

An Empirical Study of Trends of Popular Virtual Reality Games and Their Complaints

Rain Epp, Dayi Lin, and Cor-Paul Bezemer

Abstract—The market for virtual reality (VR) games is growing rapidly, and is expected to grow from \$3.3B in 2018 to \$13.7B in 2022. Due to the immersive nature of such games and the use of VR headsets, players may have complaints about VR games which are distinct from those about traditional computer games, and an understanding of those complaints could enable developers to better take advantage of the growing VR market. We conduct an empirical study of 750 popular VR games and 17,635 user reviews on Steam in order to understand trends in VR games and their complaints. We find that the VR games market is maturing. Fewer VR games are released each month but their quality appears to be improving over time. Most games support multiple headsets and play areas, and support for smaller-scale play areas is increasing. Complaints of cybersickness are rare and declining, indicating that players are generally more concerned with other issues. Recently, complaints about game-specific issues have become the most frequent type of complaint, and VR game developers can now focus on these issues and worry less about VR-comfort issues such as cybersickness.

Index Terms—Virtual Reality games, gamer complaints

I. INTRODUCTION

The releases in 2016 of consumer virtual reality (VR) headsets such as the HTC Vive and Oculus Rift has afforded developers a large and growing audience for VR computer games. The VR market is rapidly growing in size: the install base of VR is expected to grow from 14 million in 2018 to 51 million in 2022, and the VR game market is expected to grow from \$3.3B to \$13.7B in the same period [45].

The unique display and input hardware used in VR games offers new and different experiences and means of interaction for players of VR games. When comparing the VR and non-VR modes of individual games, VR has been shown to increase immersion and enjoyment, but also may be associated with worse performance in completing in-game tasks and a perceived loss of control [27, 37, 46]. Moreover, use of VR in games may lead to cybersickness [39], an issue typically absent from non-VR games. Reflecting these differences, players may have complaints about VR games which are distinct from complaints about ‘traditional’ computer games. An awareness of such complaints would enable developers to take greater advantage of the growing VR market as it continues to evolve.

While many aspects of VR computer games have been studied (e.g., their medical [15, 26] or educational [7, 48] potential, and comparisons with non-VR games [27, 37, 46]),

to the best of our knowledge our study is the first large-scale examination of user complaints of such games.

In this paper, we conduct a large-scale study of 750 VR games and 17,635 reviews on the Steam platform, one of the most popular distribution platforms for computer games. Our study provides a broad view of complaints among players of VR games. We first conduct a preliminary analysis of the metadata of VR games on Steam. We then examine the negative reviews of those games and classify them into categories to identify common complaints. We focus on the following two research questions.

RQ1: What are the characteristics of VR games on Steam? The games have a median of 3 updates each, and a median price of \$10.99, though this price has increased over time. Cross-platform support for multiple VR headsets is increasing, although support for Windows Mixed Reality headsets lags behind. Support for smaller-scale play areas is getting increasingly common.

RQ2: What do players of VR games complain about? We identify seven complaint categories. The most common complaint is that games are overpriced or lack content; however, the total proportion of such complaints is rapidly decreasing over time. Complaints reflecting issues that are most related to VR-comfort, such as cybersickness and VR control schemes, are rare. Recently, game-specific complaints have become the most common type of complaint.

This paper is organized as follows. Section II provides background information on VR technology and Steam. Section III presents the methodology of our study. Sections IV and V present the results of our investigation into VR game characteristics and user complaints respectively. Section VI discusses the implications of our findings. Section VII discusses related work. Section VIII discusses threats to the validity of our study. Finally, Section IX concludes the paper.

II. BACKGROUND

In this section we provide a brief background on the VR technology used by the games we study. We also describe aspects of the Steam Store and Community which are relevant to our data mining processes.

A. VR technology for games

The main device needed to play VR games is a VR headset, which creates the illusion of a physical world by showing each eye a separate render of the virtual game world. Orientation and position tracking are used to update the virtual world

• R. Epp and C. Bezemer are with the Analytics of Software, Games and Repository Data (ASGAARD) lab, University of Alberta, Canada. E-mail: {mdepp, bezemer}@ualberta.ca

• D. Lin is with the Center for Software Excellence, Huawei, ON Canada. E-mail: dayi.lin@huawei.com

TABLE I
AN OVERVIEW OF THE FOUR MOST POPULAR VR HEADSETS USED TO PLAY STEAM GAMES

Name	Tracking	Controllers	Controller Tracking
Oculus Rift	External sensors (outside-in)	Oculus Touch	Same as headset
HTC Vive	External markers (inside-out)	Vive controllers	Same as headset
Windows Mixed Reality	Markerless inside-out	WMR controllers	Sensors on headset
HTC Vive Pro	External markers (inside-out)	Vive controllers	Same as headset

based on the user's movements. According to the June 2019 Steam survey [47], the most popular VR headsets used with Steam games are the Oculus Rift, HTC Vive, Windows Mixed Reality¹, and HTC Vive Pro. The capabilities of these headsets are summarized in Table I.

Several headsets use external apparatus to aid position and orientation tracking, such as external sensors for the Oculus Rift and external markers for the HTC Vive and HTC Vive Pro. However, some headsets, especially those released later, implement tracking using only cameras and sensors on the VR headset itself (markerless inside-out tracking). All Windows Mixed Reality headsets use this method of tracking.

VR headsets can be roughly grouped into three categories. PC-based headsets require a tethered connection to an external desktop or laptop computer in order to offload processing and rendering from the headset. Smartphone-based headsets instead require a connection to a smartphone that is used for processing and display. Finally, standalone headsets require no connection to any external device in order to function. It is worth noting that various means exist to use a wireless connection instead of a tether for PC-based headsets, such as the Vive Wireless Adapter for the Vive and Vive Pro. Also, some standalone headsets (such as the Oculus Quest) can be connected through a wire to a PC as well.

All four of the headsets in Table I are PC-based. However, playing Steam games on other types of headsets is also possible, and various projects exist for this purpose, such as ALVR² which acts as a remote display for the smartphone-based Samsung Gear VR and the standalone Oculus Go and Oculus Quest headsets.

While some VR games can be controlled using standard input devices such as keyboards, mice, and game controllers, VR gaming systems typically have dedicated motion controllers. In most cases two controllers are used, one held in or attached to each hand. Like the headset, their position and orientation is tracked to provide a more natural and intuitive input than traditional control schemes. These controllers may be tracked using the same external markers or sensors as the headset, or by sensors on the headset itself. Controller support is generally limited to a single headset or series of headsets, but exceptions exist (e.g., 3rd-party software exists to use the Vive controllers with Windows Mixed Reality headsets).

Although game engines such as Unity and Unreal attempt to abstract away the differences in controller and headset functionality, such differences still have implications for game

developers. For example, using the headset to track controllers (as Windows Mixed Reality systems do) can lead to less reliable controller tracking than other methods, potentially breaking games which need 360° tracking of controllers.

B. VR games on Steam

Steam is a digital distribution platform maintained by Valve Corporation. It is the largest digital distribution platforms for PC games, containing over 30,000 games³ and over 90 million active users monthly⁴.

Steam is also a major source of VR games, containing over 3,000 games which support VR in some capacity. By contrast, the official Oculus Rift store "Rift Experiences" had 792 games, and the official HTC Vive store "Viveport" had 986 games, as of 2019-05-13. Valve provides a set of tools and services for users and developers of VR games known collectively as SteamVR⁵. Applications can interface with SteamVR using the OpenVR SDK. However, to publish a VR game on Steam, it is not mandatory to use SteamVR. For example, a game might use the Oculus SDK to support the Oculus Rift and the OpenVR SDK to support other headsets.

VR games on Steam are searchable through two separate filters: *VR Only* and *VR Supported*. The former describes games which require VR hardware in order to play; in the latter, use of VR is supported but optional. *VR Only* is almost entirely a subset of *VR Supported* (we observed during our data analysis that the four exceptions are likely an error by Valve or the games' developers).

For each game, developers can indicate officially supported categories of headsets and controllers as seen by users. Headset support is indicated to users as any combination of *Oculus Rift* (e.g., Oculus Rift, Oculus Rift S), *HTC Vive* (e.g., HTC Vive, HTC Vive Pro), and *Windows Mixed Reality* (all Windows Mixed Reality immersive headsets)⁶; controller support by *Gamepad*, *Keyboard/Mouse*, and *Tracked Motion Controllers*. While in theory all SteamVR games are compatible with all headsets among the possible categories (i.e., they will run), as noted previously there are still differences in practice, which make testing on individual devices important. In addition to controllers and headsets, developers can specify supported play areas for their games, indicated to users as *Seated*, *Standing*, and *Room-Scale*. Valve defines a *Standing* play area

³<https://store.steampowered.com/search/?category1=998>

⁴<https://partner.steamgames.com/>

⁵<https://www.steamvr.com/en/>

⁶There are also labels for the Oculus Rift developer kits, but these are rarely used, hence we ignore them. Additionally, after our data was collected Steam added a label for the Valve Index headset.

¹Windows Mixed Reality is not actually a discrete headset but instead refers to any VR headset following Microsoft's Windows Mixed Reality standard, known as Windows Mixed Reality immersive headsets.

²<https://github.com/polygraphene/ALVR>

as a 1-meter diameter circle and a *Room-Scale* play area as a rectangle of at least 2×1.5 meters.⁷

C. Steam Community

Once a user has recorded playtime⁸ for a game on Steam, they are able to post a user review for that game to Steam Community. Instead of a star rating system, Steam reviews are designated either *Recommended* or *Not Recommended*, corresponding with positive and negative ratings respectively. For each game, these ratings are summarized on the store page in categories (e.g., *Mostly Positive* or *Mixed*), over all reviews and over only reviews posted within the last 30 days.

Through Steam Community, developers and journalists are able to post news updates, such as promotions or update announcements, on a per-game basis. These news posts are organized into channels denoting their origin or intended purpose, such as *PC Gamer* or *Product Releases*.

III. METHODOLOGY

In this section we describe the methodology of our empirical study. Figure 1 gives a visual overview of our data collection and filtering processes. In the remainder of this section, we describe the processes in greater detail.

A. Collecting basic game data

We adapt a customized crawler from our prior work [23] to extract information for all 31,177 games on Steam as of April 4, 2019. For each game, the crawler gathers the title, release date, publisher, developer(s), price (undiscounted, in Canadian dollars), supported headsets, supported input devices, supported playing areas, VR support (the *VR Supported*, *VR Only* properties), and number of reviews. Vague or non-specific release dates (e.g. January 2018) are ignored; only release dates with daily precision are gathered.⁹

We select the games for our study according to the following criteria:

- **VR support:** The game must be a VR game.
- **Popularity:** The game must have 25 or more reviews.

We wish to investigate which complaints players have with VR games. This makes gathering reviews for games with optional VR support somewhat problematic, since we do not know whether a review describes the experience with or without VR. Unfortunately, Steam provides no mechanism to determine if a review is of the VR mode or non-VR mode of a game (or both). Therefore, we only consider games marked *VR Only* as “VR games” in this study.

There are a few instances in the paper in which it is helpful to consider the full set of games (including the non-VR and less popular ones); these cases are noted specifically when they occur.

For each studied game, we also collected all Steam Community news posts. As described in Section IV, we use these news

TABLE II
A BRIEF SUMMARY OF THE DATA COLLECTED FROM THE STEAM STORE
AND STEAM COMMUNITY

# of collected games in total	31,117
# of VR games	2,710
# of VR games with ≥ 25 reviews	750
# of collected reviews	150,887
# of unique review authors	74,899
# of English reviews	117,012
# of English, negative reviews	17,635
# of collected news posts	5,436

posts to estimate the number of updates for each VR game. Additionally, we created a crawler to collect historical pricing information of Steam games from the website *SteamDB*¹⁰. We ran this crawler for all non-free, popular VR games to capture how their prices changed over time.

B. Collecting reviews of VR games

We use another custom crawler to collect all 150,887 user reviews of the studied VR games. This crawler was executed directly after the basic characteristics crawler and also performed its data collection on April 4, 2019. For each user review, we gather the title of the game under review, the review’s content, the rating (*Recommended* or *Not Recommended*), and whether the review was an early access review.

C. Dataset description

Table II shows a brief description of the collected data. As noted in prior sections, basic characteristics were crawled for every game on Steam, but reviews and news posts, and historical prices were only collected for VR games with 25 or more reviews.

IV. RQ1: WHAT ARE THE CHARACTERISTICS OF VR GAMES ON STEAM?

Motivation: Before analyzing user reviews, we examine the basic characteristics of VR games. As our paper is the first large-scale examination of VR games on Steam, it is important to understand the dataset. We examine trends within the data to understand the evolution of VR games over time. The findings in this section contextualize the remainder of the study.

Approach: We examined the release dates, developers, prices, and updates of the collected VR games. We also examined the level of support over time for the categories of headsets, controllers, and play areas.

To estimate the number of updates for each VR game, we used a method from our prior work which extracts release notes from collected Steam Community news posts [21]. This method was evaluated in that work to have a precision of 89% and a recall of 87% for early access games. Only news posts from the *Product Updates*, *Product Releases*, *Client Updates*, and *Steam Community Announcements* channels were considered. A news post from these channels was considered to contain release notes if:

⁷<https://steamcommunity.com/app/358720/discussions/0/350532536103514259/>

⁸<https://partner.steamgames.com/doc/store/reviews>

⁹Only the *release date* is ignored. The rest of the game data is still crawled and processed as usual.

¹⁰<https://steamdb.info>

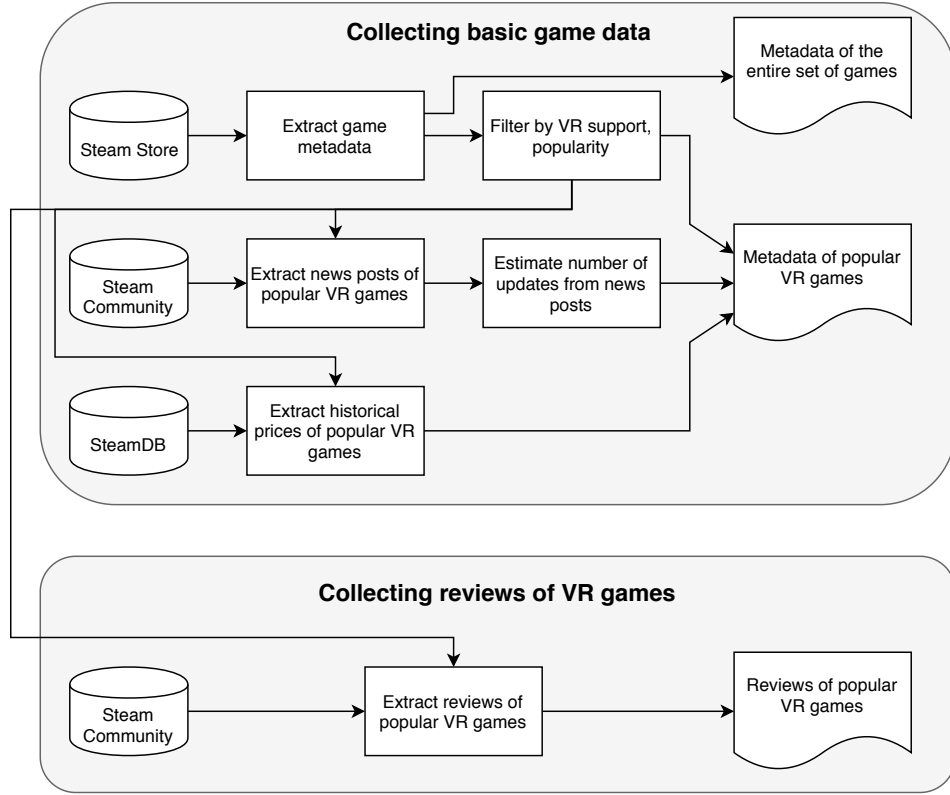


Fig. 1. An overview of the methodology of our study.

- it was published on the *Product Updates* or *Product Releases* channel; or
- its title contained the word *update*, *release*, *patch*, *hotfix*, or *change log*, ignoring case, and using the case-insensitive regex patterns `hot.*fix` and `change.*log` for *hotfix* and *change log* respectively; or
- its title contained a version number, using one of the following case-insensitive regex patterns:
 - `(version|alpha|beta|gamma|build) ?[0-9]+`, or
 - `\bv\.\.? ?[0-9][0-9.a-z]*\b`.

The number of updates per game was then considered to be the number of news posts for that game which contained release notes.

To compare the price and update distributions, we used the Wilcoxon rank-sum test, a non-parametric test which can be used to compare two independent distributions. If $p < 0.05$ we consider the distributions to be significantly different. To quantify this difference, we use Cliff's delta (d) [24], with the thresholds provided by Romano et al. [40]:

$$\text{Effect size is } \begin{cases} \text{negligible} & \text{if } |d| \in [0, 0.147] \\ \text{small} & \text{if } |d| \in (0.147, 0.33] \\ \text{medium} & \text{if } |d| \in (0.33, 0.474] \\ \text{large} & \text{if } |d| \in (0.474, 1] \end{cases}$$

Early access games were identified as in our prior work [21], using the presence of either an early access banner or an early access review.

We estimated the release price of the paid games by combining the SteamDB historical pricing data with the game's release date from Steam. If SteamDB recorded the price before or at the release date, we used the last recorded price as the release price. Otherwise, we estimate the release price using the earliest recorded price for the game, provided it was recorded no more than two days after the game's release date.

Results: The rate at which VR games are released is declining. Figure 2 shows the number of games released per week from 2016 to 2019. Throughout the year 2018, 158 of the studied games were released, corresponding with 3.04 releases per week. The release rate of VR games peaked at the beginning of 2017 and thereafter declined. However, since only games with 25 or more reviews were considered, a slight decline is to be expected nearer the present, as recent games have had less time to accumulate reviews. For this reason Figure 2 also shows the number of games per week for *all* VR games, including those with few reviews. When considering all VR games, the release rate peaked mid 2017, but again declined after that time. The initial spike coincides with the release of the HTC Vive in June 2016. The next spike, at the end of 2016, occurs in the first three-quarters of December in that year and is likely a result of developers rushing to release their games by Christmas 2016.

The median price of a VR game was \$10.99, and \$14.24 when considering paid VR games only. Paid VR games were more expensive than paid non-VR games. Figure 3 shows the price distribution of the games. The median price of the studied games was \$10.99. 184 of 750 (25%) of the studied

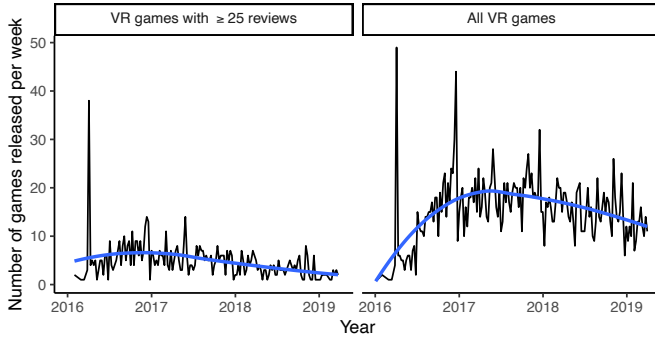


Fig. 2. Release dates of Steam VR games. The plot on the left shows the release dates of the VR games with ≥ 25 reviews; the plot on the right shows all VR games, including those with fewer reviews. Games with release dates in the future are excluded, as are 6 games with release dates before 2016. Additionally, 151 games with no release date given are not shown (including 1 with 25 or more reviews). The trend line is a LOESS trend line which provides a smoothed estimate of the rate of releases at each point in time.

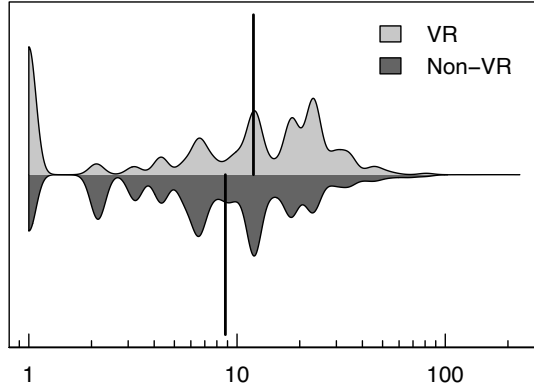


Fig. 3. Price distribution of the studied VR and non-VR games. The x-axis is $1 +$ the price of the game, in Canadian dollars. The black vertical lines indicate the median values of each distribution.

games were free or free to play. Excluding these games, the median price (of non-free games) was \$14.24. The two most costly games, with a price of \$80.00, were *Fallout 4 VR*¹¹ and *The Elder Scrolls V: Skyrim VR*¹², both developed by Bethesda Game Studios. The median price of a non-VR game (with 25 or more reviews) was \$7.79 and \$8.99 when considering paid games only. A Wilcoxon rank-sum test showed that the price distribution of the VR games was significantly different than that of (popular) non-VR games, both in the overall and paid case. The effect size was negligible ($|d| = 0.06$) overall and small ($|d| = 0.30$) for paid games.

The release price of paid VR games has increased over time. Figure 4 shows the trend of the release price of the paid studied VR games over time. This is the undiscounted price of each game at its release. The price increases for two years until just after 2018, when it levels off and even decreases slightly. Within the final six months, the median price for a paid VR game was \$22.14, compared with \$10.99 within the first six months.

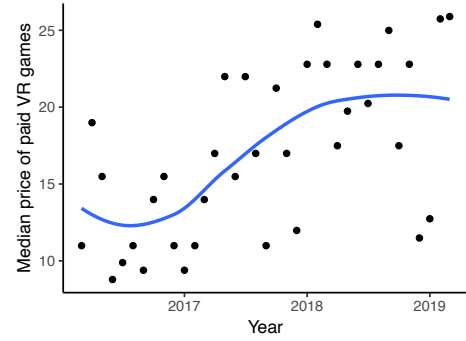


Fig. 4. Trends for the release prices of the paid VR games. The *Year* axis is grouped into months. Each data point is the median price of paid VR games released in that month. The line is a LOESS trend line, indicating a smoothed estimate of the median price of paid games over time. Games with release dates before 2016 or in the future are excluded, as are 22 games with no estimated release price from SteamDB.

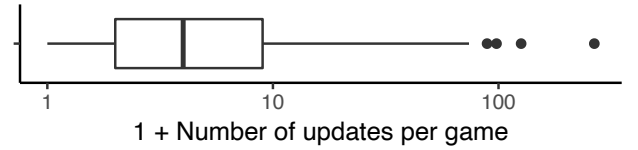


Fig. 5. Distribution of the number of updates to VR games.

The studied VR games received a median of 3 updates. Figure 5 shows the distribution of the number of updates per studied VR game. Of these games, 176 of 750 (23%) received no updates at all. In addition, 439 of 574 games with updates (76%) received only 10 or fewer updates. Early access games received a median of 8 updates, while non-early access games received only 2 (this difference was statistically significant as determined with a Wilcoxon rank-sum test, with $|d| = 0.61$ indicating a large effect size). These figures suggest that there is a low level of ongoing developer support for many of the studied games.

396 of 715 VR developers (55%) published only a single game on Steam, and 77% exclusively published VR games. These numbers take into account all games (including non-VR or less popular ones) created by the VR developers.¹³ Of the developers with multiple games, 153 of 319 (48%) have exclusively created VR games. In addition, **among the developers of the studied games with both VR and non-VR games on Steam, 72% published the non-VR game first.** Again, these numbers take into account all games by the VR developers, including non-VR or less popular ones. A total of 166 developers published both VR and non-VR games on Steam. 156 of these developers had at least one game in each category (VR or non-VR) with a well-defined release date. Among these 156 developers, 113 (72%) released their first non-VR game before their first VR game. Therefore while only a small proportion of VR developers created both VR and

¹³By *VR developer* we refer to developers which created at least one of the 750 studied VR games. Developers of *only* less-popular and/or non-VR games are not considered, since we wish to characterize the developers of the studied games only. However, *all* games created by the VR developers are taken into account, in order to properly characterize their output.

¹¹<https://store.steampowered.com/app/611660/>

¹²<https://store.steampowered.com/app/611670/>

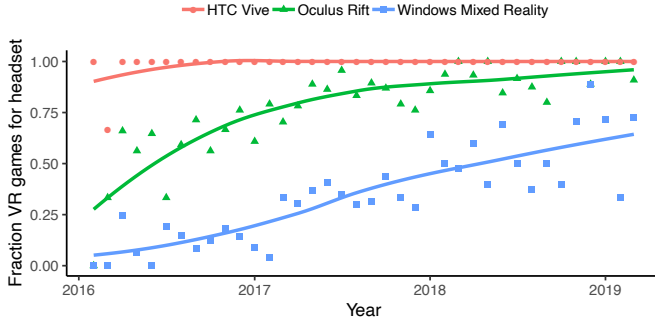


Fig. 6. Trends for headset support among the studied games. The *Year* axis is grouped into months. Each data point is the fraction of games released in the corresponding month which support a VR headset as of 2019-04-04. The lines are LOESS trend lines, indicating for each headset a smoothed estimate of the fraction of VR games released supporting that headset.

non-VR games, most developers among that group had prior experience making non-VR games before publishing their first virtual reality game on Steam.

In addition, we manually examined a random sample of 50 of these developers and determined that only 6 of their collective 199 games (3.0%) are sequels, spinoffs, or remakes of existing non-VR titles that were released by those developers on Steam. This indicates that, generally speaking, the studied virtual reality games are new material likely created specifically for VR instead of adaptations of existing games.

The HTC Vive is the most widely supported headset on Steam with more than 99% support. Among the studied games, 749 of 750 support the HTC Vive. In contrast, only 77% and 31% of the studied games support the Oculus Rift and Windows Mixed Reality categories, respectively. The high support for the HTC Vive is likely related to Valve’s close association with the HTC Vive headsets, which were developed by a co-operation between HTC and Valve.

Support for Oculus Rift and Windows Mixed Reality headsets is increasing. Figure 6 shows the fraction of games released each month with official support for each VR headset category. Support for Oculus Rift and Windows Mixed Reality headsets grows throughout the entire timeline. Within the final 6 months, support for Oculus Rift headsets has reached 97%, indicating it has essentially caught up with the Vive. In that time period, Windows Mixed Reality support still lags behind at 67%, although it continues to increase at least up until the time of data collection.

SteamVR gained support for Windows Mixed Reality headsets in November 15, 2017 [25], a month after the first launch of Windows Mixed Reality headsets. At the time, indicated support for Windows Mixed Reality headsets was disabled by default, requiring developers to manually specify that their games were compatible with Windows Mixed Reality. Therefore it is likely that the 116 games with support for these headsets prior to November 15, 2017 have had support added after their release.

712 of 750 (95%) games support tracked motion controllers, and support for other control schemes is decreasing. Figure 7 shows the proportion of games released

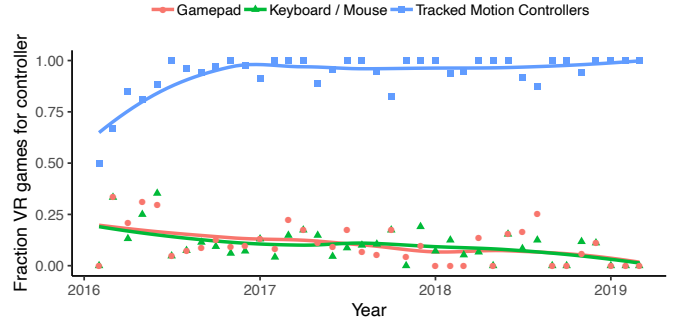


Fig. 7. Proportion of games released each month with support for the three controller categories. This graph is constructed identically to Figure 6, except it displays controllers instead of headsets.

with support for each category of controller. 98% of the studied games released in the final six months support tracked motion controllers, and only 3 of 58 games (5.2%) released in that time period support other control schemes (i.e. keyboard/mouse or gamepad). By contrast, only 87% of the VR games that were released within the *first* 6 months of 2016 had support for tracked motion controllers, and a full 25% supported other control schemes. A possible explanation could be the lack of a dedicated controller for the Oculus Rift at launch. Not only has support for dedicated VR controllers increased to near-ubiquity, but recent releases have largely stopped supporting traditional control schemes at all.

The most common play areas are room-scale and standing, supported by 79% and 76% of games respectively. Support for the seated play area is increasing. As seen in Figure 9, the proportion of games released with support for the “room-scale” and “standing” play areas has remained approximately constant since the year 2017. However, support for the “seated” play area has increased throughout the timeline from 40% of games released in the first six months to 59% in the final six months. Within the final six months the support for room-scale and standing play areas is 78% and 85% respectively, indicating that support for “standing” has also increased. The increasing support for smaller-scale play areas is likely an attempt by developers to make their games appeal to a wider audience, since the larger-scale VR setups require greater user commitment in terms of equipment and physical space, and thus may be unavailable or unappealing for some users. In addition, a smaller required play area would make the game more portable. As standalone VR headsets grow in popularity, this would allow the game to be played while travelling, such as on an airplane.

V. RQ2: WHAT DO PLAYERS OF VR GAMES COMPLAIN ABOUT?

Motivation: User reviews are a visible indication of the user-perceived quality or enjoyment of a game. It is therefore important to developers to obtain positive reviews for their games in order to maximize their chances of success. Reviews also provide valuable feedback for developers on which aspects of their games have been positively received and which need

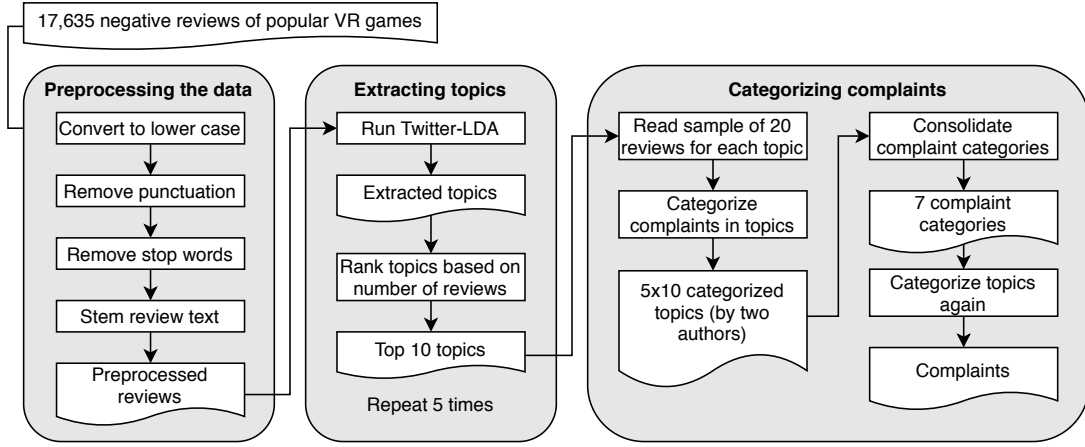


Fig. 8. An overview of our approach for extracting complaints from the negative reviews.

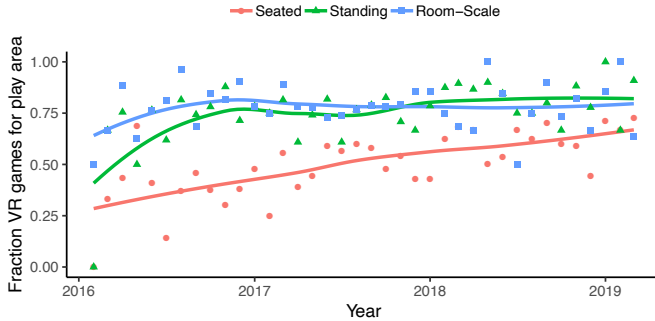


Fig. 9. Proportion of games released each month with support for the three play area categories. This graph is constructed identically to Figure 6, except it displays play areas instead of headsets.

improvement. Although both positive and negative reviews can provide useful feedback, our prior work shows that negative reviews do so more often [23]. Here we analyse negative user reviews to extract the most commonly-occurring complaints for virtual reality games. Such an overall understanding of users’ complaints of VR games is useful for developers and could enable them to tailor their development and testing efforts to most effectively address players’ concerns.

Approach: Figure 8 give an overview of our approach for extracting complaints from reviews. We detail each step below.

Preprocessing the data: All 17,635 negative, English reviews for the studied VR games were automatically classified using the Twitter-LDA [51] algorithm. Twitter-LDA is a variation of LDA that is designed to assign topics to Twitter “tweets”, and as such is capable of generating meaningful topics from smaller amounts of text than traditional LDA. One topic was assigned to each document (review). Prior to running the algorithm, each review was preprocessed as follows:

- 1) The review text was converted to lowercase and punctuation was removed.
- 2) Using the Python package `nltk` [3], English stopwords were removed from the review text.
- 3) The review text was stemmed using the Porter stemmer, as implemented by `nltk`.

Extracting topics: After running Twitter-LDA on the reviews, the generated topics were ranked according to the number of reviews assigned to each topic. Twitter-LDA is nondeterministic, and its results will vary by run. Therefore we ran the algorithm five times and combined the results by grouping the top 10 topics of each run (by number of reviews) into manually-classified complaint categories.

Categorizing complaints: Each topic was assigned one complaint category as follows:

- 1) For each run, the first and third authors were each given a random sample of 20 reviews for each of the top 10 topics of that run. Each author was therefore assigned a total of 1000 reviews over 50 topics. Using these reviews, both authors independently identified categories for each topic. Each topic was assigned exactly one category.
- 2) The categories identified by each author were consolidated into a new set of 7 complaint categories agreed upon by both authors. These complaint categories are shown in Table III.
- 3) The topics were classified by the first and third authors once more (using the same set of reviews as in step 1), this time into the predefined complaint categories. The authors disagreed on 3 of the 50 classified topics, with Cohen’s kappa $\kappa = 0.929$, indicating a strong agreement between categorizations. The remaining differences were resolved through discussion.

After this process was complete, each review had been automatically classified into a topic once in each run, and each of the top 10 topics for each run had been grouped into a complaint category. We therefore considered a review to belong to a complaint category if it was assigned a topic which was grouped into that category. For each run, we ranked the complaint categories by number of reviews, from 1 to 7 for most to least reviews. For each complaint category, we calculated the median rank across runs to summarize its relative occurrence among the reviews.

Results: Table III shows the median rank of each category across all runs. Categories with rank 1 had the most reviews within a run; categories with rank 7 had the least. Figure 10

TABLE III

MANUALLY IDENTIFIED COMPLAINT CATEGORIES FOR NEGATIVE REVIEWS OF STEAM VR GAMES. THE EXAMPLE FOR EACH COMPLAINT CATEGORY IS A FRAGMENT OF TEXT FROM A REVIEW WITH THAT CATEGORY (THE REVIEW TEXT HAS BEEN MINIMALLY EDITED FOR CLARITY).

Complaint category	Description and Example	Median rank
lacks content	The game is too short, lacks content, or is overpriced. <i>“it’s good but not worth the money”</i>	1
game-specific	A significant number of reviews in the topic are for a single game only, and contain complaints about game-specific information such as gameplay. <i>“Skyrim is cool and all, but ... it’s basically ‘Crouch-Simulator with dragons’ ”</i>	3
community	The game has been abandoned by either players or developers, or has multiplayer issues. <i>“I couldn’t find anyone online to play with”</i>	4
crashes	The game crashes, freezes, cannot start, or otherwise has major bugs rendering it unplayable. <i>“it crashes within minutes of starting”</i>	4
nausea	The game causes nausea or headaches. <i>“made me ill to the point I have to lie down”</i>	5
controls	The game has various issues with controls (locomotion and/or manipulation) or with headset. <i>“can’t correctly use the oculus touch controllers”</i>	6
optimization	The game has optimization issues such as stutters or a low framerate. <i>“starts out smooth then gets progressively choppy”</i>	7

shows the relative frequency of each complaint category over time. **The most common complaint was that games lack content or are overpriced. The frequency of this complaint is decreasing.** These complaints correspond with the `lacks content` complaint category. As can be seen in Figure 10, the relative frequency of this complaint category has decreased over time from 2016 to 2019. Interestingly, while one might expect early reviews to be more tolerant of such concerns given the early nature of the technology, this does not appear to be the case. The decreasing frequency of this complaint suggests that the VR games market has become more mature over time, with developers as a whole moving away from smaller, “tech demo”-like experiences typical of early efforts. Another explanation could be that gamers are now used to VR games being smaller than non-VR games and hence compare newer VR games to older VR games, instead of to non-VR games. It is interesting to note that the release price of VR games has increased over time (see Figure 4) even as the frequency of `lacks content` complaints have decreased. This suggests that as time goes on, players of VR games have felt they are getting better value for their money.

A major complaint was with lack of developer or player community. The corresponding complaint, `community`, had median rank 4 and is thus tied for second most-occurring (aside from the `game-specific` complaint category). The relative frequency of this complaint has largely stayed constant over time. This suggests that abandoned games (both by the developer and the players) have been and continue to be an issue for players of VR games.

Crashes, freezes, and game-breaking bugs were complained about less frequently than game design- and business-oriented issues. The category `crashes`, corresponding with such bugs, had median rank 4. The game design- and business-oriented `lacks content` and `community` complaint categories, with ranks 1 and 4 respectively, collectively occurred more frequently than complaints about bugs and crashes. This observation is consistent with our prior work [23], which was not specific to VR games and which showed that, especially among negative reviews, complaints about game design are more common than complaints about bugs (including crashes).

The relative frequency of game-specific complaints jumped in early 2018. These complaints (belonging to the `game-specific` category), are from topics with a large percentage of reviews that complaint about a single game. This increase and another in early 2019 was large enough to make these complaints the second-most-frequently occurring. This could indicate a shift to complaints about specific game content instead of other issues. As with the decline in the `lacks content` complaint category, this shift likely reflects the growing maturity of the VR games market.

The complaints reflecting issues that are most unique to VR occurred the least often. The complaint categories `nausea`, `controls`, and `optimization` had the lowest median ranks assigned, at 5, 6, and 7 respectively. These categories correspond most closely with issues specific to VR-comfort: `nausea` since it describes symptoms of cybersickness; `controls` due to the unique control schemes in VR; and `optimization` since VR games require extra processing power to render scenes to each eye, and since issues with unoptimized games (e.g. stuttering or reduced framerate) can induce cybersickness. While the absence of these issues does not guarantee user enjoyment, such issues can quickly render a game unplayable. The low rank assigned to the categories suggests that these issues were considered by users to be less important than the others which are less intrinsic to VR platforms, suggesting that developers and hardware designers have largely succeeded in reducing VR-comfort issues to an acceptable level for computer games.

VI. IMPLICATIONS

The VR games market is maturing. Complaints that games are overpriced for their content have decreased sharply, even as the price of VR games has risen over time. The median price of a paid VR game has risen from \$10.99 in the first six months of 2016 to \$22.78 in the latest six months of data collection, levelling off after 2018. Despite this increase, complaints about games which are overpriced or lack content have decreased in frequency from roughly 50% to 20% of all complaints. This decrease suggests that the quality of new games has generally increased, since players have become more satisfied with the amount of content in newer games even

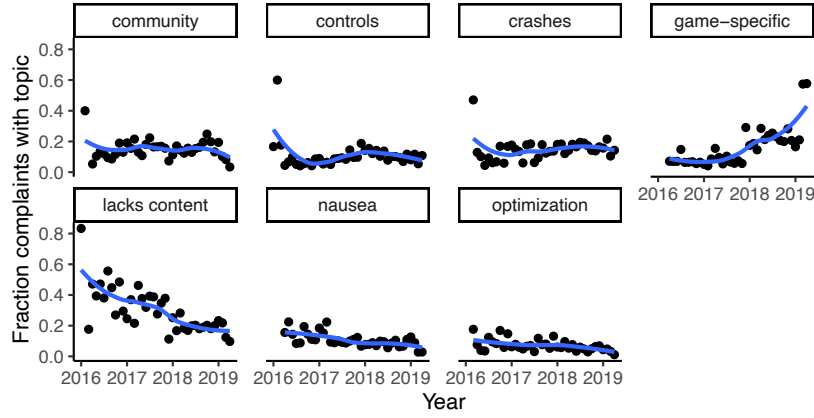


Fig. 10. The fraction of categorized negative reviews belonging to each category over time. The *Year* axis is grouped into months. Each point is the fraction of categorized complaints published in that month belonging to the given category. This fraction is calculated using the combined topics and categories of all five runs. The lines are LOESS trend lines, providing a smoothed estimate of the Fraction of categorized complaints with each topic over time.

as they are more expensive. Hence, while less VR games are released each month, their quality seems to be improving.

Developers should be prepared to deliver cross-platform experiences which support multiple play areas. The majority of recent releases have support for multiple headset categories and play areas, likely creating an expectation of such support among players. Within the final six months of our study period, 97% of the released games supported both *Oculus Rift* and *HTC Vive* headsets, with *Windows Mixed Reality* support at 67% and increasing. Similarly, within this time period 85% of the releases supported *Standing* and 78% of the releases supported *Room-Scale* play areas. *Seated* play areas have seen increasing support, reaching 59% overall within the final six months and climbing steadily, a trend developers should keep in mind when creating their VR games.

It is also worth noting that nearly all recent releases have supported tracked motion controllers: 98% of games released in the final six months have support for tracked motion controllers, and only 5.2% supported other control schemes. Therefore developers should not worry about including keyboard or gamepad support in their games.

Cybersickness is not a major issue among players of current VR games. The frequency of complaints reflecting cybersickness has decreased steadily over time, likely reflecting efforts of developers and hardware makers to mitigate the issue. However, cybersickness has never comprised a large proportion of the complaints. It was initially dwarfed by complaints about price and content, and recently by complaints about game-specific issues. Overall, it is ranked the 5th-most common complaint category. While Rangelova et al. concluded from an online survey that cybersickness is widespread among gamers [38], our results suggest that cybersickness is not a major issue with players compared with other concerns.

Developers should focus on delivering high-quality gaming experiences rather than further improving VR-comfort aspects of their games. Since approximately 2018, the most common complaint among user reviews has been with game-specific issues. Complaints of cybersickness, controls, and other elements reflecting VR-comfort issues have low and

declining frequency. In order to address the largest number of complaints, developers should focus on further improving the value of their VR experiences as games rather than addressing VR-comfort issues. Since the frequency of complaints about lack of content and overpriced games has decreased sharply over time, developers may consider improving the quality of game content, for example by fixing bugs and crashes, or by building game communities, both corresponding to complaint categories ranked above VR-comfort issues.

VII. RELATED WORK

In this section, we discuss related work about (1) virtual reality complaints, (2) game repository mining, and (3) games and software engineering.

Virtual reality complaints: Most of the work which focuses on what issues users of current VR technology have with the technology is concerned with cybersickness. Porter III et al. [36] studied discussions on Reddit about cybersickness and immersion on the *HTC Vive*. Similar to our study, they conclude that the VR games market is maturing and that the concerns of gamers have evolved over time.

A few studies have investigated the severity of VR sickness among current VR games. Munafo et al. [30] reported 22% and 56% incidence of motion sickness when participants played two VR games (*Balance Rift* and *Affected*) on an *Oculus Rift DK1*. Shafer et al. [41] conducted an experiment across three commercial VR games and two headsets (*Oculus Rift DK1* and *CV1*), and concluded that factors important to the enjoyment of other types of games were also important for VR games, and that the different headsets had no impact on the severity of cybersickness. Walch et al. [49] compared a racing game played on either flat screens or an *HTC Vive* headset. They found that the visualization method had no significant effect on SSQ scores (a measure of cybersickness), although users of the VR setup felt a significantly greater amount of discomfort. Tan et al. [46] explored gameplay experiences from the game *Half-Life 2* on the *Oculus Rift DK1*. While 8 of 10 participants experienced cybersickness at some point, it largely did not effect immersion. Rangelova et al. [38] conducted the only

large-scale examination of cybersickness among gamers of which we are aware. Using an online survey, they concluded that cybersickness is widespread among players of current VR games. While our study does not measure the incidence or severity of cybersickness among players, we do examine user reviews complaining of its symptoms, and are thus able to rank its perceived severity relative to other player complaints.

Several studies have leveraged user reviews to study complaints for other platforms such as mobile apps. Khalid et al. [17] studied 6,390 low-rated reviews for 20 free iOS apps. They found that the most frequent complaints were of functional errors, feature requests, and crashes, while the most negatively-impacting complaints were of ethical and privacy issues. Mcilroy et al. [28] found that up to 30% of low-rated reviews contained multiple complaints, and presented a classifier to identify complaints from user reviews. Fu et al. [8] presented a system which analyzed how user complaints evolved over time. Mujahid et al. [29] investigated complaints in 2,667 reviews of 19 Android wearable apps. The most frequent complaints were functional errors and cost, and the most negatively-impacting complaints were installation, incompatibility, and ethics issues. Hassan et al. [12] studied the top 250 bad updates (updates with the highest increase in the percentage of negative reviews following the update) of 2,526 free apps on the Google Play Store. They found that feature removal and user interface issues caused the highest increase in negative reviews but that developers were most likely to fix crashes and functional issues. Our study investigates PC-based VR games, and is the first to show that the occurrence of complaints about cybersickness is low amongst players of popular VR games on Steam.

Game repository mining: Several studies have mined data from digital game distribution platforms. Sifa et al. [43] analyzed the playtime distributions of over 3,000 games on the Steam platform, grouped them into four archetypes, and described the main types of games within each archetype. Poretski and Arazy [35] collected data for 45 games from the *Nexus Mods* game mod distribution platform to investigate the value added by modding communities. Our prior work has collected game and review data from Steam in order to analyze early access games [21], urgent updates [20], and user reviews [23]. In our prior work on user reviews, we studied reviews of all types of games. In addition, we did not focus on complaints of gamers. In this paper, we focus on complaints from VR gamers in particular. VR games and their complaints must be studied separately from traditional games due to the presence of hardware that may affect the gamers' complaints.

Many studies focused on the social media aspects of Steam Community and on characterizing player behaviour. O'Neill et al. [32] collected data from 109 million user accounts and 384 million owned games on Steam. They reported a low number of friendships compared to other social networks, among other findings, and emphasized the diversity of player behaviour within their results. Becker et al. [2] analyzed the Steam Community social network, user groups, and the evolution of the network over time. Sifa et al. [44] used the Steam web API to analyze cross-game behaviour of 6 million players across more than 3,000 games. Blackburn et al. [4]

observed "cheaters" within Steam Community and found that cheating behaviour spreads with a "contagion"-like effect, with players having cheating friends being more likely to later become cheaters themselves. Li et al. [19] collected user profile for 60 thousand users within Steam Community and extracted eight factors characterizing user attributes. Baumann et al. [1] analyzed so-called "hard-core" gamers.

We collect game data and reviews for VR games on Steam, and our work is the first to focus specifically on VR games.

Games and software engineering: There is a large amount of work studying computer games and various aspects of software engineering. Much of this is work concerned with the game creation process, such game as architecture [9, 18, 31] or development practices [16, 33, 34].

Numerous studies have applied data mining techniques to compute game data. Many of these have used in-game telemetry, public APIs, or game logs to analyze player behavior for a single game or small number of games, e.g. [10, 11, 13, 14, 42, 50]. Among the studies which consider a large number of games as we do, a few gather data primarily from platforms other than digital distribution platforms. Chambers et al. [5] examined a collection of 550 online games using data from *GameSpy.com* (now defunct) to investigate game workloads and the potential for shared infrastructure. Cheung et al. [6] gathered 200 reviews from a combination of Amazon and long-form reviews from gaming websites, as well as interviews from industry professionals, and recommended a focus on engagement rather than fun for the first playing hour of games. In our prior work [22], we collected metadata of Youtube game videos and trained a random forest classifier to identify videos showcasing game bugs.

VIII. THREATS TO VALIDITY

Internal validity: A potential threat to the validity of our study is that we filtered out games with fewer than 25 reviews. Recent games could be underrepresented since they have not had much time to accumulate enough reviews. Future studies should repeat our analysis in several years to understand how trends have changed since then.

Another threat is that user reviews are prone to review bombs, i.e., large numbers of negative reviews to discredit a game. While we did not observe any indications of review bombs in our data, future studies should further sanitize the collected data to reveal the impact of such bombs.

In order to categorize user complaints, we only studied reviews written in English. We also examined only negative reviews, even though our prior work has shown that positive reviews also frequently contain complaints [23]. Because of the small percentage of positive reviews which contain complaints, it is difficult to automatically extract complaints from these reviews. Future studies should investigate the complaint categories in positive reviews of VR games.

The HTC Vive headset was developed by HTC in partnership with Valve, the owner of Steam. This close relationship with Valve likely explains the large fraction of games which support that headset.

Construct validity: We estimated trends in headset, controller, and play area support by combining the release date

of each game with the stated support on Steam. However, developers may update these values at any time, and so the stated headset, controller, and play area support among earlier games may not reflect their values at launch. In particular, we found 116 games which were released before the launch of Windows Mixed Reality headsets, but which officially supported those headsets at the time our data was collected.

We estimated the number of updates to each game by filtering the game's news posts for release notes. In our prior work with early access games, we evaluated this method to identify release notes with a precision of 89% and a recall of 87% [21]. Since posting such release notes is not mandatory, our method could underestimate the number of updates.

Twitter-LDA only assigns one topic per review, even though a review might contain multiple complaints. Our manual categorization also assigned only one category per topic. Assigning only one topic per review might cause complaints of secondary importance within user reviews to be disregarded. To assess the impact of this threat, we manually studied a randomly selected sample of 70 negative reviews to verify what portion of the reviews contains multiple complaints. We found that approximately only 23% of the reviews contained complaints from more than one category. However, most of these reviews were categorized into the game-specific and the controls category. During the manual categorization in the rest of the paper, we always favoured the other complaint categories over the game-specific category. Hence, we do not expect that the fact that we assigned only one category per topic influences the eventual outcome of our study. Regardless, future studies should further investigate the impact of allowing multiple complaint categories per review.

External validity: We only examine games with mandatory VR support. This excludes games which have both VR and non-VR modes, and so our results may not generalize to those games. Unfortunately, this restriction was unavoidable since games with optional VR support may contain reviews both of the VR and non-VR content, with no method to determine which is which. Likewise, we only study popular PC-based VR games on Steam and our results may not generalize to VR games for other platforms, such as PlayStation VR, or PC-based VR games that are not available on Steam.

Our findings may not hold for the entire potential audience for VR games. For example, players sensitive to cybersickness may be less likely to own a VR headset and to review VR games on Steam, possibly making them underrepresented in our results. Finally, our results apply to PC-based headsets since, as explained in Section II, these comprise the majority of VR headsets used with Steam. Our results may not generalize to other types of VR headsets, such as standalone headsets.

IX. CONCLUSIONS

The unique display and input hardware used by VR games allows these games to offer different experiences than non-VR games. Understanding what complaints players have about VR games may allow developers to tailor their development and testing efforts to most effectively address user concerns.

We performed an empirical study of 750 PC-based VR games and 17,635 user reviews on the Steam platform. We ex-

amined trends in release frequency, price, headsets, controllers, and play areas, and also examined the games' developers and updates. We extracted seven common categories from the negative user reviews in order to understand player complaints, and examined the frequency of these complaint categories over time. Our most notable findings are:

- 1) The majority of recent releases have support for multiple headset categories and play areas. In particular, support for smaller-scale playing areas has increased.
- 2) Complaints that games are overpriced for their content have decreased sharply, even though the median price of VR games has more than doubled over time.
- 3) The frequency of complaints reflecting cybersickness is low and has decreased steadily over time.
- 4) Since approximately 2018, the most common complaint among user reviews has been with game-specific rather than VR-comfort issues.

Our findings show that the PC-based VR games market is maturing. In particular, we observe that VR-comfort issues which have traditionally received widespread attention in research, such as cybersickness, are uncommon within complaints and have become less prevalent. There are several possible explanations for this reduction in complaints about VR-comfort. The first possible explanation is that PC-based VR technology and design patterns for VR have greatly improved over the years, allowing developers to create VR experiences that are more comfortable and less nausea-inducing. A second possible explanation is that players who suffer from VR-comfort issues when playing VR games simply no longer play these games, thereby reducing such complaints (even though the issues may still exist). Future studies should further investigate these possible explanations for the reduction in complaints about VR-comfort.

Overall, our findings are an indication that developers of PC-based VR games now no longer should have improving VR-comfort as their sole focus. Instead, they can start focusing more on improving the design and gameplay of VR games to attract a wider range of players (while, of course, maintaining the same level of VR-comfort).

REFERENCES

- [1] F. Baumann, D. Emmert, H. Baumgartl, and R. Buettner, "Hardcore gamer profiling: Results from an unsupervised learning approach to playing behavior on the Steam platform," in *Knowledge-Based and Intelligent Information & Engineering Systems: Proc. of the 22nd International Conference KES-2018, Belgrade, Serbia, 3-5 September 2018.*, 2018, pp. 1289–1297.
- [2] R. Becker, Y. Chernihov, Y. Shavitt, and N. Zilberman, "An analysis of the Steam community network evolution," in *IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, Nov 2012, pp. 1–5.
- [3] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. O'Reilly Media Inc., 01 2009.
- [4] J. Blackburn, N. Kourtellis, J. Skvoretz, M. Ripeanu, and A. Iamnitchi, "Cheating in online games: A social network perspective," *ACM Trans. Internet Technol.*, vol. 13, no. 3, pp. 9:1–9:25, May 2014.
- [5] C. Chambers, W.-c. Feng, S. Sahu, and D. Saha, "Measurement-based characterization of a collection of on-line games," in *Proc. of the 5th ACM SIGCOMM Conference on Internet Measurement*. Berkeley, CA, USA: USENIX Association, 2005, pp. 1–1.
- [6] G. K. Cheung, T. Zimmermann, and N. Nagappan, "The first hour experience: How the initial play can engage (or lose) new players," in *Proc. of the First ACM SIGCHI Annual Symposium on Computer-human Interaction in Play*, 2014, pp. 57–66.

- [7] L. Freina and M. Ott, "A literature review on immersive virtual reality in education: State of the art and perspectives," 04 2015.
- [8] B. Fu, J. Lin, L. Li, C. Faloutsos, J. Hong, and N. Sadeh, "Why people hate your app: Making sense of user feedback in a mobile app store," in *Proc. of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2013, pp. 1276–1284.
- [9] A. W. B. Furtado, A. L. M. Santos, G. L. Ramalho, and E. S. de Almeida, "Improving digital game development with software product lines," *IEEE Software*, vol. 28, no. 5, pp. 30–37, Sep. 2011.
- [10] A. R. Gagné, M. Seif El-Nasr, and C. D. Shaw, "Analysis of telemetry data from a real-time strategy game: A case study," *Comput. Entertain.*, vol. 10, no. 1, pp. 2:1–2:25, Dec. 2012.
- [11] E. Harpstead, T. Zimmermann, N. Nagapan, J. J. Guajardo, R. Cooper, T. Solberg, and D. Greenawalt, "What drives people: Creating engagement profiles of players from game log data," in *Proc. of the Annual Symposium on Computer-Human Interaction in Play*. ACM, 2015, pp. 369–379.
- [12] S. Hassan, C. Bezemer, and A. E. Hassan, "Studying bad updates of top free-to-download apps in the Google Play store," *IEEE Transactions on Software Engineering*, pp. 1–1, 2018.
- [13] J. Huang, T. Zimmermann, N. Nagappan, C. Harrison, and B. Phillips, "Mastering the art of war: How patterns of gameplay influence skill in Halo," in *Proc. of the International Conference on Human Factors in Computing Systems*. ACM, April 2013.
- [14] K. Hullett, N. Nagappan, E. Schuh, and J. Hopson, "Data analytics for game development (NIER Track)," in *Proc. of the 33rd International Conference on Software Engineering*. ACM, 2011, pp. 940–943.
- [15] D. Jack, R. Boian, A. S. Merians, M. Tremaine, G. C. Burdea, S. V. Adamovich, M. Recce, and H. Poizner, "Virtual reality-enhanced stroke rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 9, no. 3, pp. 308–318, Sep. 2001.
- [16] J. Kasurinen, M. Palacin-Silva, and E. Vanhala, "What concerns game developers?: A study on game development processes, sustainability and metrics," in *Proc. of the 8th Workshop on Emerging Trends in Software Metrics*. IEEE, 2017, pp. 15–21.
- [17] H. Khalid, E. Shihab, M. Nagappan, and A. E. Hassan, "What do mobile app users complain about?" *IEEE Software*, vol. 32, no. 3, pp. 70–77, May 2015.
- [18] J. Kienzie, A. Denault, and H. Vangheluwe, "Model-based design of computer-controlled game character behavior," in *Proc. of the 10th International Conference on Model Driven Engineering Languages and Systems*. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 650–665.
- [19] X. Li, C. J. Lu, J. Peltonen, and Z. Zhang, "A statistical analysis of Steam user profiles towards personalized gamification," in *GamiFIN*, 2019.
- [20] D. Lin, C.-P. Bezemer, and A. E. Hassan, "Studying the urgent updates of popular games on the Steam platform," *Empirical Softw. Engg.*, vol. 22, no. 4, pp. 2095–2126, Aug. 2017.
- [21] —, "An empirical study of early access games on the Steam platform," *Empirical Softw. Engg.*, vol. 23, no. 2, pp. 771–799, Apr. 2018.
- [22] —, "Identifying gameplay videos that exhibit bugs in computer games," *Empirical Softw. Engg.*, Jun 2019.
- [23] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the Steam platform," *Empirical Softw. Engg.*, vol. 24, no. 1, pp. 170–207, Feb. 2019.
- [24] J. D. Long, D. Feng, and N. Cliff, *Ordinal Analysis of Behavioral Data*. American Cancer Society, 2003, ch. 25, pp. 635–661.
- [25] J. Lyons, "New today: Experience SteamVR on Windows Mixed Reality," 2017, <https://steamcommunity.com/games/719950/announcements/detail/1466347451617826043>, Accessed: 2019-05-13.
- [26] K. M. Malloy and L. S. Milling, "The effectiveness of virtual reality distraction for pain reduction: A systematic review," *Clinical Psychology Review*, vol. 30, no. 8, pp. 1011 – 1018, 2010.
- [27] E. Martel, F. Su, J. Gerroir, A. Hassan, A. Girouard, and K. Muldner, "Diving head-first into virtual reality: Evaluating HMD control schemes for VR games," in *Proc. of the 10th International Conference on the Foundations of Digital Games*, 2015.
- [28] S. McIlroy, N. Ali, H. Khalid, and A. E. Hassan, "Analyzing and automatically labelling the types of user issues that are raised in mobile app reviews," *Empirical Softw. Engg.*, vol. 21, no. 3, pp. 1067–1106, Jun. 2016.
- [29] S. Mujahid, G. Sierra, R. Abdalkareem, E. Shihab, and W. Shang, "An empirical study of Android Wear user complaints," *Empirical Softw. Engg.*, vol. 23, no. 6, pp. 3476–3502, Dec 2018.
- [30] J. Munafo, M. Diedrick, and T. A. Stoffregen, "The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects," *Experimental Brain Research*, vol. 235, no. 3, pp. 889–901, Mar 2017.
- [31] T. Olsson, D. Toll, A. Wingkvist, and M. Ericsson, "Evolution and evaluation of the model-view-controller architecture in games," in *Proc. of the Fourth International Workshop on Games and Software Engineering*. Piscataway, NJ, USA: IEEE Press, 2015, pp. 8–14.
- [32] M. O'Neill, E. Vaziripour, J. Wu, and D. Zappala, "Condensing Steam: Distilling the diversity of gamer behavior," in *Proc. of the Internet Measurement Conference*. ACM, 2016, pp. 81–95.
- [33] A. Osborne O'Hagan, G. Coleman, and R. V. O'Connor, "Software development processes for games: A systematic literature review," in *Systems, Software and Services Process Improvement*. Springer Berlin Heidelberg, 2014, pp. 182–193.
- [34] C. Politowski, L. Fontoura, F. Petrillo, and Y. Guéhéneuc, "Are the old days gone? a survey on actual software engineering processes in video game industry," in *2016 IEEE/ACM 5th International Workshop on Games and Software Engineering (GAS)*, May 2016, pp. 22–28.
- [35] L. Poretski and O. Arazy, "Placing value on community co-creations: A study of a video game 'modding' community," in *Proc. of the Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 2017, pp. 480–491.
- [36] J. Porter III and A. Robb, "An analysis of longitudinal trends in consumer thoughts on presence and simulator sickness in VR games," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY)*. ACM, 2019, pp. 277–285.
- [37] J. Porter III, M. Boyer, and A. Robb, "Guidelines on successfully porting non-immersive games to virtual reality: A case study in Minecraft," in *Proc. of the 2018 Annual Symposium on Computer-Human Interaction in Play*. ACM, 2018, pp. 405–415.
- [38] S. Rangelova, D. Motus, and E. André, "Cybersickness among gamers: An online survey," in *Advances in Human Factors in Wearable Technologies and Game Design*, T. Ahram, Ed. Cham: Springer International Publishing, 2020, pp. 192–201.
- [39] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Real.*, vol. 20, no. 2, pp. 101–125, Jun. 2016.
- [40] J. Romano, J. D. Kromrey, J. Coraggio, J. Skowronek, and L. Devine, "Exploring methods for evaluating group differences on the NSSE and other surveys: Are the t-test and Cohen's d indices the most appropriate choices," in *Annual Meeting of the Southern Association for Institutional Research*, 10 2006.
- [41] D. Shafer, C. Carbonara, and M. Korpi, "Factors affecting enjoyment of virtual reality games: A comparison involving consumer-grade virtual reality technology," *Games for Health Journal*, vol. 8, 09 2018.
- [42] K. B. Shores, Y. He, K. L. Swanenburg, R. Kraut, and J. Riedl, "The identification of deviance and its impact on retention in a multiplayer game," in *Proc. of the 17th Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2014, pp. 1356–1365.
- [43] R. Sifa, C. Bauckhage, and A. Drachen, "The playtime principle: Large-scale cross-games interest modeling," in *Conference on Computational Intelligence and Games*, Aug 2014, pp. 1–8.
- [44] R. Sifa, A. Drachen, and C. Bauckhage, "Large-scale cross-game player behavior analysis on Steam," in *Proc. of the Eleventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2015, pp. 198–204.
- [45] SuperData Research Holdings, Inc., "State of the consumer XR market," 2019, <https://www.superdataresearch.com/state-of-consumer-xr-2019/>, Accessed: 2019-06-11.
- [46] C. T. Tan, T. W. Leong, S. Shen, C. Dubravs, and C. Si, "Exploring gameplay experiences on the Oculus Rift," in *Proc. of the Annual Symposium on Computer-Human Interaction in Play*. ACM, 2015, pp. 253–263.
- [47] Valve Corporation, "Steam hardware & software survey: June 2019," 2019, <https://store.steampowered.com/hwsurvey>, Accessed: 2019-07-19.
- [48] M. Virvou and G. Katsionis, "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE," *Comput. Educ.*, vol. 50, no. 1, pp. 154–178, Jan. 2008.
- [49] M. Walch, J. Frommel, K. Rogers, F. Schüssel, P. Hock, D. Döbelstein, and M. Weber, "Evaluating VR driving simulation from a player experience perspective," in *Proc. of the CHI Conference Ext. Abstracts on Human Factors in Computing Systems*. ACM, 2017, pp. 2982–2989.
- [50] G. Wallner, "Sequential analysis of player behavior," in *Proc. of the Annual Symposium on Computer-Human Interaction in Play*. ACM, 2015, pp. 349–358.
- [51] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li, "Comparing Twitter and traditional media using topic models," in *Proc. of the 33rd European Conference on Advances in Information Retrieval*. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 338–349.