Crowdjump: Investigating A Player-Driven Platform Game

Pascal Lessel
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus
pascal.lessel@dfki.de

Maximilian Altmeyer
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus
maximilian.altmeyer@dfki.de

Nicolas Brauner
Saarland University
Saarland Informatics Campus
s9nsbrau@stud.uni-saarland.de

ABSTRACT
We present Crowdjump, a player-driven online platform game: players could post ideas on how to improve the game, the platform, or any related aspect. Continuously, all players could then choose ideas which were implemented and released on a daily basis. We conducted an exploratory study with 25 players over the course of 23 days and aimed at releasing two features per day. We analyzed idea types and could show that players did not change an idea voting scheme based on up- and down-votes and were more focused on changing the game than any other component. In contrast, features that would improve the community feeling or would help to improve ideas of others were not often suggested or selected. Nonetheless, the experience was rated as quite enjoyable by players, showing the appeal of such a player-driven game design approach and the relevancy for further research in this context.

CCS Concepts
•Human-centered computing → Empirical studies in collaborative and social computing;

Author Keywords
Game design; “Bottom-up”; User-led design; Modding

INTRODUCTION
Large modding communities exist in which people create and offer new content for games [23]. Thus, players can alter their game experience in ways not intended by the original developers [23]. Although these often provide tools with which the game can be modded, modders need skills (e.g., artistic or programming skills) and time to implement their mods [23]. While some communities allow users to provide feature requests (e.g. [29]), whether these are realized depends on the developers/modders, i.e., some ideas may never be realized, limiting the actual impact the community can have.

An interesting approach in this respect is please be nice :) [1]. In 2014, a relatively simple version of this game was released: a red rectangle only needed to be moved to a designated area (see Figure 1, left). The first player who reached this goal was allowed to submit a feature request for the next game version. As soon as the feature was implemented (they aimed for daily releases), the next cycle started in the same fashion. 133 versions were released in roughly five weeks and altered the game considerably (see Figure 1, right).

We started the scientific exploration of a similar, player-driven game design approach by developing an online platform game, called Crowdjump on which the community was able to submit and select textual ideas. In contrast to please be nice :) , every registered user was able to submit ideas. Ideas could be submitted for both the game and all related aspects, e.g., the online web page or the way ideas are selected. Our goal was to see how this platform game evolves when the players have direct influence options and know that we aim to release new features (based on their suggested and selected ideas) daily.

This paper contributes the design of Crowdjump and the results of a pilot and a 23-day main study. During the latter, 138 ideas were suggested and 43 new features were released (see Figure 2 for how Crowdjump evolved visually). We found that most of the suggested and selected ideas focused on the game itself. Players kept a voting scheme based on up- and down-votes for the idea selection and were satisfied with two idea realizations a day. Although players had the chance to evolve the game in a truly community-driven way, almost no collaborative features to improve ideas of others were selected and picked. Nonetheless, the players reported having a high enjoyment, even those that dislike platform games, indicating the appeal of such a player-driven game design.

RELATED WORK
Besides please be nice :) , the commercial roleplaying game RuneScape [10] also provided their player-base design options [20], such as feedback cycles, sharing design documents and polls before features the designers envisioned were implemented. In addition, with RuneLabs ideas could be submitted
Figure 2. First version of *Crowdjump* and three example versions during the main study. Images also show the idea as provided by the user (i.e., quoted as provided, or translated if provided in German), the version number and day (in respect to the study).

that could be voted on by the players. According to [20], 1/3 of the game features realized were based on this (which is a difference from our approach in which every feature is decided on by the players). While the positive effects were mentioned through the different integrative options (e.g., that players appear to get more involved with the game), to our knowledge, these were not scientifically reported. Nonetheless, both examples show that there are already non-academic approaches that consider how to give players more influence.

A motivation for our research also comes from game modding: Based on, for example, developers that allow others to modify parts of their game [22, 23] (e.g., by offering Software Development Kits), modding communities have established themselves. Resulting modifications are offered to a large audience (i.e., other players), and thus allow individuals, even those who are not capable of creating their own modifications, to alter their game experience. Nicholson calls this “player-generated content” and postulates that allowing for such is beneficial to extend the life of a game and “allows designers to see how creative users can be with the toolkits provided” [16].

Scacchi [25] states that this is a form of user-led innovations and an approach to tailor games. Poretski and Arazy [22] showed that companies that support the modding of their games attract many people that are willing to spend hours in creating modifications and that those affect sales positively. Nonetheless, creation of a mod takes effort and considerable time investment from the modders [23]. The wish for identification with the game and the desire to create one’s own experience is a strong motivation for doing this [23]. Based on, for example, the Self-Determination Theory [24], it can be assumed that similar needs are also present for the players who do not have the skill or time to create mods themselves. With a player-driven approach such as the one considered and investigated in this paper, we also give these players a chance to express their wishes, which could have a positive impact on their enjoyment and perception of the game.

Approaches to empower users to create their own games or game-related experiences are also considered scientifically: for example, in the work of Guy et al. [7], employees of a company received the option to create their own (relatively simple) games through a wizard component to allow for an easy game creation. In a three-month study, 34 games were created that attracted 339 players in the company. Interestingly, the player base demanded more choices in the creation of their games, underlining the above aspects of users who want to have an impact. Several commercial games also offer their players easy ways to shape and create their “own” experience today (e.g., *Super Mario Maker* [18], *Minecraft* [15] or *Dreams* [14]), showing that this is a movement that is currently an important theme in the domain.

Recently, in the gamification domain, several approaches also provided strong customization options for the users of these systems: Siu and Riedl [28], for example, allowed some of their users to change the reward scheme in a game (leaderboards, customizable avatars, unlockable narratives, or a global progress tracker); others received a reward randomly.
As shown in a study, this choice led to better task completion times. In the work of Lessel et al. [12, 13], users even received options to set up all gamification elements in a system through a simple wizard system. In user studies, the authors could also show that people not only are interested in setting up their own gamification, but that this also has positive effects on the perception of the intervention and task completion times in a microtask setting. With the study of Crowdjump, we contribute one further option to approach the question of how to enable a direct user participation in the creation of game experiences. The approach chosen in this paper can be seen as a direct form of participatory design [26], in which individuals can provide ideas and the community can pick the most relevant ones that are made available in a system.

CROWDJUMP

We implemented a 2D platform game to increase the chance that players could easily relate to it and have suggestions for how to shape it. This was based on the fact that, for example, the platform game Super Mario Bros. [19], sold over 180 million copies [21] and has already been used in games research (e.g., [27]). The base version of Crowdjump is rudimentary (see Figure 2, upper left): an avatar needed to cross static platforms to reach a flag, after which the (only) level could be re-played. We integrated a jumping sound and a sound when reaching the flag. The time was counted, shown in the game and after a run, but not further used. Users could control the avatar with the arrow keys.

In contrast to please be nice [1], in which every version needed to be downloaded, we designed Crowdjump as a game on a web page to lower the participation threshold. On this page, we allowed (registered) users to post new ideas for the game (consisting of a title and a description). Furthermore, every user could up- and down-vote other ideas (with one vote per idea) to indicate that they want them to be implemented (i.e., the sum after subtracting the down-votes from the up-votes would then be the relevant score). Users saw new ideas first, but could filter these by version number (i.e., for which game version they were submitted), the submitter’s user name, the idea’s title and description, and the state of the idea (i.e., not realizable, realized, suggested). They could also adapt the sorting of the ideas: either by date or by the amount of up- and down-votes. As soon as an idea was suggested, we estimated the time for realizing it and added this time to the idea. Ideas that we assessed as too time-consuming to realize (within our goal of frequently releasing a new update) we stated that it is “not realizable” and provided an explanation for this. Selected ideas that were currently implemented were highlighted prominently on the ideas page as well (including who submitted them). Figure 3 shows an excerpt of this page.

The web page also offered a history (i.e., a version overview) to allow the community to inspect how Crowdjump had evolved. We always highlighted the current state of the game textually (on this sub-page as well as in emails when a new version was released), by putting emphasis on the game mechanics, the graphics of the game, the idea selection method and the web page on which the game was running. For the starting version of the main study, the statements shown were:

- Just a very simple platform game. You could request more game related features!
- The game is black, white and not that pretty. Maybe design requests would help!
- You can vote for an idea to be implemented! If you want another way of selecting the features, submit it!
- There is not much on the website. If you have ideas to alter the website, submit them!

In contrast, the text for the last version of the main study was:

- Teleporters give some alternative routes! Question mark boxes can give you all kinds of items like the classic double jump, the time power-up as well as coins for more lives which should help you to pass the more difficult levels, filled with cannons and enemies! If not, there is now the possibility to practice levels you already completed! You could still submit some new game play features!
- Power-ups now explode when you collect them! Platforms have color, as well as the flag! Decorations and background change depending on the time you play! Background music, animations and new characters, the game gets more and more vivid! You can still ask for more animations or decorations!
- You can vote for an idea to be implemented! If you want another way of selecting the features, submit it!
- There is a new page where you can report bugs! There is now an explanation of the items, random statistics, you can see the current “Master of Masters”, the high score list and comment on ideas! If you have more ideas to alter the website, submit them!

The main purpose of these statements was to give users, even if they are not experienced players, hints on what types of ideas they could submit. Also, older Crowdjump versions could be inspected on the history page, with their associated statements. Overall, the web page was served in English.

PILOT STUDY

We conducted a pilot study to learn about the types of ideas suggested, how fast they could be realized and which potential issues could arise with the web page or the study setting. To this end, we also integrated questionnaires to mimic the method that we planned for the main study. In this section, though, we will not elaborate on all the details, due to the small sample size and thus limited expressiveness.
Method
We sent the link to our page to the (unpaid) participants. A survey had to be completed, covering questions on their video game experience/preferences and previous experiences with game design and please be nice :)\(^1\). The page and the game were then available as described in the previous section, with the only exception that there was no voting mechanism. Instead, a suggested idea was randomly picked to see how the participants behave in this case, and whether a selection scheme is suggested, when and what kind. The pilot ran for 15 days, with the aim to release a new version every day at 7pm. This time was chosen to increase the chance that participants would have finished their workday and would have time to try out the new version. At the same time, the idea for the next version was chosen and announced. Overall, 13 versions were released. Every release was accompanied with an email notification to the participants also highlighting the new feature. After the 15 days, participants filled out a questionnaire, consisting of the System Usability Score (SUS) [3], the Game Experience Questionnaire (GEQ) [9], the Social Presence in Gaming Questionnaire (SPGQ) [5] and statement-based questions to learn about the players’ perception of Crowdjump. This was complemented by interviews (about 10 minutes) with every participant individually. The interviews were transcribed and thematically coded [8] afterward. Both studies presented in this paper were approved by the Ethical Review Board of Saarland University (Department of Computer Sciences).

Results
Five German participants, following [17], (4 male, all between 18–26 years old) with a good command of the English language participated and were recruited via the social circle of one of the authors. Four participants reported prior experiences with software development and three stated that they would like to participate in the development of a video game. Every participant claimed to interact with computers at least 10 hours a week. The sample tended to dislike platform games (7-point scale, mean M=3.6, standard deviation SD=1.1, median Mdn=4). No one knew about please be nice :)\(^1\).

Usage of Crowdjump
Overall, the game was played 4374 times during the study and 34 different ideas were suggested, of which 13 were implemented (one idea took two days to implement). Table 1 shows how many ideas were provided per participant and how involved they were with the game. One idea took two days for its realization, two ideas took one day, three realizations four to six hours and seven took less than two hours. We compared the 13 manual time estimations for the implementations that we provided on the ideas page (when an idea was submitted) with the actual time that the realization needed. Seven times the estimation and realization matched, five times we were faster and only one time we needed longer than estimated (four instead of two hours). From a player perspective this shows that our estimates were reliable: 92.3% of the time we were at least as fast as estimated and in all the cases the promised releases of a feature happened as announced (at 7pm).

\(^1\)Questions used in the pilot and main study can be found in the supplementary material.

<table>
<thead>
<tr>
<th>ID</th>
<th>Ideas Suggested</th>
<th>Ideas Chosen</th>
<th>Game Time</th>
<th>Tries</th>
<th>Platformer Preference (PP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>4</td>
<td>34</td>
<td>133</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>270</td>
<td>590</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>284</td>
<td>1894</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>4</td>
<td>197</td>
<td>1552</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>2</td>
<td>82</td>
<td>205</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1. How many ideas a participant (ID) suggested (Suggested), how many of these were implemented (Chosen), how long (Time, in minutes) and how often (Tries) the participant played, and the participant’s rating of the platform games genre (PP; single item 7-point scale question). We counted a try as one life of the avatar, ending either when the avatar died, the level was reset, or all levels were completed.

Perception of Crowdjump
The SUS (questions were framed with “the system”, i.e., covering both the game and the web page) was on average 86.5 (SD=8.8, Mdn=87.5), indicating an excellent usability [2]. In the 5-point-scale statement-based questions in the closing questionnaire, participants reported liking the idea of Crowdjump (M=4.2, SD=.8, Mdn=4) and being able to submit new ideas (M=4, SD=0, Mdn=4). They also agreed to the statement that the game (M=4.6, SD=.5, Mdn=5), the website (M=4.6, SD=.5, Mdn=5) and the idea selection process (M=4.2, SD=.4, Mdn=4) evolved positively during the study. Fitting with the low number of community-related features (see below), the players disagreed with the statement that they had formed a community (M=1.8, SD=1.3, Mdn=1) and were indifferent as to whether other players had interfered with the development of the game (M=3, SD=1, Mdn=3).

None of the participants complained about the overall study methodology in the interviews and the positive perception of Crowdjump could be seen there as well: every participant talked about the fact that compared to the initial version, the game changed considerably and that they were content with the changes made to the game and the web page. They also stated that they could imagine continuing playing and participating in Crowdjump, and that they would recommend the project to others as well. In addition, every participant had the feeling that although not all of their ideas were realized, the important ones were, and those that were realized were perceived mostly positively.
<table>
<thead>
<tr>
<th>Version</th>
<th>Idea</th>
<th>Participant</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>The game has low risk reward ratio. By adding a floor made out of lava or something similar material, you could punish a player for failing a jump.</td>
<td>4</td>
<td>Game</td>
</tr>
<tr>
<td>3</td>
<td>Show the highscores on the website.</td>
<td>2</td>
<td>Game, Community</td>
</tr>
<tr>
<td>4</td>
<td>Provide multiple levels.</td>
<td>5</td>
<td>Game</td>
</tr>
<tr>
<td>5</td>
<td>Up- and downvote ideas by users, best idea should be chosen.</td>
<td>1</td>
<td>Community</td>
</tr>
<tr>
<td></td>
<td>Voting system for new ideas. The ideas with the most upvotes are going to be implemented in the next version.</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Like in Mario, make the level larger such that the screen scrolls to the left and right.</td>
<td>1</td>
<td>Game</td>
</tr>
<tr>
<td>7</td>
<td>Enemies run around the level and kill you.</td>
<td>3</td>
<td>Game</td>
</tr>
<tr>
<td>8</td>
<td>Restarting from the beginning is frustrating, if a level is won, the alien should restart at the first level that was not completely done.</td>
<td>1</td>
<td>Game</td>
</tr>
<tr>
<td>9</td>
<td>Add moving game parts to the level to increase difficulty.</td>
<td>2</td>
<td>Game</td>
</tr>
<tr>
<td>10</td>
<td>Offer the possibility to replay various levels.</td>
<td>5</td>
<td>Game</td>
</tr>
<tr>
<td>11</td>
<td>Commenting ideas of other users could further develop and improve them.</td>
<td>4</td>
<td>Community</td>
</tr>
<tr>
<td>12</td>
<td>There could be items which, when collected, increase speed or jump power or fire/lava resistance.</td>
<td>1</td>
<td>Game</td>
</tr>
<tr>
<td>13</td>
<td>When you enter a checkpoint, save the time and reload this time, if you die and respawn at this checkpoint.</td>
<td>2</td>
<td>Game</td>
</tr>
<tr>
<td>14</td>
<td>Add different platform types. Platforms that switch between normal and lava - Platform made out of ice you can slide on - sticky walla you can jump off...</td>
<td>4</td>
<td>Game</td>
</tr>
</tbody>
</table>

Table 2. The realized ideas (quoted as provided by the participants, or translated if provided in German) in the pilot study. Ideas for version 6+ were selected by the community; before this, an idea was chosen randomly. The idea for version 5 was provided by two participants.

**Evolution of Crowdjump**

Table 2 shows the ideas that were realized. Before version 6, the ideas were chosen randomly. Version 5 realized a voting mechanism with up- and down-voting options (as explained in the Crowdjump section), i.e., the community could choose which ideas should be implemented next. This was suggested by two participants (on day two and day four). In total, 56 up- and 7 down-votes were provided, showing that this feature was used. Most of the implemented features were related to the game itself. The time necessary to complete a game run increased considerably: in version 1, the game could be played through in roughly five seconds; in the last version a (perfect) run through all levels of the game took up to 50 seconds. While version 1 started with the game as explained in the previous section, Figure 4 shows a screenshot from one of the levels of version 14. Throughout the versions, the game got more difficult. While in the first versions 60% of all tries were successful (i.e., participants completed all levels of the game), this dropped to 2.5% in the end. Only the first three game versions were completed by every participant. For the later parts of this and the main study, we concluded that we needed to make the level designs easier: shortcuts were implemented that would reduce the time needed to complete a level (see Figure 5 for an example), but were quite challenging and aimed for the more experienced player. With this, better players could still compete on the high score lists, but less experienced players would still be able to finish the game.

Table 2 indicates that ideas were only described roughly (character length of all ideas, including spaces, was Min=24, Max=367, M=100, SD=62, Mdn=91), potentially to give developers room to realize an idea and/or because players did not want to go into the full details. For example, the player who suggested the extension to multiple levels (version 4) did not elaborate on the number of levels, the size or the differences between the levels. In the closing questionnaire the participants agreed with the statement that ideas were realized as requested (M=4.2, SD=.8, Mdn=4) and disagreed with the statement that their opinion was not heard (M=1.8, SD=1.1, Mdn=.9). This indicates that they were content with how their ideas were realized and that although many ideas were not selected, that seemed not to be an issue for them.

In the interviews, two times, the increase in the difficulty level was mentioned as a negative side effect of some ideas, and also twice, it was mentioned that the checkpoint ideas (version 8 and 13) were a design decision that those participants assessed as acceptable, although they would not have wanted it in the game. Nonetheless, all agreed in the interview that they did not have a better idea for selecting features and that the voting system was reasonable. Three participants also explicitly mentioned that it was better than just picking a random idea (as was done in the beginning of the study). All also elaborated on their feeling that they did not think that the idea selection was “unfair”, i.e., it appeared to them that everyone else had the same number of ideas realized, without anyone being clearly favored. One participant raised concerns about potential trolling tendencies in a large community.

![Figure 5. Multiple paths per level: red is faster, but also more difficult. Blue is easier because of the higher jumps power-up, but slower.](image)
Participants reported that they rarely used the ideas filter; two did not even recognize that they could see the “highest” rated one by using it. We also asked why they did not comment on ideas (as only one comment was written overall), and two participants stated that comments were not necessary as ideas were clear enough. They provided mixed thoughts on the community-driven game design of Crowdjump: one participant raised concerns about the initial state of the game and that such a process should be adopted only later in the design process, after more features of the game were already available. Three raised the question of what would happen in a larger community (again, highlighting trolling tendencies and “non-optimal choices” by the community) and demanded moderation options for the developers or experts in the crowd.

Lessons Learned

We learned that the player-driven game development idea of please be nice (; could be investigated within a study setting (effort- and time-wise). The sample was not only quite content with the features that were realized, but they could imagine continuing their participation, even after the 15 days. For the main study, we decided to start directly with the voting feature implemented. There were two reasons for this: first, it allowed us to see how the community wanted to evolve the game right from the beginning (through their voting), instead of waiting until the voting feature was suggested and picked randomly. Second, every participant in the pilot study highlighted that the voting was a reasonable approach for feature selection.

As we had seen that many of the feature suggestions could be implemented in a couple of hours, we decided to aim for two feature realizations per day in the main study. In addition, a larger user base should help to get a broader spectrum of suggestions. As many of the ideas were only game related, we also decided to highlight during idea submission that features for other areas (voting, web page, ...) could be suggested as well. Based on the increased difficulty, and as stated above, for the main study we decided to more carefully design the levels, to allow more people to finish the game. We also added a headline to the sorting/filter area to make it more prominent.

MAIN STUDY

The goal of the study was to investigate how Crowdjump is perceived, to see which ideas were suggested and to get insights into how players shape the game design.

METHOD

The game feature state was reset, i.e., the study started as described in the Crowdjump section. Players did not receive any compensation for their participation. We used the same method as in the pilot, but changed the frequency of selecting ideas (one was picked at 9pm, one at 7pm). Every day at 7pm we also released a new version of Crowdjump, with the features that were ready at that point. The idea voting based on up- and down-votes was available directly from the beginning (i.e., the players alone decided which new features were implemented) and in the case of a draw, we took the idea of a user who had fewer ideas realized so far. We informed all participants via email when a new release was available, and included screenshots if significant visual changes had occurred. Based on [11] we also decided to exchange the GEQ and SPGQ with the GAMEX [6]. Furthermore, we added statement-based questions (5-point scales, similar to the other questions) on the web page/game performance, the amount of bugs players encountered and the idea realization frequency to rule out negative perceptions being due to such issues. The study ran for 23 days. We recruited the participants via emails and over social media, and people could join the study any time. Participants from the pilot study did not participate.

PARTICIPANTS

25 users interacted with Crowdjump, but 13 of them were only active on up to 5 days (7 on only one day) and they did not fill out the closing questionnaire. Thus, we will primarily focus on the other twelve participants (1 female; 18-26: 8x, 27-39: 3x, >40: 1x), which we call “active participants”2. One participant reported using a PC 5–10 hours a week, two 10–15 hours and the rest more than 25 hours. While ten participants reported to have designed an application, only one also had designed a video game, and four participants stated that they would like to be integrated in the design process of a video game. The sample reported mixed responses for the statement that they like platform games (7-point scale, M=4.2, SD=1.1, Mdn=4) and no one knew about please be nice (; already.

Usage of Crowdjump

Overall, the game was played 12988 times (12470 times by the active participants), 138 ideas were suggested of which 43 were implemented and 479 up- and 121 down-votes were cast. Three ideas (sophisticated map creation tool, 2x multiplayer setup on the same maps in which players could actively hinder each other) were assessed as “not realizable” in the study as their estimated implementation time was too long for the study goals, four ideas were duplicates of other ideas and three ideas requested a feature which was already implemented.

Table 3 shows an overview of the active participants, similarly to Table 1. We saw a difference in behavior for those who provided a score below 5 regarding the statement that they like platform games and those that provided a score above 5 (i.e., those who like platform games): the former group submitted more ideas (M=12.3, SD=11.6, Mdn=9 vs. M=7.4, SD=6.7, Mdn=6); test statistic U=10, standardized test statistic z=1.22, p=.222, effect size r=.35), while the latter played more games (M=316, SD=342, Mdn=139 vs. M=2050, SD=2598, Mdn=808; p=.211, U=30, z=-2.030, p<.05, r=.58), i.e., Crowdjump was used differently. Although the descriptive numbers show a large difference, only the number of games played was significantly different according to a Mann-Whitney U test. Concerning our estimates for realizing ideas, 17x estimation and realization time were equal, 19x the estimation was too conservative, and 7x a realization needed longer. Overall, we estimated 137 hours for the selected ideas, but only needed 106 hours. From a version release point of view, on 19 of the 23 days, the released version contained the intended number of features (one for day 1, and two for all other days); on two days just one feature was integrated, but the day after the version contained three features to compensate for this.

2As ideas of non-active participants were selected, we need to consider all participants for the idea-related results.
As the pilot study already indicated that no community was built, and besides up- and down-voting not much was done to improve individual ideas as a community, we analyzed the answers to the statement in this respect: the sample tended to disagree that they formed a community with other players (M=2, SD=1, Mdn=2) or that they worked as a team (M=2.3, SD=1.1, Mdn=2). Although a comment feature for ideas was implemented at day ten of 23, only 25 comments were posted. As will be illustrated below, only a low number of other community-related ideas were suggested. The participants did not have the feeling that others interfered with the development (M=2.3, SD=1.1, Mdn=2) and they disagreed with the statement that their opinion was not heard (M=1.9, SD=1, Mdn=2). Taken together this leads to R2: The players were not interested in altering/developing ideas as a community and were content with acting individually.

Evolution of Crowdjump

Figure 2 shows example screenshots of versions during the study to illustrate how the game was changed visually. In the supplementary material, we also uploaded a video that showcases all versions. While the first version could be played through (again) in five seconds, the record for the final version of the game was above 3 minutes. In the first three versions of the game 84% of all tries were successful, in the last three only 6% (which coincides with submitted ideas that aimed specifically to make the game harder; see below).

Idea Length

The longest suggestion description had 495 characters (including spaces) with the title "Voting time stamp":

There are some really nice ideas I would like, but I can't vote them at the moment, because I find other things more important. Things like forgotten pw or design. Would be great if u could give your votes priorities. Maybe shh like u vote everything and all positive votes u can watch and sort by your own priority. The first x (maybe 5) become a vote up and the other not. When one of the top 5 is done, the sixth goes up and becomes the vote up. Or u just make the option "reminder nice idea"

and the shortest one only 18 (and had the title "Double jump"):

Maybe as a powerup

The average character count (including spaces) of selected ideas was 90.4 (SD=69.9, Mdn=66), of non-selected ideas 143.76 (SD=112.6, Mdn=113.5). This difference was significant as a Mann-Whitney U test showed (U=1383.5, z=-3.127, p<.01, r=.27). This leads to R3: Ideas expressed in fewer characters were more likely to be selected by the community. Whether or not the length also has an impact on the idea quality was not investigated within this study.

Table 3. Overview of active participants; see caption of Table 1. In addition, the numbers of up- and down-votes (Votes) are shown.

<table>
<thead>
<tr>
<th>ID</th>
<th>Suggested</th>
<th>Chosen</th>
<th>Time (day)</th>
<th>Tries</th>
<th>PP (day)</th>
<th>Up</th>
<th>Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>37</td>
<td>8</td>
<td>1985</td>
<td>435</td>
<td>2</td>
<td>75</td>
<td>20</td>
</tr>
<tr>
<td>02</td>
<td>8</td>
<td>7</td>
<td>484</td>
<td>126</td>
<td>4</td>
<td>41</td>
<td>21</td>
</tr>
<tr>
<td>03</td>
<td>13</td>
<td>3</td>
<td>1619</td>
<td>370</td>
<td>3</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>04</td>
<td>8</td>
<td>3</td>
<td>710</td>
<td>139</td>
<td>3</td>
<td>49</td>
<td>11</td>
</tr>
<tr>
<td>05</td>
<td>19</td>
<td>6</td>
<td>373</td>
<td>808</td>
<td>5</td>
<td>52</td>
<td>24</td>
</tr>
<tr>
<td>06</td>
<td>2</td>
<td>1</td>
<td>1285</td>
<td>737</td>
<td>5</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>07</td>
<td>6</td>
<td>2</td>
<td>1035</td>
<td>150</td>
<td>5</td>
<td>46</td>
<td>11</td>
</tr>
<tr>
<td>08</td>
<td>6</td>
<td>4</td>
<td>144</td>
<td>2022</td>
<td>6</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td>09</td>
<td>0</td>
<td>0</td>
<td>358</td>
<td>1017</td>
<td>4</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>0</td>
<td>1530</td>
<td>6536</td>
<td>5</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>3</td>
<td>631</td>
<td>31</td>
<td>4</td>
<td>34</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>3</td>
<td>88</td>
<td>99</td>
<td>4</td>
<td>29</td>
<td>5</td>
</tr>
</tbody>
</table>

Suggested Chosen Time PP (day) Votes

Perception of Crowdjump

Similar to the pilot study, Crowdjump was perceived quite positively. The SUS score was on average 80 (SD=12, Mdn=82.5), indicating a good to excellent usability [2]. In the closing questionnaire, the active participants agreed that the performance of the game/page was good (5-point scale, M=4.3, SD=.7, Mdn=4) and they disagreed with the statement that there were too many bugs in the game (M=2.1, SD=.7, Mdn=2).

The active participants also agreed that they liked the idea of Crowdjump (Mdn=4.5, SD=.7, Mdn=5). On the enjoyment sub-scale of the GAMEX, the sample scored on average 5.4 (averaged from six 7-point scales; SD=1, Mdn=5.6) also showing a clear tendency that the experience was enjoyable. Again, by comparing those providing a lower (<5) score on the genre preference statement (answers to Crowdjump idea: M=4.6, SD=.5, Mdn=5/GAMEX enjoyment scale: M=5.4, SD=.1, Mdn=5.5) with those that provided a higher score (>4; M=4.4, SD=.9, Mdn=5/M=5.5, SD=1.1, Mdn=5.8), no significant difference between these two groups could be found (Mann-Whitney test, U=16.5, z=1.85, p=.05/U=19, z=.244, p=.807, r=.07), i.e., it seems that the a priori preference for the genre did not affect the enjoyment of Crowdjump.

The sample tended to like being able to submit new ideas (M=3.4, SD=.9, Mdn=3.5) and clearly agreed with the statement that the game (M=4.8, SD=,4, Mdn=5), the web page (M=4.5, SD=.5, Mdn=4.5) and the idea selection process (M=3.9, SD=.8, Mdn=4) were developed in a positive direction. The participants also clearly agreed with the statement that the features were realized as requested (M=4.6, SD=.5, Mdn=5), but less so with the statement that the implemented features met their wishes for Crowdjump (M=3.5, SD=.5, Mdn=4). Taking all this together, this leads to the result R1: A platform game that allows its players to directly influence the features in short release cycles is enjoyable.

As the pilot study already indicated that no community was built, and besides up- and down-voting not much was done to improve individual ideas as a community, we analyzed the answers to the statement in this respect: the sample tended to disagree that they formed a community with other players (M=2, SD=1, Mdn=2) or that they worked as a team (M=2.3, SD=1.1, Mdn=2). Although a comment feature for ideas was implemented at day ten of 23, only 25 comments were posted. As will be illustrated below, only a low number of other community-related ideas were suggested. The participants did not have the feeling that others interfered with the development (M=2.3, SD=1.1, Mdn=2) and they disagreed with the statement that their opinion was not heard (M=1.9, SD=1, Mdn=2). Taken together this leads to R2: The players were not interested in altering/developing ideas as a community and were content with acting individually.
An active participant cast 34.5 up- and 9.75 down-votes on average (see also Table 3). While the sample disagreed that they preferred new ideas over old ones (M=2.1, SD=.8, Md=2), nearly all votes were cast on newer versions (which were shown first on the page), although we offered search, filter and sorting options for ideas. Of the 600 up- and down-votes, only 30 (5%) were provided for ideas that were older than one day. In addition, only seven old ideas were deleted by the suggesting users. Taken together, this leads to R4: Recent ideas were more interesting for the community than browsing through old ideas to find suitable ones for the current version. Selected (non-selected) ideas had on average 5.5 (2.6) up- and 0.2 (1.2) down-votes and in 110 cases the submitting user also up-voted his/her own idea (and users never down-voted their own ideas). On average, the time between an idea being submitted, selected and implemented was 4.6 days. Given that old ideas (even those already having many up-votes) were not browsed by players who had not yet considered them or new players (according to R4), moderately up-voted ideas thus had to wait until ideas with even more up-votes were implemented first, explaining this delay.

Suggested Ideas
Three people independently categorized every provided idea with one or more categories and deviations were solved through discussions. The categories were Game (covering ideas on game design, game play, game controls, level design and story aspects), Web page (covering ideas aiming at altering the web page on which Crowdjump was offered), Social (covering ideas on social features that connect one player to other players), Game Aesthetics (covering ideas on audio or visual changes), and Ideas (covering changes to how ideas can be submitted or selected).

- **Game**: 73 ideas of the 138 suggested (52.9%) were assigned this category. Ideas covered the addition of collectibles (e.g., coins and power-ups for the avatar), static (e.g., spikes) and movable obstacles (e.g., enemies), new platform types affecting the game play (e.g., ice platforms), the level design (e.g., more and larger levels), alteration of how the time per level is counted, adjustments to the difficulty level, new game mechanics (e.g., teleporters or pressure plates to open doors), fixes to existing features (e.g., fixing some collision detections), controls (e.g., being able to control the character with other keys) and adding capabilities for the avatar (e.g., abilities). Of the 43 implemented ideas, 24 were in this category (55.8%).

- **Game Aesthetics**: 35 ideas (25.4%) were assigned this category. Ideas included altering/extending the sound scape (e.g., adding background music), altering the available game assets (e.g., adding background images), adding animations (e.g., coins should spin), adding game assets (e.g., adding non-game-play-related assets such as trees) and allowing for customizations (e.g., selecting the avatar’s appearance). 11 implemented ideas fell in this category (25.6%).

- **Web page**: 35 ideas (25.4%) were assigned this category. Ideas included adding a high score page on the web site, a profile page with all scores reached in the versions, changes to existing pages (e.g., altering the button appearances for up- and down-voting), adding a password recovery function, adding new communication means (e.g., a forum), adding gamification elements to the web page (e.g., a badge for the best player visible to everyone or an inventory for collectibles that you gain after completing a level), sending e-mails for different incidents (e.g., as a reminder or if a player reaches a new top score), and adding user-generated features (e.g., collaboratively, describing how levels should look). 8 implemented ideas were in this category (18.6%).

- **Social**: 19 ideas (13.8%) were assigned this category. Ideas included multiplayer options (e.g., playing on the same map or being able to sabotage runs of other players), community-generated content (e.g., a user builds a map and provides it to all other players), communication options (e.g., a forum or comments that can be posted for idea suggestions) and options to asynchronously connect to others (e.g., being able to congratulate the top player by sending an email). 3 implemented ideas were in this category (7%).

- **Ideas**: 4 ideas (2.9%) were assigned this category. Two ideas aimed at improving the existing voting system: the space limit (which was at 500 characters) should be removed and more search options implemented (e.g., search by user name, although this was already available on the page). The two other ideas acknowledged the large amount of ideas and suggested changes: the first idea was to add categories to the ideas to make it easier to keep track of suggestions. The second idea stated that there are many ideas that seem more important than other interesting, but not so relevant ideas. Therefore, a voting priority was suggested, i.e. five up-votes could be spent, and all other up-votes of this user afterward only become active when one previously up-voted idea was implemented. None of these ideas were implemented (0%).

Selected Ideas
How Crowdjump was changed over the 23 days, i.e., which ideas were chosen by the player base and when they were released, will now be described in detail (see the supplementary material for a video showcasing all changes).
Early-study changes (Day 1-Day 8): Of the 13 features realized in this time span, only two were not directly game-related and covered aspects of the ideas page (day 6 and 7: improve the title representation and make the up-vote and down-vote buttons visually more vivid). The first realized idea (day 2) made the game fairer by only counting the time after the first keystroke, followed by a high score list (which would start empty each release) to compare and give the timings meaning (day 3). As the players did not have of much of a chance to show their skill on the leaderboard with just one level, it was not surprising that the next feature was the demand for bigger levels with secrets to be found (day 4). A similar aspect was introduced (on the same day) through the option to collect coins. The collected secrets and coins were shown on the high score list as well. Also on day 4, a feature to mute the sounds of the game was selected. On day 5, an idea was chosen to allow the avatar in the game to be visually changed: the players could select one of several pre-defined assets or could upload their own avatar asset. On day 6, the game was made more challenging by adding lava and obstacles on the maps that would kill the avatar. Additionally, an idea to add power-ups was selected: when consumed, higher jumps or moving through walls became possible. On day 7, enemies (that could be killed by jumping on them) that move in pre-defined patterns on the platforms were added, further increasing the difficulty. On day 8, an implemented idea was extended: now collecting coins also reduced the time score in a level, i.e., a player was further rewarded to collect coins. Finally, an idea to also be able to control the character with the W, A, S, D keys was realized.

Mid-study changes (Day 9-Day 15): On day 9, background music was added and the white background was exchanged with a colorful background image, so both ideas aimed at aesthetic changes. On day 10, two more levels were released (when dead, a player needed to re-start at the first level), and the option to comment on ideas. The day after, a badge for the best players was introduced which was always visible to all players, showing the user’s name. The game was made more difficult, with the idea to add cannons that shoot projectiles through the levels. On day 12, pressure plates were added that open gates in the levels. In addition, another idea realized on that day aimed at having longer levels so that the impact of the coins on the time could not lead to negative scores anymore. On day 13, more lives for the avatar were realized, i.e., a death would only respawn the character in the current level and another aesthetic idea was selected: the coins should spin. Another power-up (slow-down for moving obstacles) and the option to practice a level after it was completed once was added on day 14. A further aesthetic idea was realized on day 15, by adding decorative, non-game-play-related assets (e.g., trees) in the levels, and destroying obstacles was rewarded with coins that spawn at the position of the obstacle.

Late-study changes (Day 16-Day 23): Day 16 brought no new game mechanics, but a day and night mode, i.e., the background changed depending on the day time (a pure aesthetic change) and a FAQ to the game elements was added below the game window. An interesting event happened on day 17: one realized idea added two further levels. The user who submitted this idea explicitly stated that the game is “too easy” and the new levels should make it harder. The second chosen idea on this day was the introduction of multiple difficulty levels to select from. Here, the creator of this idea stated that the game was “too hard”. This indicates that a heterogeneous community can lead to contradicting ideas, but that the player-driven approach can also lead to solutions for this. On day 18, an aesthetic change was implemented by animating the lava obstacle and a power-up (the double jump) was added. Day 19 added a “back to menu” button when the player died and the platforms were graphically adapted (again, a purely aesthetic change). On day 20, a question-mark box was added: when a player jumped into it, coins spawned. Again, earlier idea was further adapted: now the option to practice a level was added as soon as a player reached that level (instead of only being able to practice it after it was completed once). One idea on day 21 added statistics to the web page (e.g., how many times a level was tried by the player base) and the other idea demanded a graphical change for the flag that needed to be reached (as it was apparently hard to recognize given the other aesthetic changes). Day 22 added a teleporter as a new mechanic and an aesthetic change, as power-ups were now also animated and explode when collected. The final two features (day 23) added a bug report page and a permanently saved setting for whether or not the game was muted.

Considering the suggested and selected ideas, we can derive two results: R5: A voting mechanism based on up- and down-votes is a proper decision tool in player-driven game design, or players have no better idea for decision finding, and R6: To improve the game with features is more important than peripheral aspects. R5 can be derived as only four suggestions for the idea creation/selection were provided (not counting the purely aesthetic changes to the ideas page). These suggestions, though, would not have changed the nature of the voting scheme. R6 can be seen by the number of selected ideas considering the game or game aesthetics (35) in contrast to the sum of all other feature categories (11). Further support for R2 can be derived, as the number of social suggestions (19) and selections (3) were not particularly high either.

DISCUSSION

A player-driven game design with frequent release cycles (in our case, daily) was seen as enjoyable, although many participants reported that they did not particularly like the genre of platform games (R1). This is an encouraging result raising the question whether the autonomy through the option to decide on features is responsible for this; perhaps the game became more appealing based on the feeling of ownership, or the resulting game might simply be more appealing than a typical platform game, based on the player-base adding the features that they liked. This is a question to be targeted in upcoming studies: for example, we could give a new participant group access to the web page and the game, but without the option to add ideas. Instead, we would keep the daily release cycle and provide the same features (on the same days) as during this study, and compare the results. It will be also interesting to investigate whether the generation of ideas alone (with developers selecting the ideas to be realized) or the selection of these (with developers providing ideas alone) would be (similarly) enjoyable for the players.
Another important aspect was that the participants were content with how we realized their ideas, although many were only formulated in a coarse-grained way and thus it was up to the developer to decide how to actually implement them. By comparing the realized pilot study ideas with those of the main study, we see overlap, suggesting that games/genres have specific features that they need to have realized. Finding out which features are directly associated with a genre in such an open design task is also promising for future research.

In light of the three ideas that we marked as “not realizable” in time (within the goal to release two features per day) a question of scalability arises. Larger development teams would have had more resources and could potentially target such ideas more easily. We thus hypothesize that scalability on the developers’ side is not an issue and would be advantageous for player-driven approaches. In contrast, the amount of feature suggestions might increase if more players were active (and was already been to be slightly overwhelming for participants in our study). While we think that the player base would then suggest features to mitigate this effect (in a self-regulatory way), this study did not allow us to draw conclusions about this yet (see limitations below). Through the voting aggregation of opinions, decision-making seems possible independent of the actual player base size, but it remains to be shown whether self-regulatory ideas would be decided on to enable handling the potential larger amount of ideas properly as well.

It was interesting to see that no community feeling arose (R2). While some of the suggested ideas involved other players (e.g., the high score, the badge for the best player, or the wish for a multiplayer option), features that would lead to a mutual, direct collaboration to change Crowdjump were not often suggested nor picked by the user base. The comment feature, which can be seen as one that would have allowed for this, was implemented, but not often used. Considering R5, a potential explanation for this is that through the voting system an user could already express his/her interest in an easy way and that adding comments to other ideas might be too tedious. The latter explanation also receives support from R3 and R4, as mainly shorter ideas were chosen and players were not interested in browsing through old ideas to either delete (if not suitable anymore) or re-consider them, given the new state of the game. The results also suggest that the frequency of the releases and the number of realized ideas was sufficient for the player base, i.e., realizing more features is not necessary in upcoming studies. R5 in itself requires future work to find out whether players are content with the voting scheme, or have only limited design ideas as they are not game designers and simply have no better approach for decision-making. This would also have implications for other suggestions as well.

Not surprisingly, although every aspect of the Crowdjump web page could be altered, the players focused on the game (R6). Here, an explanation could be that people knew that if they were to up-vote another, non-game-related idea, the progress of the game development would be slowed down. In a future study, it might be advisable to test an altered voting scheme, with multiple “buckets”: instead of taking the two highest-rated ideas, developers could always choose the highest game-related and non-game-related ones to see how the peripheral features change. Additionally, we also saw that the player base adapted to the game continuously and considered the new state for their ideas. For example, after the coins were added to the game, a later realized feature connected the coin collections with the time score (day 8). On day 12, longer levels were added to mitigate the fact that a good player’s score can become negative through this. Similarly, we saw a non-uniformity of the player base at day 17: for some players the game was too hard, for others too easy. Nonetheless, the player-driven approach also allows players to regulate these aspects, although it will be interesting to see what would happen if there is a clear dominance for one solution. Upcoming studies can investigate this by simulating a player base (in a Wizard-of-Oz [4] style) selecting contradicting features.

The conducted study had limitations. First, the low number of participants and the gender bias towards men might be an issue. Also, the selection of our participants covered mainly IT-affine participants with a background in programming. Whether or not this had an influence also cannot be derived from this study, and thus a sample with a different background should be considered next as well. In addition, we only considered the genre of platform games. How the player-driven approach would unfold in other genres thus also cannot be determined through this study. It must also be noted that we had players that started to interact with the web page but were not very active and did not fill out the final questionnaire. Maybe, for them, the scenario was less interesting or they were simply not motivated to spend their time in an unpaid experiment. Thus, the enjoyment ratings might be biased. Finally, due to the short study runtime, some of the results might also be explained by novelty effects, requiring a longer follow-up study.

CONCLUSION

In this paper, we presented our first exploration of a player-driven platform game, called Crowdjump. Within daily release cycles, features were implemented that were suggested and selected through up- and down-voting by all players. We found that players (even those that do not like platform games in particular) enjoyed Crowdjump, although no feeling of being a community arose. While players used up- and down-voting for ideas (in particular for game-related aspects), other community-based features such as providing comments on ideas were not often used, indicating that this was less interesting to them. With its study, this paper contributed a first set of results that needed to be considered for such player-driven design approaches and several opportunities for future studies, as elaborated in the last section.

Working with such participatory design approaches appears interesting to allow players to shape the future of “their” games, and especially allow those who have neither time nor the required skill for modding to make an impact. For developers, such approaches can also be useful to learn what players find important in a particular genre.

In summary, it seems worthwhile to further explore this paradigm with the eventual goal to derive concepts and guidelines on how to administer, carry out and use player-driven game design approaches.
REFERENCES


