Rapid game jams with fluidic games: A user study & design methodology

Gaudl, Swen

http://hdl.handle.net/10026.1/15087

10.1016/j.entcom.2018.02.007
Entertainment Computing
Elsevier

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.
Rapid Game Jams with Fluidic Games: A User Study & Design Methodology

Swen E. Gaudl\textsuperscript{a,}, Mark J. Nelson\textsuperscript{a,}, Simon Colton\textsuperscript{a}, Rob Saunders\textsuperscript{a}, Edward J. Powley\textsuperscript{a}, Blanca Perez Ferrer\textsuperscript{a}, Peter Ivey\textsuperscript{a}, Michael Cook\textsuperscript{a}

\textsuperscript{a}Falmouth University, Games Academy, Penryn Campus, Treliever Road, Penryn TR10 9FE, UK

Abstract

We introduce rapid game jams, a style of game jam that takes only 1–2 hours and is focused on design experimentation rather than on programming and technical implementation. To support that kind of rapid game-design experimentation, we have designed a class of games that we call \textit{fluidic games}. These are mobile games in which the game mechanics and other aspects of the games are editable on the fly, directly on the device, allowing for frequent play/design context shifts. We have conducted four rapid game jams with 105 participants from a local Girlguiding organisation, in order to gain real-world experience with this concept. We analyse results from a survey instrument completed by 69 participants in two of these rapid game jams. In order to guide future work in addressing questions left open by this study, we did a qualitative analysis of the designed games to gain additional insights into participants’ design practice.

Keywords: Game Design, User Study, Game Jams, Digital Games, Mobile Games, Design Tools

1. Introduction

Game jams are creative, social events in which participants, usually working in teams and physically colocated in a space with other jam participants, rapidly produce a game. Typical game jams last 24-72 hours; for example, the Global Game Jam, an annual event in which hundreds of locations around the world simultaneously host game jams, lasts 48 hours \cite{1}. Kultima summarises them as “accelerated, constrained and opportunistic game creation events with public exposure” \cite{2}.

\textsuperscript{*}Corresponding author

Email addresses: swen.gaudl@gmail.com (Swen E. Gaudl), mjn@anadrome.org (Mark J. Nelson), s.colton@gold.ac.uk (Simon Colton), Rob.Saunders@falmouth.ac.uk (Rob Saunders), edward.powley@falmouth.ac.uk (Edward J. Powley), blanca.perezferrer@falmouth.ac.uk (Blanca Perez Ferrer), peter.ivey@falmouth.ac.uk (Peter Ivey), michael.cook@falmouth.ac.uk (Michael Cook)
We have been experimenting with a much more severely time-constrained style of game jam, which we call rapid game jam. These last only 90 minutes in the events we’ve held so far. That time is split into: 30 minutes of familiarisation with a design space, 30 minutes of rapidly experimenting with multiple design concepts in that space, and 30 minutes of sharing and critiquing designs. In order to support such rapid experimentation, we have developed the notion of fluidic games, mobile apps that embody parameterised design spaces (on iPhone and iPad currently). Fluidic games come with a number of regular, playable games from a given design space, which players can play initially to understand the possibilities of games within the space. All of these games can however be edited to produce other games in the design space; in addition, there are features to allow new games in the space to be automatically generated for brainstorming purposes.

We have held four rapid game jams thus far, with 105 participants from Girlguiding Cornwall, who visited Falmouth University as part of a larger Girls Can Code event. We analyse quantitative survey results from two of these game jams, with 69 participants, in order to understand their experiences and refine how we run rapid game jams. We furthermore collected the games that were produced during these game jams, and present a qualitative analysis of the design choices.

2. Game Jams

Rapid game jamming on mobile devices takes inspiration from traditional game jams, but compresses the experience into a much shorter period of time, aimed at enabling design by end users who may not think of themselves as game developers – not even necessarily as aspiring indie game developers. In addition to being tightly time constrained, our rapid game jams focus more on free-form design experimentation, with participants encouraged to explore many different designs, rather than on prototyping an implementation of a specific game idea.

Traditional game jams last between 24 and 72 hours, often over the course of a weekend. In addition to providing a forum for prototyping game ideas, they serve a community-building role in the indie game community, somewhat reminiscent of the role LAN parties serve among gamers, but focused on developing new games rather than playing existing ones. Game jams are often promoted as a forum for free-form experimentation that is quite different from traditional, more methodical and constrained game-design processes, such as those used at large companies, or even at smaller indie studios during “normal”, non-game-jam development periods.

However, while game jams are indeed open ended relative to those reference points, what actually takes place at jams is still largely software development, albeit rapid, messy software development along the lines of a hack session. Teams participating in game jams generally start with an idea that, after some initial refinement, is implemented over the course of the jam, without major changes in design direction. The concept may be scaled down or modified if aspects turn out to be infeasible, and some directions may be pursued opportunistically, but the time is mainly spent on producing a working prototype of that idea, rather than more free-form design experimentation.

Due to the large proportion of time at typical game jams spent implementing a prototype, the experimentation that takes place is therefore largely of the style described
as opportunistic programming, a form of design experimentation focused on improvisation within the context of technical implementation [9]. This is understandable, since producing working software is hard; much of the advice to game jam participants focuses on how to simply finish a playable game in the allotted period of time [10]. The results are interesting in their own right, but this does put the focus of game jams more on technical implementations of specific game prototypes, rather than brainstorming or playing with design spaces. This code-first style of design experimentation therefore emphasises specific facets of the much larger set of techniques that go into game design in its full breadth.

The starting question for our research is whether the new form of hybrid mobile game and game-design app described here, dubbed a fluidic game, can enable a new style of game jam that shakes up these norms. We would like to both make game jams much more rapid (hours rather than days) and shift their emphasis to design experimentation that foregrounds aspects of game design such as systems thinking, aesthetics and balance, rather than coding and technical implementation. This form of faster, less implementation-focused design experimentation is intended to retain some of the advantages of a game-jam style setting, but shift the kind of design that takes place. Such a shift brings us somewhat in the direction of existing approaches such as brainstorming sessions and low-fi design, which focus on ideation rather than development [11], but importantly, our goal is that the output of the process should be fully playable digital games, which is the case with game jams but not with brainstorming/ideation approaches. That is possible with the new approach described here.

3. Fluidic Games

There are many video games which could be described as maker-games, where some of the appeal comes from creating new game levels. One of the well known console games of this type – where players have made millions of levels – is Little Big Planet (Media Molecule, 2008). On hand-held devices, which is our focus, there are a number of maker-games, including Big Bang Racing (Traplight, 2016) where players have created and shared millions of tracks for this driving game. While the options for customisation of levels in such games can be extensive, they usually extend only to the content and layout of levels. That is, game elements such as the scoring, spawning, killing or progress mechanisms, or aspects of the underlying physics simulations are not exposed for players to experiment with.

We are interested in the development of maker-games which empower the creation of casual games with genuinely new game mechanics, but which require no programming, and can be carried out directly on the target device (i.e., a mobile phone or tablet). Our motivation here is to lower the barrier to entry to game making, to levels enjoyed by other creative domains, such as drawing/painting, composing music or writing stories. To this end, we are investigating what we call fluidic games [12][13][14], where the line between game play and game design is blurred as in other maker-games. With fluidic games, substantially more control over the game world is afforded to the player and novel game mechanics can be designed directly or be found emerging from the game setup and exploited accordingly.
We have preliminarily identified the following requirements for such apps:

- It should be possible to make meaningful edits to games in seconds, including the time taken to navigate the design UI as required and the re-starting of the game being created;
- Making meaningful edits should be as much fun as playing the game in terms of user-experience;
- It should be possible to see how game edits have altered the gameplay in near-realtime;
- The app should provide inspiration for players who need help getting started with an idea for a game;
- The app should provide a clean-slate option (a bare bones game template) for people who have a specific design in mind.

3.1. Designing fluidic games

Since fluidic games are a designed space of games, but also intended to enable end users themselves to design games [15], they fall into the category of designing for design-after-design [16] [17]. This is a catchy, if unwieldy, name for a broad class of approaches that focus on designing open-ended, even “unfinished” systems that enable the systems’ users to themselves continue the design process after the initial design of the product is ostensibly finished and the product shipped. This approach arose out of the older field of participatory design, but rather than focusing on involving users in the design process up front, as in classic participatory design, design-after-design focuses on building systems where the design process is not closed when the system is done from the initial designer’s point of view – instead the system is designed so that “there is design (in use) after design (in the design project)” [16].

The design-after-design approach is more of a design philosophy than a specific implementation approach. Thus, in order to build fluidic games, we draw on two more specific fields of research: our technical approach for constructing fluidic games is parametric design [18], and our user-interaction style is that of casual creators [19]. Parametric design [18] is a well-established technique, in which sets of possible solutions to design problems are defined as explicitly parameterised design spaces, i.e., multi-dimensional spaces defined by a number of parameters that can be varied independently or jointly. Individual designs are then specific points in this design space. The parametric design approach gives rise to natural metaphors for exploring design spaces such as navigation of the space, where one can travel from design to design by varying parameters. It also provides an underlying framework for automated or semi-automated design, as suitable parameters (for some design goal) can be searched for using a variety of artificial intelligence or optimisation methods [20] [21] [22]. In recent years, it has become so commonly used in architecture, for example, where parametric design methods are incorporated into standard CAD tools, that “parametricism” has been called the dominant contemporary architectural practice [23].
A casual creator is “an interactive system that encourages the fast, confident, and pleasurable exploration of a possibility space”, aimed at supporting autotelic creativity rather than supporting task completion [19]. Key to a casual creator for games is that it should be enjoyable to explore the design space just as it’s enjoyable to play games within that space, with easy switching between those modes. Of the casual-creator design patterns Compton and Mateas identify in [19], we focus most on limiting actions to encourage exploration and saving and sharing. We give players the ability to change anything within a limited parametric design space, and save and share the results (as well as modify games shared by others).

There is admittedly some tension here in balancing the enjoyable, casual user-interaction style of casual creators that is often gained by radically limiting the design space, with the design-after-design approach, which argues that end users should be able to reconfigure the design in quite open-ended ways qualitatively different from what the original designers intended.

Understanding this tension between design freedom and enjoyability of the experience and how it impacts the creative process of designing is the key point we want to investigate in this article because in a parametric design setting, the challenge in balancing these goals is to identify a parametric space that can be made fluid and fun to explore through UI mappings and feedback on design changes, but large enough that it contains surprising designs that are quite different from the games initially included with the app as exemplars.

3.2. Wevva, a fluidic game

Wevva is the first fluidic game we have produced and released. It comprises a design space of 2d, physics-based games for the iOS platform. There is a loose theme based around the weather, as games are set in outdoor scenes and the characters – which are the kind of things found in the sky, such as insects, animals, flora and vehicles such as helicopters – are blown about by the weather. Players interact with the game through tapping the characters, or by dragging a controller which is either another character or a kind of shape for collecting/repelling the characters. An example controller is the shape in Figure 1(a).

Wevva originated as a subspace of a much larger parametric space of games that can be represented with our underlying parametric game technology platform, Gamika. The Gamika platform exposes over 180 parameters representing game components and game-engine parameters, as well as graphical editing tools and more. The space of games afforded by Wevva was derived by cutting back the game-editing options of our in-house research prototype to target a more manageable and coherent design space.

Wevva contains a subset of the full set of Gamika parameters, both to reduce the cognitive load [24] when thinking about design problems as navigation of a large parameter space, as well as to allow us to design an easy-to-use, natural interface for players/designers to be able to easily edit games on mobile devices (the full Gamika

---

1The Wevva app used for our study was an alpha version and differs from the release version in terms of interface and also a smaller subset of parameters constituting the design space. However, most of the games designed during our study can be reproduced in the release version.
Figure 1: (a) An in-game screenshot of the user-designed game "Snowy Terff" made with Wevva (b) The design screen of the app version used by the participants during the game jams. The design screen is encoding the essential design space parameters to define the game shown in (a). (c) The sub-menu for defining what happens when players tap on one of the game elements, in this case the snowflake (d) The controller sub-menu showing different controller types such as the inverse bucket visible in (a). The controller is a game object the player can move around on screen using touches (e) Players can pick a general background for their game, select a music track, control the speed of movement of game all game objects and change the AI play-support (a feature which was disabled for the studies).

editor also runs on mobile devices, but is more of a power-user tool, full of hundreds of sliders and options. At the same time, the app should still enable users to make qualitatively different games, not just minor tweaks on built-in games, so the parameter space exposed by Wevva still allows quite divergent games to be designed. For the game jams conducted in this article, we used a version of the app which is presented in Figure 1. This version contains only two characters – snow and rain – along with a set of 81 different controller shapes and a variety of scoring, win-condition, and physics parameters.

Figure 1(a) shows a screenshot of one of the games designed during one of our game jams. In this game the goal is to move the controller on the screen in a way to increase the number of collisions snowflakes while minimizing the collisions of raindrops. Those rules are described in (b) through the design screen, which captures all crucial parameters in a set of nine categories on a single screen. This was designed to allow the designer/player to grasp all explicit rules of a game without the need to switch into a sub-menu. In the top row, the three categories describe the interaction between the two character types, namely what happens if two or more snowflakes, raindrops, or snowflake-raindrop collide. The first menu from the left on the middle row opens up menu (c) which allows the user to define what happens when the player taps on one of the characters. The options in this category include exploding the tapped character, switching character type and more. The remaining two categories in the middle row allow the designer to specify the size and amount of characters from any of character types. The last row contains specifying the controller, such as the ⊓, adding noise to the movement and setting as the first category from the right the spawning and scoring

The relationship between the full Gamika design space and in-house app, and the more polished fluidic games such as Wevva that are based on Gamika subspaces, is discussed further in [14, 12, 13].
zones. The controller category in the used version contains a set of grid controllers of 8 different categories, visible in Figure 1(d). Controllers can have different sizes, orientations and granularities. To complete the design of a game, users have the option to switch backgrounds and pick an audio track which fits their style of playing, see (e).

The explicit rules of the game “Snowy Terff” encoded in Figure 1 are:

The player receives 12 points if a cluster of three snowflakes form. The player loses 2 points if two raindrops collide. If raindrops and snowflakes collide nothing happens, they just bounce off. The player cannot tap any character and can only move the large ⊓ controller. There are always only ten snowflakes and 15 raindrops on screen. Snowflakes spawn at the top edge of the screen. Raindrops spawn at the bottom edge of the screen.

Instead of numerical parameters with textual descriptions, Wevva uses iconography for the parameters. The usage of iconography instead of textual descriptions was chosen to keep the interface as clean and comprehensive as possible. Initial tests and observations showed that users try to understand the the parameters through trial and error even if an a-priori description is given. During the development we also experimented with different textural ways to show further information or help in textual form but have not identified a good approach which users seem to use. Including a help system requires more research and potential for future work. The combination of the different options results in a fairly large design space which is exposed to players/designers. This allows the user to create quite different games, beyond visual difference. The app was built with the intention to allow users to rapidly test their changes as they explore this space. To foster this, the play button in Figures 1(b,c,d) can be used at any point in the design to instantly explore changes by playing the modified game.

Games are playable at any time during the design process, with parameters changed incrementally as the player/designer undertakes design exploration by navigating to new parts of the space. Therefore there needs to be a starting point for any given design session. Wevva gives the option of three types of starting point. The first option is to start from one of the built-in games and begin to modify it. The app ships with a number of existing games we have already designed in the app; there is nothing special or hardcoded about these games, as they are simply games we have ourselves made in Wevva and included with it by default. But they serve as initial games to play before the player has designed any of their own, and also illustrate different points in the design space from which the player can start any given session of design exploration. The second option is to choose a more randomised starting point by pressing the “inspiration” button. This button loads a randomly chosen existing game and furthermore randomly perturbs up to 50% of the game’s parameters, thereby starting off the design session in a previously unseen part of the design space. Players can either press the button once and start from there, or continue to press the inspiration button repeatedly until they find a randomised starting point they like. Finally, there is a third option of a “clean slate” starting point. This is a starting point designed to be as close to an empty template as possible, with all parameters set to default values.
4. Methodology

With the previously introduced fluidic game prototype, we held four rapid game jams with a total of 105 participants from a local Girlguiding organisation. From the four game jams, we analyse two, with 69 responses. One session was excluded because a large fraction of participants had to leave halfway through the game jam due to a scheduling misunderstanding, which led to confusion with the remaining participants and likely impacted their experience. We also excluded another session from the analysis to have a better balance between the two age groups of participants. The second excluded session only consisted of younger participants. We include only fully completed questionnaires in the analysis, which amount to roughly 90% of all handed out questionnaires. Therefore the analysis in the following sections is based on survey results from the following two game-jam sessions:

- Session 1: 40 girls age seven to 10 with an average age of eight years.
- Session 3: 29 girls age 11 to 14 with an average age of 12 years.

Participants only participated in one game jam, the two sessions took place a week apart. Participants had no prior experience with rapid game jams or our app.

4.1. Procedure

The game jams serve as opportunities to gather feedback on the effectiveness of the rapid game jams concept, and of fluidic games in enabling them, as well as the design of Wevva. To evaluate the app as to get a general measure of how well people could explore the parametric space and how much they enjoyed it, we conclude our game jams with a general feedback round using paper questionnaires.

For the game jams, we begin by asking the participants to play the included games for ten minutes. A pair of two participants share one device. This means participants take turns when designing their games. Next, we provide a brief introduction to the design interface and give them about an hour to design their own games, again paired. This is followed by a period where they shared their games with other participants.

After wrapping up each game jam, we ask our participants to individually fill in a short anonymised questionnaire, to document their experience with the app and some general information about themselves. The feedback sessions generally took ten to 15 minutes at the start of which we handed out pencils and questionnaires to give them some extra time to wrap up their games and reflect about their experience. Our questionnaires utilise the Visual Analog Scale (VAS) [25] to better express the participants’ tendencies towards or away from certain statements about their experience. We chose VAS over the more traditional Likert scale as, according to Tucker [26], participants seem to find it easier or more comfortable to express themselves using it.

The surveys contained 21 questions, with an extra question present in Session 3, which was added as a result of running the first two sessions and observing the participants’ reaction to the games they created. Questions 7, 13, 14 and 17-20 were open ended containing general feedback but after an initial analysis the answers seemed not particularly revealing, which could be due to young age of the participants or that
the time for the feedback session was too short. Hence, those answers are not taken into account hereafter.

4.2. Discussion of Questionnaire Items

• **Q0**: "How often do you play games on a phone or tablet?" This item is used to get a measure of the participants’ game literacy. We hypothesise that, if people play games less often, it may be more difficult for them to understand and navigate the complex game space, adding the extra hurdle of learning the controls and understand the difference between games before being able to modify them.

• **Q1**: "How often—if at all—have you made games?" Here, we wanted to get an idea how much exposure our participants have had to game design before using our app.

• **Q2**: "How interested are you in game design as a future career?" With this item, we want to see if there is a pre-disposition towards game design. We hypothesise that if people want to become game designers, they might have already done some design or would explore the design space of our app more in-depth fashion which should result in more critical in-depth feedback.

• **Q3**: "Did you enjoy playing Let It Snow / Rain Rain / Jack Frost / Slush Slosh?" The Wevva prototype used in the study has a set of four crafted games – named above – showing examples of what is possible to design. With this item, we wanted to see how much participants enjoyed playing crafted games.

• **Q4**: "How often did you start from an empty game by shaking the iPad or starting from the clean slate when designing games?" When designing new games, users can start either from an existing game or a vanilla version which contains the bare minimum of functionality. This clean slate allows an exploration from a fixed point in the fluidic space, but requires more knowledge about how to navigate the space.

• **Q5**: "How often—if at all—did you use the inspiration button?" The app contains a game generator which can be used to move to a new point in the fluidic game space. This option presents an alternative to the clean slate. The generator takes a template game within the fluidic space, and then randomly modifies up to 50% of it, so that participants can start designing from this point.

• **Q6**: "How useful were the inspiration games for your design?" The choice of using game generation is rarely made explicit in other apps so we wanted to understand its usefulness to the experience and enjoyment when using the app.

• **Q8**: "Did you make many games during this Game Jam using 'Make A Game'?" Instead of asking for the amount of games which can vary between users and game jam time drastically, we were more interested in understanding if the users felt that they were able to create a decent amount of output in the given time.

• **Q9**: "Are you satisfied with the games you created?" The item stands for the perceived quality of the produced games relating to the tool and approach used to design it.

• **Q10**: "How interesting were the games you produced?" In contrast to the quality of the game from Q9, it is also of interest to find out if the participants found points in the fluidic space which they were interested in. It was an assumption that the participants could create games which they are not satisfied with as a game in general but found some interesting dynamics which they liked but could not yet fully utilise.
• Q11: "Do you feel that the app has enabled you to be creative?" Due to the exploration of the game space, we wanted to know if the users could relate to this activity as a creative one and if this made them feel more creative afterwards.
• Q12: "How much do you feel that the games you created are your design?" How much do the participants feel ownership over the created games? This item was introduced in the third game jam after surveying the first results and observing the participants reaction during the game jam. We wanted to closer examine whether participants engaged in the design process develop some form of ownership over the games they produced.
• Q15: "How much did you enjoy using the app?" As part of a typical game valuation we included this item to see whether the participants liked using the app for the duration of the game jam. We can also use this item to identify a specific target audience.
• Q16: "How familiar do you now feel with designing games in the app?" During the design of the app we concentrated on ease of use and accessibility of the design space and we wanted to check whether the app is sufficiently explaining the design elements. We also assumed that Q16 might correlate with Q9.
• Age: To get some insights into the demographics of potential game jam participants and app users, we included age as one of the items, but we did not ask for any other demographic information to minimise personal data acquired by us.

<table>
<thead>
<tr>
<th></th>
<th>Q0</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q15</th>
<th>Q16</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>total (n=69)</td>
<td>Min</td>
<td>0.70</td>
<td>0.00</td>
<td>0.00</td>
<td>1.15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.60</td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>10.90</td>
<td>9.60</td>
<td>9.55</td>
<td>10.00</td>
<td>9.75</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>14.00</td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td>6.08</td>
<td>2.28</td>
<td>2.85</td>
<td>6.58</td>
<td>3.70</td>
<td>2.67</td>
<td>4.29</td>
<td>5.72</td>
<td>7.21</td>
<td>6.63</td>
<td>6.63</td>
<td>7.44</td>
<td>6.68</td>
<td>9.97</td>
</tr>
<tr>
<td></td>
<td>STDDEV</td>
<td>2.74</td>
<td>2.28</td>
<td>3.08</td>
<td>2.79</td>
<td>2.76</td>
<td>3.13</td>
<td>3.50</td>
<td>3.33</td>
<td>2.80</td>
<td>2.63</td>
<td>2.98</td>
<td>2.46</td>
<td>2.62</td>
<td>2.11</td>
</tr>
<tr>
<td>Session 1 (n=40)</td>
<td>Min</td>
<td>0.70</td>
<td>0.00</td>
<td>0.00</td>
<td>1.15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.60</td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>10.90</td>
<td>9.60</td>
<td>9.15</td>
<td>10.00</td>
<td>9.75</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td>6.34</td>
<td>1.89</td>
<td>1.86</td>
<td>6.92</td>
<td>3.83</td>
<td>2.05</td>
<td>4.80</td>
<td>4.36</td>
<td>6.82</td>
<td>6.32</td>
<td>6.71</td>
<td>8.11</td>
<td>6.86</td>
<td>8.38</td>
</tr>
<tr>
<td></td>
<td>STDDEV</td>
<td>2.81</td>
<td>2.47</td>
<td>2.78</td>
<td>2.92</td>
<td>2.96</td>
<td>3.04</td>
<td>3.52</td>
<td>3.53</td>
<td>3.20</td>
<td>3.10</td>
<td>3.31</td>
<td>3.26</td>
<td>2.75</td>
<td>0.98</td>
</tr>
<tr>
<td>Session 3 (n=29)</td>
<td>Min</td>
<td>0.95</td>
<td>0.55</td>
<td>0.00</td>
<td>1.80</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.10</td>
<td>3.00</td>
<td>3.75</td>
<td>2.29</td>
<td>1.30</td>
<td>1.90</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>10.00</td>
<td>7.75</td>
<td>9.55</td>
<td>8.60</td>
<td>9.55</td>
<td>9.70</td>
<td>10.00</td>
<td>10.00</td>
<td>9.75</td>
<td>9.85</td>
<td>9.75</td>
<td>10.00</td>
<td>9.70</td>
<td>14.00</td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td>1.73</td>
<td>2.83</td>
<td>4.22</td>
<td>6.11</td>
<td>3.52</td>
<td>3.51</td>
<td>3.59</td>
<td>7.54</td>
<td>7.75</td>
<td>7.03</td>
<td>6.53</td>
<td>6.42</td>
<td>6.51</td>
<td>6.43</td>
</tr>
<tr>
<td></td>
<td>STDDEV</td>
<td>1.65</td>
<td>1.88</td>
<td>3.04</td>
<td>2.60</td>
<td>2.51</td>
<td>3.10</td>
<td>3.40</td>
<td>1.98</td>
<td>2.10</td>
<td>1.76</td>
<td>2.40</td>
<td>2.80</td>
<td>2.46</td>
<td>2.47</td>
</tr>
</tbody>
</table>

Figure 2: Results of the 17 items of our survey. In the first row bracket, the total results of the two groups (Session 1&3) are given. In the subsequent two row brackets, the results are split between the conditions of being in Session 1 which consisted of younger participants and Session 3 which consists of older participants. The additional item Q12 was only asked in Session 3, and hence is left empty for the total results and Session 1, as it was not yet added to the questionnaire at that point.

5. Results

An overview of the results of our two analysed games jams is given in Figure 2. The results concentrate exclusively on the items that use the visual analogue scale (VAS). Based on [26], we assume that VAS is a good alternative to capture opinions as participants feel that they are not bound to either writing down their own ranking or selecting between different elements, i.e. using the Likert scale. From our observations, we found that the scale did not create any issues during the assessment, even the young participants in Session 1 (aged 7 to 10 years) were able to complete the questionnaire in time. For the VAS, we used a standard 10cm scale and included an example on the first
side of the questionnaire to familiarise the participants with it. For the question texts, we decided on short statements of a single line, asking about the participants agreeableness with the item. Using a paper based questionnaire with short statements assessed through VAS items proved to work fairly well, with only 10% of the forms being partially or incorrectly filled.

In our total sample of 69 female participants, all of them indicated that they had played games before. On average, the participants felt they play a more than moderate amount of games on mobile devices ($\text{avg}(Q0) = 6.08$). While there were no participants in the older group who had not made games in some form or the other before ($Q1$), the average result for having made games before was in the lower third of the scale. One of their motivations for attending a game jam could have been to get first insights into game design and to enhance their understanding of game making. The results also showed relatively low interest in a professional career in game design ($Q2$), but given the average age of participants of ten years, career development is probably not a major concern. Albeit designing games could be a great starting point when motivating children for STEM.

The clean slate, as well as the generator, were used less frequently than we expected: the results are in the lower third of the scale. This goes hand in hand with the fact that generated games were judged slightly less useful ($Q6$) by all participants, ($\text{avg}(Q6) = 4.29$). Also, all participants felt that they made a decent amount of games during the game jam ($Q8$) which, for the duration of one hour, seems encouraging and shows potential for using rapid game jams. Participants in both sessions were also more than moderately happy with their resulting games and felt that it made them more creative. All participants enjoyed using the app $\text{avg}(Q15) = 7.44$, which at the early stage of the app was also very encouraging for continuing the development. In both sessions, on average, the participants didn’t feel that they were entirely familiar with the design interface, but the average of 6.86 and 6.43 for the responses to $Q16$ indicates a more than moderate feeling of familiarity with the interface at the end of the game jam.

After looking at the global results of our survey, we wanted to check if there are any meaningful correlations between the different results. To identify those correlations, we employ the Pearson correlation $r_{xy}$ shown in the correlation tables in Figures 3, 4 and 5, with a probability of being wrong of $p = 0.05$ for all our t-value tests. We also use a two-tailed significance test to remove outliers. The correlations are coloured coded by the strength of their correlation, where red items are negatively correlated, and green items are positively correlated.

Figure 3 presents the global correlations spanning both sessions. The table only contains significant correlations found in the data using Pearson correlation and a two-tailed significance test. All empty cells represent non-significant correlations. The full list correlation table is part of the supplementary material. Two interesting correlations instantly present themselves in Figure 3, namely the negative correlation between playing games and using the clean slate ($r_{Q0Q4} = -0.3$) and the positive correlation between playing games and how interesting the designed games are perceived to be.

---

3 The t-values for all correlations can be found in the supplementary material as well as the data set for all VAS items.
Figure 3: The significant correlations for the combined results of Session 1 and Session 3 using the Pearson correlation \( r_{xy} \) over the 69 samples with an error probability of 5% and a critical t-value \( t_{crit} = 2.00 \). The table contains 13 items excluding item Q12 which is only relevant to Session 3. The t-values can be found in the supplement material. (\( n = 69, p = 0.05, t_{crit} = 2.00 \))

\[
\begin{array}{cccccccccccc}
Q0: Playing Games & Q1 & Q2 & Q3 & Q4 & Q5 & Q6 & Q7 & Q8 & Q9 & Q10 & Q11 & Q12 & Age \\
1 & -0.3 & 0.26 & \\
Q1: Making Games & 1 & 0.32 & -0.3 & -0.3 & -0.4 & 0.35 & \\
Q2: Game Career & 0.32 & 1 & 0.25 & 0.34 & 0.28 & 0.29 & 0.35 & 0.41 & \\
Q3: Included Games & 1 & 0.39 & 0.28 & 0.38 & 0.62 & 0.29 & \\
Q4: Clean Slate & -0.3 & 1 & \\
Q5: Inspiration Used & 1 & 0.49 & 0.31 & 0.3 & \\
Q6: Inspiration Useful & 0.49 & 1 & -0.25 & \\
Q8: Amount of Games & 0.25 & 0.31 & 1 & \\
Q9: Satisfying Games & -0.3 & 0.34 & 0.39 & 0.25 & 1 & 0.71 & 0.47 & 0.45 & \\
Q10: Interesting Games & 0.26 & 0.3 & 0.28 & 0.28 & 0.71 & 1 & 0.49 & 0.47 & \\
Q11: Creativity & 0.29 & 0.38 & -0.3 & 0.47 & 0.49 & 1 & 0.61 & 0.37 & \\
Q15: Enjoyment & -0.4 & 0.62 & 0.45 & 0.47 & 0.61 & 1 & 0.32 & -0.3 & \\
Q16: Familiarity & 0.35 & 0.29 & 0.37 & 0.32 & 1 & \\
Age & 0.35 & 0.41 & 0.46 & -0.3 & 1 &
\end{array}
\]

\( r_{Q0Q10} = 0.26 \). Additionally, the negative correlation between having made games before and how much the participants enjoyed using the app (\( r_{Q1Q15} = -0.4 \)), is clear. There is also a cluster of correlations involving the items Q9 to Q15. In addition to that, the table also contains more negative correlations between having designed before (Q1) and other items for the combined results of sessions 1 and 3 in Figure 3. Another negative correlation is defined between the participant age and the general enjoyment of the app (AgeQ15).

Figure 4 presents the correlations of the subset defined by the participants of Session 1 only. These correlations were derived from 40 participants aged 7 to 10. In contrast to Figure 3, there is no correlation \( r_{Q0Q4} \) but \( r_{Q0Q3} \) shows a strong significant negative correlation. Age has a negative correlation in combination with how satisfied the participants were with their designed games (AgeQ9). The dominant cluster from the global correlation table centred around Q10 is also visible in the subset, the only missing correlation in that cluster is the one linking creativity and familiarity with the app, \( r_{Q11Q16} \). Besides this, the negative correlations within Session 1 are stronger than in the global table, suggesting that the younger participants, which are also the larger sample, drove some of the correlations in Figure 3.

The last correlation table, Figure 5, is the subset of participants constituted by those of Session 3. These participants were between 11 and 14 years of age and also part of the same Girlguiding group. However, this session took place on a different day than Session 1. Within this subset, there are a lot more positive correlations between playing games (Q0) and other factors such as age or how satisfied participants were with their creations. The cluster around Q10 now extends from Q9 to Q16 spanning a large number of significant positive correlations. Q2 and Q3 are (in contrast to Figure 4) stronger correlated to other items. The main difference to Session 1, is that the clean slate starting point Q4 for new designs is strongly negatively correlated to other items.
6. Discussion of Results

In the previous section, we presented the results of our survey on two rapid game jam sessions using our app. The two game jams happened on different days and had a total sample size of \( n = 69 \). In our sample, how often participants play games on mobile devices has a significant negative correlation with the usage of the clean slate. This suggests that the participants who play games less frequently, i.e. novice players, prefer a stable, well-defined starting point for their exploration whereas participants that play frequently prefer to dive into the fluidic space. The clean slate can be interpreted as such a stable starting point, because the user can always reset the game to that same state. This correlation is not significant for Session 1 (which contained the younger set of participants), suggesting that for this group the usage of the clean slate was not impacted by how often the participants played games. For this group there is also no correlation between age and the amount of games played. This could point towards a saturation in game literacy or reached expertise at some later age that is not contained in the younger age bracket.

Interestingly, the correlation between age and having made games before (Q1) is present on a global scale and for Session 1 but is not significant for Session 3. This suggests that making games is at some point introduced within the younger group, but the number of games they make flattens out within the age bracket of 11 to 14. In Session 1 and on the global scale, there is also a strong negative correlation between having made games before and how the participants evaluate their own designs, which could hint at a more critical reflection on their designed games, judging them against previous experiences. In Session 3, as the responses to Q1 have no significant correlation with those to any other question, we can assume that the number of games participants play did not affect their judgement of interacting with Wevva for the older participants, which was not the case for the younger participants. It also does not affect the judgement

---

**Figure 4:** The significant correlations for the results of Session 1 using the Pearson correlation \( r_{xy} \) over the 40 samples with an error probability of 5% and a critical t-value \( t_{crit} = 2.024 \). The t-values can be found in the supplement material. \( (n = 40, p = 0.05, t_{crit} = 2.024) \)
Figure 5: The significant correlations for the results of Session 3 using the Pearson correlation $r_{xy}$ over the 29 samples with an error probability of 5% and a critical t-value $t_{crit} = 2.05$. The table also contains the extra item Q12. The t-values are given in the supplement material. ($n = 40, p = 0.05, t_{crit} = 2.05$)

of their enjoyment when using the app for Session 3, which suggests that they have used different apps before that are not necessarily game design apps or casual creators [19].

As expected, the idea of becoming a game designer (Q2) is related to the age of the participants, especially when looking at our age group of 7 to 14. This result is only significant when combining both age groups. The interesting part is that even though in Session 1 there is no significant correlation to age, there is a significant correlation between designing games and wanting to become a game designer. This supports the previous assumption that STEM activities at the age of 7 to 9 year old girls could affect their future goals. The correlation is also not significant for the older participants, suggesting that making a game is not the driving force behind their goals of becoming a game designer anymore. Looking only at Session 3 and the global scale also shows that participants that want to become game designers did not prefer to the use the clean slate option and generally were more satisfied with their own design. They also felt that they were more familiar with the design interface, which could be based on their previous use of similar tools and their existing expertise in navigating game spaces. Interestingly, there is only a significant positive correlation between the game design career and app enjoyment in Session 3, which could be based on how they compare our app with other apps in which you can design games.

Another interesting finding is the correlation between the time investment in playing games (Q0) and how the participants liked the included games (Q3). This correlation is not significant for Session 3 or the entire cohort but strongly negatively correlated for Session 1. As Q3 shows no significant correlation with participant age, we assume that participants who play games less often enjoy playing the included games and using the app more. This could also be an indication that the app was at the time of testing not polished enough yet to compete with other games but it is generally interesting as the overall enjoyment in Figure 2 for all participants is relatively high. There is also the cluster of positive correlations relating the included games (Q3) to enjoyment and satisfaction of their own designed games. This suggests that the dynamics of our app
were attractive to the participants and navigating a fluidic game space was a novel or interesting experience for them.

Taking a closer look at the usage of the clean slate as a starting point for designing a new one, focusing on Session 3 we see that a set of items negatively correlates with the clean slate. These correlations do not exist for Session 1 or in the global scope. As the clean slate is a fixed point in the fluidic space, which by design is not an interesting game, the participants reported that when using the clean slate their new games were not that interesting and they also felt a lower ownership over them. The reason for this might be grounded in the number of modifications the clean slate requires to arrive at an interesting point in the fluidic space, as the clean slate is an empty game template.

As an alternative to the clean slate, we provide a generator which samples points in the game space. (The third possible starting point for design is to take an existing game and modify it.) Figure[2] shows that the clean slate was used with similar frequency by both groups; the generator on the other hand was used slightly more often in Session 3 without a correlation (Q5) to age or previous experience with games. Interestingly, there is a significant negative correlation within Session 1 between the generator and feeling more creative after using the app (rQ5A11). This could indicate that in Session 1, participants were either modifying other games to explore the space or were using all approaches but were less happy with the generator, as it gave them more complete but very different games each time, which they did not modify much, thus feeling less creative.

Looking at whether participants felt they made a lot of games during the rapid game jam (Q8) shows a clear difference between sessions. This is supported by the fact that Q8 is also positively correlated with age. Additionally, the usage of the generator is positively correlated with the the feeling that the participants made more games. This suggests that older participants were navigating the fluidic space using the generator as they also found the generator more useful in Session 3 than in Session 1, i.e., rQ5Q6(Session1) < rQ5Q6(Session3). During the entire game jam, the participants saved around 100 games which is roughly one per person. However, participants did not save all their games and sometimes used the same game to continue navigating the space. To gain further insights into the design process we analysed those recorded games and we will have a closer look at those games in Section[7]. However, further instrumentation and logging is needed to gain more detailed insights into how users navigate the design space. We believe that having to work with a new game design tool and finishing a new game within an hour long game jam at the age of 7 to 14 years is an achievement which indicates a successful proof of concept. This aligns perfectly with the idea of rapid game jams and more game jams with older participants might enrich our understanding of how to aid the frictionless navigation of fluidic games more.

By looking further into the resulting findings, the large cluster of positive correlations around Q9 – Q15 suggests that if the participants were happy with their design, then they enjoyed using the app and also felt more creative. Interestingly, the cluster spread to how familiar the participants are with the app (Q16) for Session 3 mainly because there are significant positive correlations with the user age for how satisfied participants are with their design and how interesting their games are. This could support the earlier assumption that the older participants were able to understand the complex design space slightly better. A conclusion we derive from the correlations could be that the beta
app we used during both game jams was not supportive enough yet for the user group of 7 to 10 years. In addition to this, we found the significant negative correlation of participant satisfaction with their games and age $r_{\text{Age}} = -0.32$ over all participants, which is an indication of a more critical analysis of their games with increased age while at the same time younger girls enjoyed their designs more. Due to the positive correlation between age and playing games, older participants are probably more critical when judging their own contribution leading to a lower satisfaction. We would like to investigate this finding further, with other age groups.

Based on the feedback, as the participants generally enjoyed their experience and engaged with the design space, the rapid game jams seemed to have worked well both as a testing tool and as a new approach for faster design sessions when creating new games within complex design spaces. After looking through the quantitative feedback we received from the participants we will take a closer look at the game jam output.

7. Qualitative Analysis of Games

In addition to the quantitative survey results discussed above, we conducted a qualitative analysis of the games that participants saved on their devices during the game jams. Participants were asked to save any games they thought would be interesting to share with others, and the last portion of the rapid game jams consisted of participants playing each others’ games. We later collected all these participant-selected saved games from the devices, which totalled to 72 games, in order to better understand their use of the design space. As a major goal of rapid game jams is to encourage free-form design experimentation, we were particularly interested in the variety of game mechanics used in participants’ games.

Table 1 presents one way to understand the design space explored in the game jams, by coding each saved game according to the game mechanics it uses, and tallying the totals for each mechanic to see which types of designs were more popular. We labeled each game with the primary game mechanic it makes use of and one or more secondary

<table>
<thead>
<tr>
<th>Mechanic</th>
<th>Used as primary</th>
<th>Used at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Herding to collect</td>
<td>21</td>
<td>39</td>
</tr>
<tr>
<td>Tap-em-up</td>
<td>13</td>
<td>42</td>
</tr>
<tr>
<td>Keeping separate</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>Catching to collect</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Batting away</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Spawning flow</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Toy-like</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Protect sprites</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Protect zones</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Steady hand</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Fast reaction</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Classification of the game mechanics used in the 72 games saved by participants in the rapid game jams.
mechanics. We developed the set of game mechanics used for this coding in a bottom-up iterative manner, playing each game in turn, and adding any new mechanics observed that we had not yet seen in a previous game. From playing through the 72 games, we identified 11 mechanics used in participants’ games:

- Herding to collect: group together clusters of sprites.
- Tap-em-up: tap as rapidly as possible on certain sprites.
- Keeping separate: keep some sprite types apart.
- Catching to collect: catch sprites with the controller.
- Batting away: knock some sprite types off the screen.
- Spawning flow: try to manipulate spawning patterns.
- Toy-like: focus on enjoyable interaction, not scoring.
- Protect sprites: keep a sprite type from exploding.
- Protect zones: keep sprites from going off screen at certain places.
- Steady hand: make careful movements, e.g. to thread through a narrow gap.
- Fast reaction: make rapid, precise movements or taps.

As can be seen in Table 1, the herding to collect and tap-em-up mechanics featured in some form in more than half of the games. This is not surprising, since scoring by forming clusters and scoring by tapping are two of the more straightforward options in the design space. The least common mechanics, seen in only three games and not in any case as the primary mechanic, were those based on skill in controlling the controller or sprites, whether the steady-hand or fast-reaction kind of skill. We had designed a number of games using these mechanics in our own 10-minute game jams, so that was an interesting difference to notice.

One aspect of automation our current apps do not have is automatic fix-up of games to balance them and avoid exploits, although we have done research on a version of automatic tweaking that runs server-side [27]. To see whether such a feature would be important to add, we further classified the 72 games according to whether we, as expert players, were able to quickly find an easy exploit in the game design. We were able to do so in 31 of the games. Of these, the two most common exploits were being able to win by indiscriminately tapping (seen in 22 games) and being able to win by doing nothing at all for a short tap-em-up games were one of the two most common mechanics used, it’s not clear that winning by indiscriminate tapping would actually be considered an exploit by the designers. We observed, for example, some pairs of users sharing an iPad taking turns playing a very easy game requiring rapid tapping, but competing to beat each others’ best scores.
8. Conclusion

We introduced the notion of rapid game jams, 90-minute jams that differ from traditional game jams by being both much shorter, and more focused on design experimentation than implementation. To support these jams, we have developed the novel concept of fluidic games. Fluidic games are a set of games in a design space that can be modified on the fly to produce other games in the design space, enabling users to navigate that space from within the game to discover, design, and share new games.

We tested our first prototype fluidic game app, Wevva, in multiple rapid game jams and present a user study with 69 participants investigating the concept, followed by a qualitative analysis of 72 games the participants designed and chose to save.

Based on the quantitative feedback from the surveys, Wevva seems to attract people who tend to frequently play games; those frequent players also feel more creative using the app than novice users. Those who are interested in becoming game designers or more experienced in game design used the clean slate less as they seem to feel that the games they produce from that point in the fluidic space are less interesting. Those not interested in a career in game design or younger participants explore more around the clean slate, perhaps because it gives a familiar starting point for exploring the game space. Also, there is a clear trend that younger participants seem to explore the design space less freely and need more guidance. The resulting game mechanics from the qualitative analysis are more mixed, as participants saved only about one game on average to share (though they designed more intermediate games), and a large fraction of the saved games use a few similar mechanics, so further research is needed in understanding the design process and its bottlenecks. However, a reasonable portion of interesting new game mechanics were developed during the rapid game jams and the participants were enjoying the design exploration. Thus, the concept of rapid game jams itself and the beta app worked well enough to not only allow users to express themselves to a certain degree but also enjoy the process.

The next steps in our development will involve further and deeper instrumentation of the design process to identify bottlenecks and problems in the process of exploring complex design spaces. In addition, we plan to release Wevva on the iOS app store as the first in a series of fluidic games, in which the level of sophistication of games that can be made increases, in line with the level of computational support provided for the design process. To this end we will draw upon Artificial Intelligence techniques to (a) increase the sophistication of the automated generation process (b) provide automated playtesting so that fluidic games can give player/designers feedback on their games, e.g., finding a simple exploit (c) provide automated levelling, whereby multiple variations are produced from a base game, with each increasing the difficulty level. To help introduce the notion of fluidic games to gaming and broader communities, we will continue to organise rapid game jams of the type described here, and endeavour to study whether they help towards our goal of democratising game design.

Acknowledgments

This work is funded by EC FP7 grant 621403 (ERA Chair: Games Research Opportunities). Thanks to Girlguiding Cornwall and the organisers of the Falmouth University Girls Can Code event.
References


