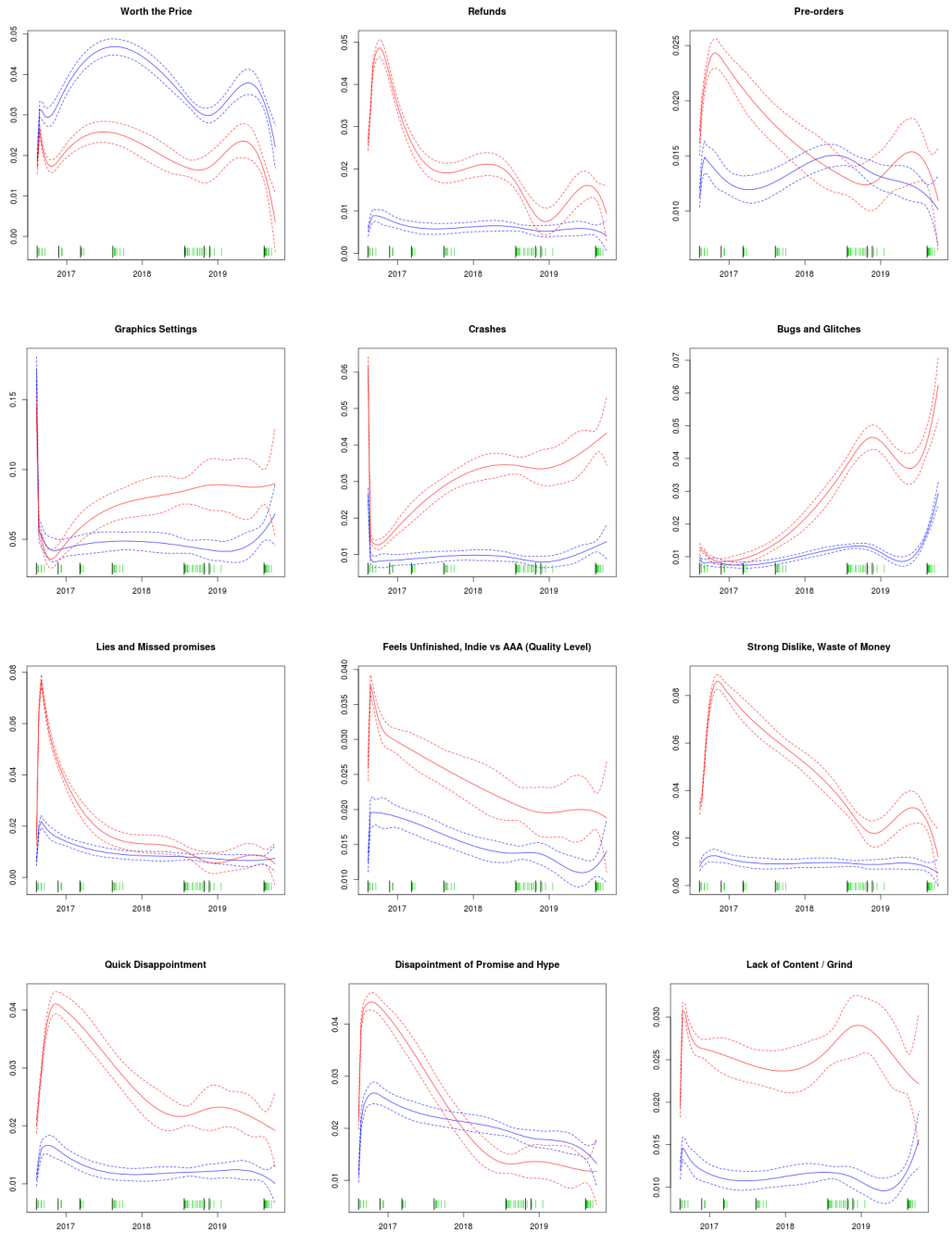

Figure 1: Selected topic prevalence over time. Part 1. Blue: Recommend; Red: Do not recommend; Dot lines: 95% interval.; Black Bar: Time of the updates

whole set of plots. One kind of such topics include players' negative perception and comments on the game, such as **Strong dislike, Waste of money**, **Lies and Missed promises**, **Feels Unfinished, Indie vs AAA (Quality Level)**, **Quick Disappointment**, **Refund**, etc.

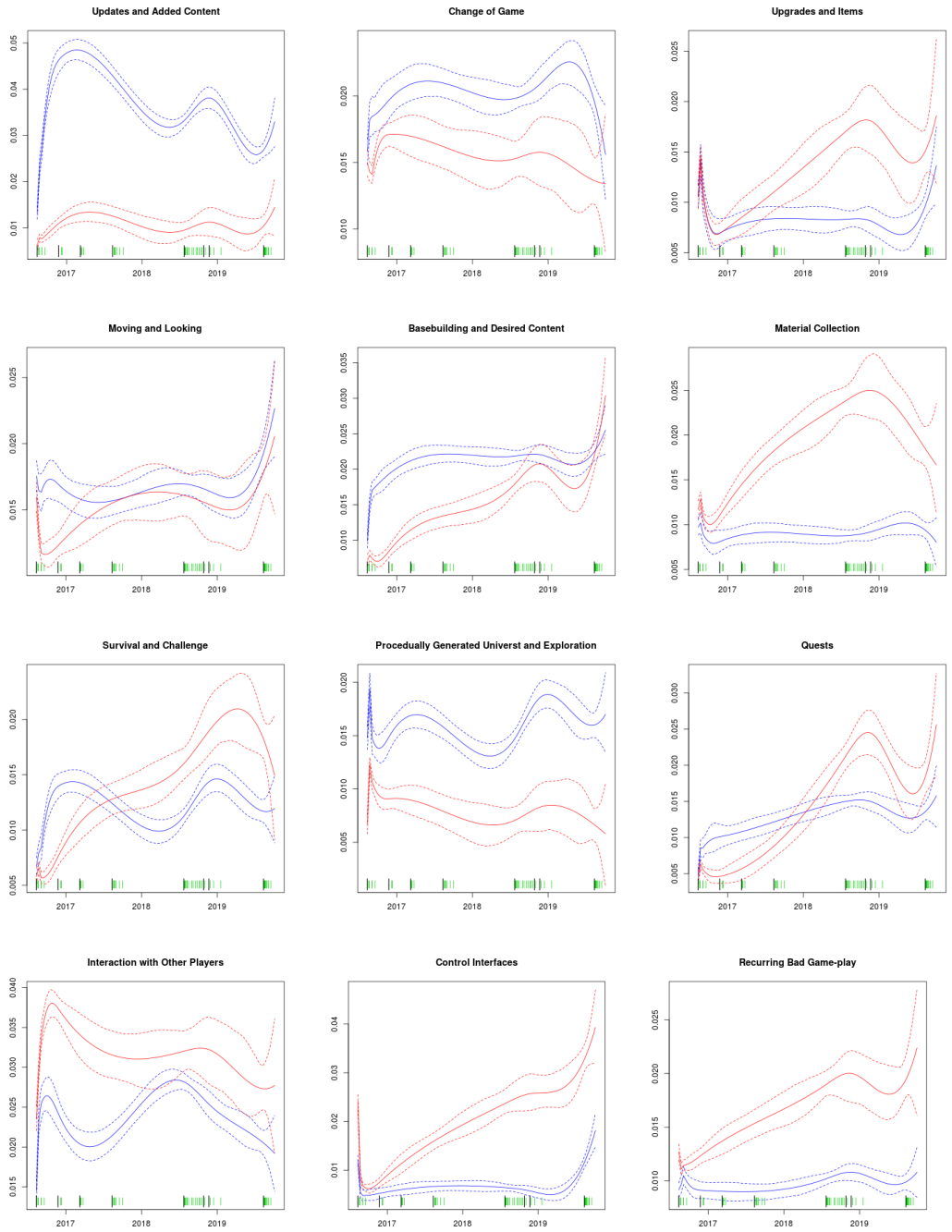

Figure 2: Selected topic prevalence over time. Part 2.

Specifically, the prevalence of these topics immediately reaches the peak after the game was released, and starts declining after the first major release called Foundation (update 1.10) on Nov. 26 2016, and continues declining over the follow-up major releases, which indicates that the game studio's effort on releases do relieve players' strong dissatisfaction. Besides, there are also topics of complaints of game contents or features, and they are **Material Col-**

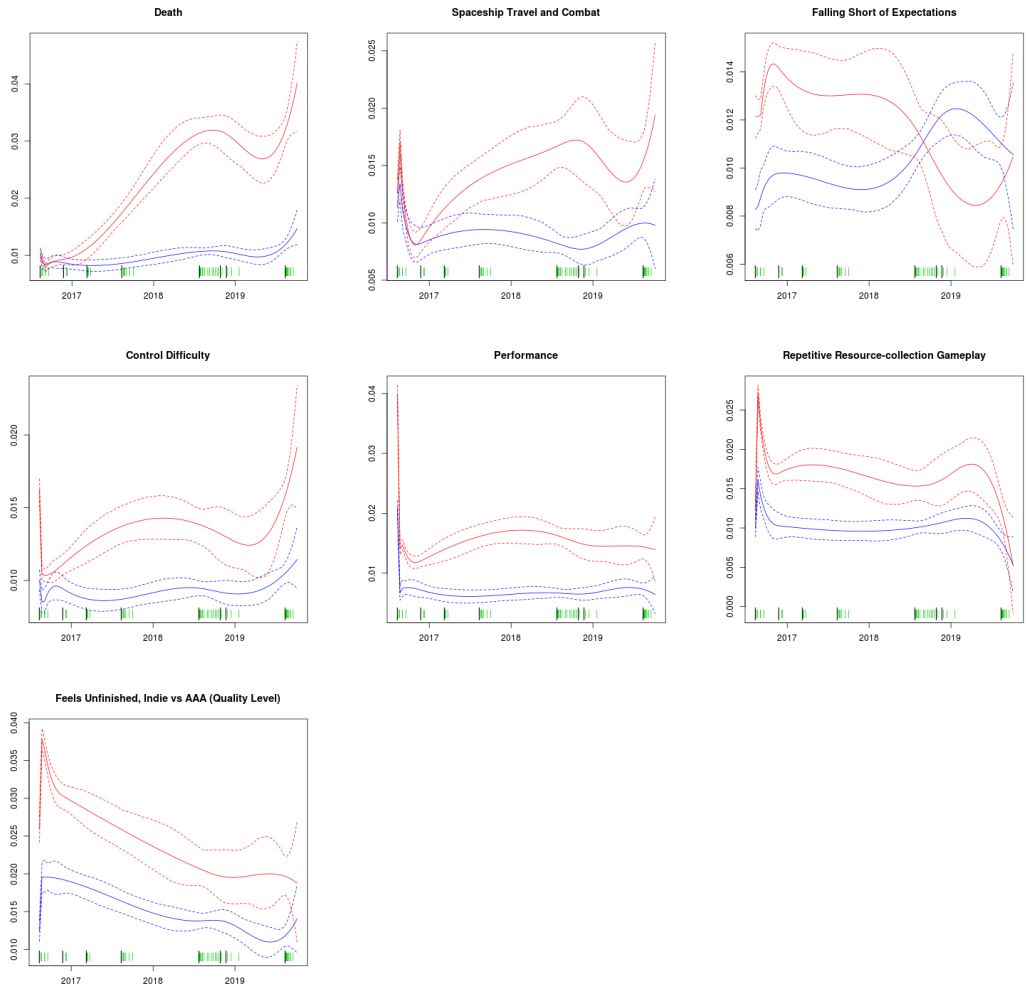

Figure 3: Selected topic prevalence over time. Part 3.

lection, **Control Interfaces**, **Graphics Settings**, **Crashes**, **Bugs and Glitches**, **Recurring Bad Gameplay**, **Spaceship Travel and Combat**, **Death**, etc. Their prevalence increases among players not recommending the game, indicating the bugs and issues remain or are introduced throughout the releases. Some of them might be mitigated in a specific release. For example, prevalence plots of topics such as **Material Collection**, **Bugs and Glitches**, **Recurring Bad Gameplay**, **Spaceship Travel and Combat**, **Death** have a small hump after two major updates in October, 2018 (1.70) and November, 2018 (1.75). Their prevalence declines for a certain period before rising again.

Despite the controversy, there are 19 topics showing more game recommendation since it was released. The topics are related to understanding, appreciation, and acknowledgement of the game and its continuous improvements despite the failure to present the promised features to meet some players' high expectations when the game came out. Some players also defend the game against the complaints and disappointment expressed in other reviews.These clearly implies the game studio's effort on keeping improving games, and

the improvements are appreciated. In addition, the prevalence plots such as **Falling Short of Expectations** and **Disappointment of Promise and Hype** show a decline among players not recommending the game and a rise among players recommending the game. Both support the observation of a gradual increase in players' satisfaction, along with the game updates.

Among game purchase topics, for the topic **Worth the Price**, there is a clear distinction between players who recommend and do not recommend the game, on the other hand, the topics **Refunds** and **Pre-orders** have reached the peak in the beginning and the prevalence of those topics in general has a going-down trend after the peak. The topic **Refunds** might have been driven by the different kinds of early disappointment and reached the peak, One the other hand, **Pre-orders** is a timely topic so the discussion capacity has been done after certain amount of time.

The temporal pattern of the topics that reflect the disappointment varies. **Lies and Miss Promises**, **Feels Unfinished, Indie vs AAA (Quality Level)**, **Strong Dislike, Waste of Money**, **Quick Disappointment** and **Disappointment of Promises and Hype** are frequently discusses among players who do not recommend the game only in the early stage, the prevalence dropped in different time. One the other hand, it seems that **Lack of Content / Grind** is a constantly lasting issue especially among players who don not recommend the game.

When it comes to technical related topics (**Graphics Settings**, **Crashes** and **Bugs and Glitches**), the trend of the prevalence in general has been growing especially among players don't recommend the game. One potential reason is that, due to more and more new features added into the game over time, there is a higher possibility for players to encounter technical difficulties, especially bugs and glitches and result in negative perceptions.

Temporal trend of topics related to game updates (in Figure 2) show different trends in terms of temporal dynamics. The topic **Updates and Added Content** and **Changes of Game** show a overall positive perception and the topic **Upgrades and Items** shows a growing negative perceptions.

Topics related to specific game-play experiences, including **Moving and Looking**, **Basebuilding and Desired Content**, **Survival and Challenge**, **Procedurally Generated Universe and Exploration**, **Quests**, and **Material Collection**, each shows a different pattern. The topic **Moving and Looking** shows a mixture of reviews both from players recommend and do not recommend the game. **Basebuilding and Desired Content** was more prevalent in the beginning among players who recommend the game but the prevalence has grown among players who do not recommend the game and in the end prevalence is roughly the same in both kinds of players. The topic **Material collection** has a growing prevalence among players who do no recommend the game and it reached the peak around the beginning of 2019. The topic **Survival and Challenge** was more prevalent in the beginning among player who recommend the game and in the end it has become more popular in players who who do not recommend the game. The topic **Procedurally Generated Universe and Exploration** has been always popular among players who recommend the game and the topic **Quests** has grown in both kinds of players but the growing changing was more

obvious among players who do not recommend the game.

Examples of Relationships between Updates and Review Content

As already mentioned in the previous subsection, the temporal trends in Figure 2 for topics directly related to game updates already show clear changes over time, with the **Updates and Added Content** and **Changes of Game** topics showing a rising presence in positive (recommending) reviews near major updates, and the topic **Upgrades and Items** showing a growing presence in negative (not recommending) reviews near the updates towards the end of 2018; and several other topics had changes of prevalence associated with times of the updates as discussed above. We next discuss the influence of updates on the reviews in more detail.

The players opinions can be potentially affected by the updates. For example, updates 1.50 (24 July 2018) have added more missions including Real time missions, Scheduled missions, New mission types including freighter combat so on. However, it turns out such content-adding updates aroused complaints. This can be seen in the growing prevalence of the topic **Quest** among players who do not recommend the game. For example, in a review written on 30 July 2018, a user wrote

"I really wanted to like this game, but after 12 hours of game-play and about 10 attempts at the freighter mission. I'm still unable to obtain the freighter. When doing the missions I either recieve wanted level from stray shots at the freighter. Or when I do complete it and recieve the ships transmission to land. I land and then the game hard crashes when I press the button accept the freighter. I spent about 4 hours redoing the mission with the same results on mulitple occasions."

Another example is related to the topic **Survival and Challenge**. Updates 1.10 (26 November 2016) had aroused positive reviews, for example one review written on the next day (27 November 2016) said:

"The new 1.1 update and survival mode add alot of promise to the game, i had uninstalled it but playing survival mode is actually alot more fun than the original game mode !"

And another example review written on the same day said

"New update made this a playable game. If it was too easy before try survival mode, it's pretty brutal."

One example of updates that brought both positive and negative reviews is seen in reviews strongly featuring the topic **Interact with Other Players**, one review written on 29 July 2018 said

"Honestly, this gmae is MIND BLOWING now. Playing it alone or with a friend is so much fun. Give it a chance."

yet another review written on 12 December 2018 said:

"Multiplayer update !! Ok Let's give it a chance, let's buy it. It can't be that bad. Except

that it's not multiplayer. It's observer mode.I was in the middle of a sandstorm walking back toward my ship, while my shields were failing rapidly. My friend was a few steps away, no sandstom, his shields are fine. Can't share anything. This is not multiplayer. Don't call this multiplayer. If this is multiplayer then watching a game on Twitch is multiplayer."

Topic Prevalence and Play Hours

Figure 4 shows the interaction between topic prevalence and play hours; the STM topic model which uses the play hours as a covariate allows us to extract this influence from the model and plot it. The horizontal axis displays how much the topic prevalence of a review increases or decreases when the writer of the review has played one hour; the dot shows the mean increase and the bar shows the 95% confidence interval. Thus, compared to an average review, an increase of one playing hour tends to happen for reviews having around .001 % more content of **Procedual Content Generation of Maps** and .001 % less content of **Worth the Price** and correspondingly for the other topics.

In general, topics related to game-play details such as **Exploration and Discovery**, **Spaceship Travel and Combat**, **Survival and Challenge**, and **Procedurally Generated Universe and Exploration** are positively associated with the play hours. Some other topics such as **Moving and Looking** and **Material Collection**. Although lean to in average higher playing hours, the associations are not significant

When it comes to disappointments, the associations with play hours vary. **Lack of Interest and Reward** leans to higher play hours whereas **Strong Dislike, Waste of Money**, **Quick Disappointment**, and **Disappointment of Promise and Hype** is associated with lower playing hours. Other topics such as **Lack of Content/Grind**, **Lies and Missed Promises**, **Falling Short of Expectations**, and **Feels Unfinished, Indie vs AAA (Quality Level)** do not show either significant positive or negative association with the playing hours.

The variation can be also found in topics related to positive feelings. **Enjoyment Despite Flaws** and **Appreciation** are significantly associated with playing hours, on the other hand, **Appreciating Improvements** leans to lower playing hours.

The topics related to updates also show both directions of association with the playing hours. Topic prevalence of **Upgrades and Items** and **Updates and Added Content** are associated with higher playing hours whereas **Change of Games** and **Gradual Improvements** are negatively associated with the playing hours.

DISCUSSION

One worth mentioning phenomena is that the players' perceptions are indeed changing over time and the change can potentially affected by updates and of the game. There are extracted topics such as **Updates and Added Content**, **Changes of Game** and **Upgrades and Items** that are directly related to the game updates. Besides, the corresponding difference of topic prevalence between players who recommend the game and player who do not recommend the game shows that the overall evaluation can be either positive of negative once the player has experienced the updates.

Another phenomenon is how the playing hours affect the playing experiences that are reflected in the game reviews. The positive association of topics related to game-play details

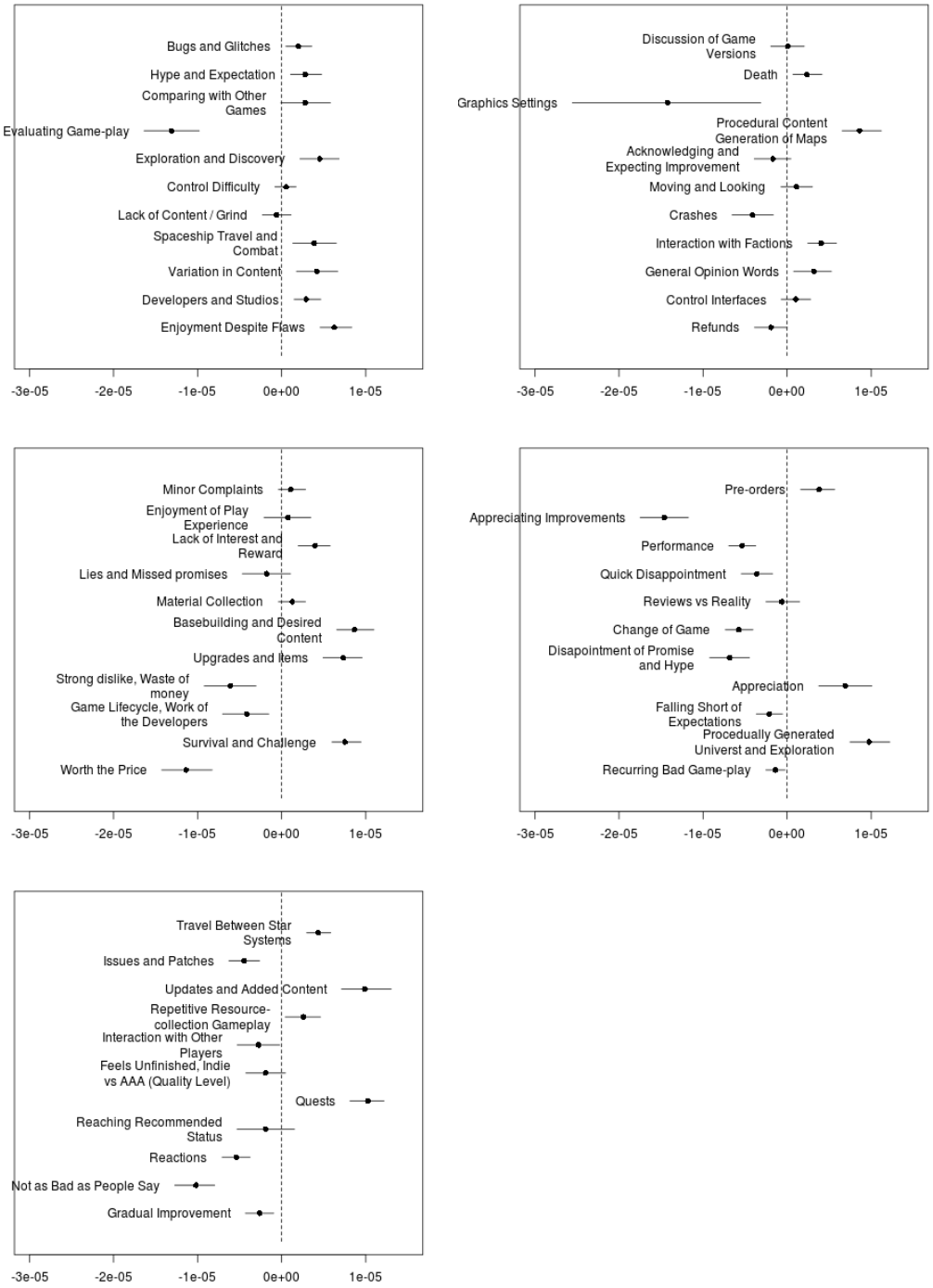

Figure 4: Influence of play hours on topical content of reviews.

with the play hours shows that, compared to other topics, it requires players to spend enough time on playing the game so that the game-play details can become a part of their experience and can be written in the reviews. The association of the topic **Lack of Interest and**

Reward with higher playing hours can potentially reflect the opinions from the players who were not satisfied even after updates. The association of the topics **Strong Dislike, Waste of Money**, **Quick Disappointment**, and **Disappointment of Promise and Hype** with lower playing hours can result from the players who were disappointed in the beginning when the game was launched and chose to complain about the game in their reviews or even gave up to play the game.

The updates have played an important role in the life-cycle of *No Man's Sky* especially after the game was launched. Some of them did potentially affect the opinions of players (e.g. topic **Survival and Challenge**) in a positive way. However, our analysis showed that the updates do not always bring positive feedback. Apparently the players had differnt perceptions to topics **Upgrades and Items**, **Updates and Added Content**, **Change of Games**, and **Change of Gradual Improvements**. Besides, the update related to the topic **Interact with Other Players** especially the multiplayer feature is one example. Some players felt even ´´MIND BLOWING" (see the quote in Section) but some players were not satisfied and left a negative review.

CONCLUSIONS, LIMITATIONS AND OPPORTUNITIES

In this research, an analysis of a large collection of over 85000 game reviews of the game *No Man's Sky* is conducted. The results reveal a large variety of topics that were discussed by players, answering RQ1; the results also reveal clear temporal dynamics of topic prevalence over time, answering RQ2; the results also revealed differences of temporal dynamics between reviews that recommended the game and reviews that did not, answering RQ3; and the results further revealed how such temporal changes coincided with updates to the game, with concrete examples how the updates can potentially affect the discussions, answering RQ4.

The reviews were collected from the Steam platform. Despite the large amount of reviews we were able to gather, reviews from this platform can only reflect the opinions of PC players. Some findings of this research might be applicable directly to players in other platforms such as PlayStation. However, some platform-specific topics, especially technical related topics, might not be appropriate to be imposed on players in other platforms.

This work can be beneficial not only for researchers to study and model players' expectations of game content and reactions to game releases and updates, but also for game industry practitioners when it comes to maintaining players' perceptions; game companies can draw insights from the issues and reactions of players found in this work if the release of another game leads to a similar situation as in *No Man's Sky*.

ACKNOWLEDGMENTS

This work was supported by Academy of Finland decisions 312395, 313748 and 327352.

ENDNOTES

1 https://nomanssky.gamepedia.com/Patch_notes

2 https://scrapy.org/

3 https://www.crummy.com/software/BeautifulSoup/

BIBLIOGRAPHY

Alves, Carina, Geber Ramalho, and Alexandre Damasceno. 2007. "Challenges in requirements engineering for mobile games development: The meantime case study." In *15th IEEE International Requirements Engineering Conference (RE 2007),* 275–280. IEEE.

Ampatzoglou, Apostolos, and Ioannis Stamelos. 2010. "Software engineering research for computer games: A systematic review." *Information and Software Technology* 52 (9): 888–901.

Blei, David M, and John D Lafferty. 2006. "Dynamic topic models." In *Proceedings of the 23rd international conference on Machine learning,* 113–120. ACM.

Blei, David M, Andrew Y Ng, and Michael I Jordan. 2003. "Latent dirichlet allocation." *Journal of machine Learning research* 3 (Jan): 993–1022.

Burnett, Margaret, Curtis Cook, and Gregg Rothermel. 2004. "End-user software engineering." *Communications of the ACM* 47 (9): 53–58.

Chang, Jonathan, Sean Gerrish, Chong Wang, Jordan L Boyd-Graber, and David M Blei. 2009. "Reading tea leaves: How humans interpret topic models." In *Advances in neural information processing systems,* 288–296.

Chen, Ning, Jialiu Lin, Steven CH Hoi, Xiaokui Xiao, and Boshen Zhang. 2014. "AR-miner: mining informative reviews for developers from mobile app marketplace." In *Proceedings of the 36th International Conference on Software Engineering,* 767–778. ACM.

Cho, Yooncheong, Il Im, Roxanne Hiltz, and Jerry Fjermestad. 2002. "An analysis of online customer complaints: implications for web complaint management." In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences,* 2308–2317. IEEE.

Ciurumelea, Adelina, Andreas Schaufelbühl, Sebastiano Panichella, and Harald C Gall. 2017. "Analyzing reviews and code of mobile apps for better release planning." In *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER),* 91–102. IEEE.

Fu, Bin, Jialiu Lin, Lei Li, Christos Faloutsos, Jason Hong, and Norman Sadeh. 2013. "Why people hate your app: Making sense of user feedback in a mobile app store." In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining,* 1276–1284. ACM.

Greer, Des, and Guenther Ruhe. 2004. "Software release planning: an evolutionary and iterative approach." *Information and software technology* 46 (4): 243–253.

Guzman, Emitza, and Walid Maalej. 2014. "How do users like this feature? a fine grained sentiment analysis of app reviews." In *2014 IEEE 22nd international requirements engineering conference (RE),* 153–162. IEEE.

Hello Games. 2016. *No Man's Sky.* [PlayStation 4, Microsoft Windows, Xbox One] Hello Games, Guildford, UK.

Holzer, Adrian, and Jan Ondrus. 2011. "Mobile application market: A developer's perspective." *Telematics and informatics* 28 (1): 22–31.

Kang, Ha-Na, Hye-Ryeon Yong, and Hyun-Seok Hwang. 2017. "A Study of Factors Influencing Helpfulness of Game Reviews: Analyzing STEAM Game Review Data." *Journal of Korea Game Society* 17 (3): 33–44.

Kanode, Christopher M, and Hisham M Haddad. 2009. "Software engineering challenges in game development." In *2009 Sixth International Conference on Information Technology: New Generations,* 260–265. IEEE.

Kasurinen, Jussi, Andrey Maglyas, and Kari Smolander. 2014. "Is requirements engineering useless in game development?" In *International Working Conference on Requirements Engineering: Foundation for Software Quality,* 1–16. Springer.

Ko, Andrew J, Robin Abraham, Laura Beckwith, Alan Blackwell, Margaret Burnett, Martin Erwig, Chris Scaffidi, Joseph Lawrance, Henry Lieberman, Brad Myers, et al. 2011. "The state of the art in end-user software engineering." *ACM Computing Surveys (CSUR)* 43 (3): 21.

Li, Xiaozhou, Zheying Zhang, and Kostas Stefanidis. 2018. "Mobile App Evolution Analysis Based on User Reviews." In *The 17th International Conference on Intelligent Software Methodologies, Tools, and Techniques,* 773–786.

Lim, Chong-U, and D Fox Harrell. 2014. "Developing Social Identity Models of Players from Game Telemetry Data." In *AIIDE.*

Lin, Dayi, Cor-Paul Bezemer, and Ahmed E Hassan. 2017. "Studying the urgent updates of popular games on the steam platform." *Empirical Software Engineering* 22 (4): 2095–2126.

———. 2018. "An empirical study of early access games on the Steam platform." *Empirical Software Engineering* 23 (2): 771–799.

Lin, Dayi, Cor-Paul Bezemer, Ying Zou, and Ahmed E Hassan. 2019. "An empirical study of game reviews on the Steam platform." *Empirical Software Engineering* 24 (1): 170–207.

Lu, Chien, Jaakko Peltonen, and Timo Nummenmaa. 2019. "Game postmortems vs. developer Reddit AMAs: computational analysis of developer communication." In *Proceedings of the 14th International Conference on the Foundations of Digital Games,* 1–7.

Lu, Chien, Jaakko Peltonen, Timo Nummenmaa, Xiaozhou Li, and Zheying Zhang. 2020. "What Makes a Trophy Hunter? An Empirical Analysis of Reddit Discussions." In *International GamiFIN conference.* In press.

Mimno, David, Hanna M Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. "Optimizing semantic coherence in topic models." In *Proceedings of the conference on empirical methods in natural language processing,* 262–272. Association for Computational Linguistics.

Nayebi, Maleknaz, Bram Adams, and Guenther Ruhe. 2016. "Release Practices for Mobile Apps–What do Users and Developers Think?" In *2016 ieee 23rd international conference on software analysis, evolution, and reengineering (saner),* 1:552–562. IEEE.

Newman, James. 2012. "Ports and patches: Digital games as unstable objects." *Convergence* 18 (2): 135–142.

Nummenmaa, Timo, Annakaisa Kultima, Kati Alha, and Tommi Mikkonen. 2013. "Applying Lehman's Laws to Game Evolution." In *Proceedings of the 2013 International Workshop on Principles of Software Evolution,* 11–17. IWPSE 2013. Saint Petersburg, Russia: ACM. ISBN: 978-1-4503-2311-6. `http://doi.acm.org/10.1145/2501543.2501546`.

Palomba, Fabio, Mario Linares-Vasquez, Gabriele Bavota, Rocco Oliveto, Massimiliano Di Penta, Denys Poshyvanyk, and Andrea De Lucia. 2015. "User reviews matter! tracking crowdsourced reviews to support evolution of successful apps." In *2015 IEEE international conference on software maintenance and evolution (ICSME),* 291–300. IEEE.

Roberts, Margaret E, Brandon M Stewart, and Edoardo M Airoldi. 2016. "A model of text for experimentation in the social sciences." *Journal of the American Statistical Association* 111 (515): 988–1003.

Ruhe, Günther, and Moshood Omolade Saliu. 2005. "The science and practice of software release planning." *IEEE Software* 2005:1–10.

Saliu, Omolade, and Guenther Ruhe. 2005. "Supporting software release planning decisions for evolving systems." In *29th Annual IEEE/NASA Software Engineering Workshop,* 14–26. IEEE.

Scalabrino, Simone, Gabriele Bavota, Barbara Russo, Massimiliano Di Penta, and Rocco Oliveto. 2017. "Listening to the crowd for the release planning of mobile apps." *IEEE Transactions on Software Engineering* 45 (1): 68–86.

Sifa, Rafet, Christian Bauckhage, and Anders Drachen. 2014. "The Playtime Principle: Large-scale cross-games interest modeling." In *CIG,* 1–8.

Sifa, Rafet, Anders Drachen, and Christian Bauckhage. 2015. "Large-scale cross-game player behavior analysis on steam." *Borderlands* 2:46–378.

Slivar, Ivan, Mirko Suznjevic, and Lea Skorin-Kapov. 2015. "The impact of video encoding parameters and game type on QoE for cloud gaming: A case study using the steam platform." In *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX),* 1–6. IEEE.

Stamolampros, Panagiotis, Nikolaos Korfiatis, Panos Kourouthanassis, and Efthymia Symitsi. 2019. "Flying to quality: Cultural influences on online reviews." *Journal of Travel Research* 58 (3): 496–511.

Steamworks Documentation: Early Access. Accessed 8 Dec 2019. `https://partner.steamgames.com/doc/store/earlyaccess`.

Svahnberg, Mikael, Tony Gorschek, Robert Feldt, Richard Torkar, Saad Bin Saleem, and Muhammad Usman Shafique. 2010. "A systematic review on strategic release planning models." *Information and software technology* 52 (3): 237–248.

Toivonen, Saara, and Olli Sotamaa. 2010. "Digital distribution of games: the players' perspective." In *Proceedings of the International Academic Conference on the Future of Game Design and Technology,* 199–206. ACM.

Villarroel, Lorenzo, Gabriele Bavota, Barbara Russo, Rocco Oliveto, and Massimiliano Di Penta. 2016. "Release planning of mobile apps based on user reviews." In *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE),* 14–24. IEEE.

Wu, Yingcai, Furu Wei, Shixia Liu, Norman Au, Weiwei Cui, Hong Zhou, and Huamin Qu. 2010. "OpinionSeer: interactive visualization of hotel customer feedback." *IEEE transactions on visualization and computer graphics* 16 (6): 1109–1118.