

User Research Retrospective for 'Flowrunners'

DADIU 2019 Team 2 (Seaside Resort), Graduation Game Project, December 6th 2019

QA & UR Manager Mikkel Mathiasen

Overview

The user research and UX testing for Flowrunners was based on Sanders' *generative tools* (E. Sanders, 2002; E. B.-N. Sanders, 2000; L. Sanders & Stappers, 2012) and a three-day workshop with Alessandro Canossa on October 7th, 8th & 9th 2019.

The section on *methodology* is an excerpt from the test plan, and adds nothing new. The sections on *process* and *protocols* add some reflections to the methodology.

Sections on *benchmarking* and *findings* describe some of my findings from the research and tests, and more is available in the Google Drive folder below; notes, protocols, a log and other documents are available from this link: https://drive.google.com/open?id=1lCZsVqCMYJ_Ev6s3ItvAtM7apwqNUefa.

Contents

Overview.....	1
Methodology (Test Plan Excerpt)	1
Process Review	3
The Protocols	4
Benchmarking.....	5
Some Findings.....	6
References.....	9

Methodology (Test Plan Excerpt)

This section (The methodology section) is taken directly from the Test Plan (pages 8-10), and describes the general methodology for UX testing and user research. The whole Test Plan is available here:

<https://drive.google.com/open?id=1EjwsWJNuo4vkb0T6KcaVEa8ZSMNtWDBa>.

This chapter covers the UR & UX aspect of QA & UR manager's responsibility. Details will be light, as each session will be planned and documented separately (if at all), but the purpose and methods of user testing are important to establish.

UX testing is, unlike QA testing, not for finding bugs, but for finding killjoys and generating feedback and ideas about mechanics, controls, graphics, difficulty, social aspects, etc. The aspects investigated through user- and playtesting will likely go beyond the scope of quality assurance as describe in *Bounds*.

UX testing can be done *with* and *without* a working build. In early stages of development tests should be conducted on concepts and sketches, and findings would be generative, i.e. participants produce maps and sketches which may inform design. In later stages of development, testing with users should be done on the basis of a working build, accompanied by some framework for expressing and capturing their opinions and reactions; these later tests would be far more evaluative than generative.

Since QA testing is focused on the functional aspects of the game, it makes sense for tests with users to focus on their experiences and the experiential aspects of the game. For this purpose generative and visuals based methods are useful; the methodology includes Generative Tools/Toolboxes (Sanders), Visually Stimulated Discussion (Weber), Metaphor Elicitation Technique (Zaltman) & Mental & Conceptual Mapping

Protocols and testing mentality

Tests with users will be conducted on the basis of a protocol made beforehand. Some protocols may be very specifically directed at a single test, and some may be used for several separate tests. A protocol should state the purpose of a given test, and contain some semblance of a script or procedure; a sequence of questions and activities.

While protocols may be very specifically planned, it is not intended to be followed meticulously; each participant will have their own opinions and a particular player profile which impact how they play and their thoughts on the experience. Performing UX tests will require agility and responsiveness from the researcher, more so than requiring that the protocol be followed with precision.

Reporting on playtesting

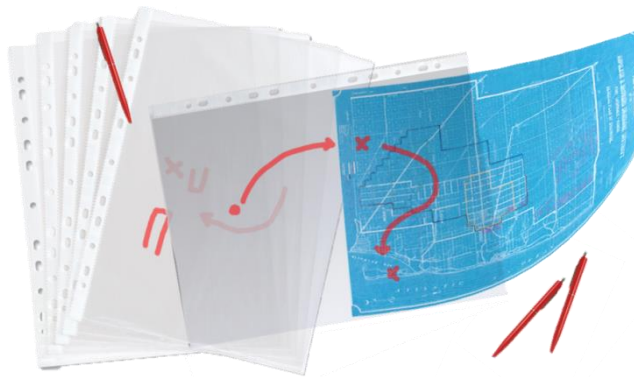
Findings from tests with users may be compiled into reports, synthesized so that they are easily and quickly read. These reports could take the form of slides, but more than likely they will be verbal.

User testing may also feed into market research and QA, e.g. terms uncovered with ZMET may form the basis of a benchmark.

Generative tools

User testing could make use of generative tools. After (or in conjunction with) playing the game, participants could be presented with a generative toolset consisting of printouts of props and levels, illustrations and drawing tools, which would then serve the participant in expressing their experience in narratives and/or visual compositions. A short, post-playtest, generative “session” may just be a single question or reflection which is visualized on a physical board or model.

A generative toolset *could* consist of level schematics and concepts, markers and plastic envelopes; participants could mark the schematics/concepts through the envelopes with concrete changes and suggestions, or with experiential notes. It could look like this ↓



Visually Stimulated Discussions

Instead of generating visual material, participants may engage with the QA & UR manager over visual content from the game’s development. Concept artwork, screenshots, videos of prototypes, etc. would be shown to the participant before or after playing the game. These visual materials would serve as an outset for a discussion and reflective session.

This could also be combined with a generative toolbox.

This method is based on: Weber, S. (2008). Visual Images in Research. In J. Knowles & A. Cole, *Handbook of the Arts in Qualitative Research: Perspectives, Methodologies, Examples, and Issues* (pp. 42–54).

Mental & Conceptual Mapping

The three above mentioned methods – generative tools, visually stimulated discussions and ZMET – will be conducted with the purpose of forming a visual and conceptual map of players understanding of and associations with particular concepts. The map will likely be the primary product of these tests, alongside the generated materials.

ZMET

Zaltman's Metaphor Elicitation Technique (ZMET) is a type of generative tool based on images, concepts and *constructs*. This method asks participants to bring ten pictures relating to a specific concept, followed by a process of sorting, reflecting and combining the participants' images. The method is intended to *elicit metaphors* and associations through the process of narrating and organizing the collection of images. There are ten steps ↓

1. Storytelling: describe the content of each picture
2. Missed images: describe pictures that was impossible to obtain and explain relevance
3. Sorting: sort pictures in meaningful piles and label them
4. Constructs: use pictures to derive constructs (Kelly Repertory Grid and Laddering)
5. Most representative picture and "widening" (what other elements could be part of the picture)
6. Opposite images: user describes pictures that are the opposite of the topic
7. Sensory images: describe tastes, smells, touch, sounds, colours and emotions attached to the topic
8. Mental map & vignette: bring the participant's unconscious thoughts to a level of awareness
9. Collage: a summary image or montage expressing the topic under investigation
10. Consensus map: highlights common themes and threads showing relationships among constructs*

Appreciation Testing

An appreciation test (or *user acceptance test*) is intended to gauge the players' overall impression of the final game; its playability, learnability, memorability and players' enjoyment. This will be the latest form of test and should be conducted with a version of the game as close to gold as possible. Unlike previous tests, appreciation test will only be evaluative.

Selecting Participants

Users will not be selected thoroughly based on demographics or a questionnaire, instead, participants will fill out a psychographic survey, i.e. Quantic Foundry's *Gamer Motivation Profile* survey, so that findings can be compared on the basis of their player profile.

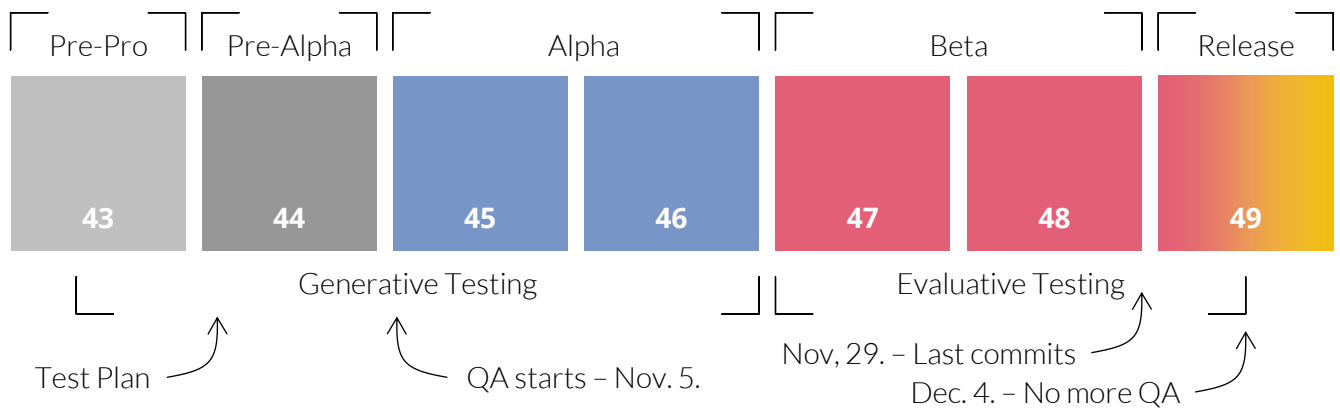
Process Review

I split the user research process for Flowruners in two parts: *generative testing* and *evaluative testing*. In this process there was also a third element of *benchmarking*, which is not inherently part of either, but I fit it with *generative testing* for planning purposes.

Generative testing took place in the first half of the production process, and is most thoroughly covered in the methodology section. The underlying principle was to assist game design and art with insights and concrete ideas while there was no actual game prototype to test on.

In the later half of the process testing would transition into an evaluative mindset, where exploration of new gameplay ideas and art would be phased out in favor of evaluating the existing gameplay. Later in the process there would be more levels and more refined gameplay to evaluate, and changes to core mechanics and art would be difficult.

Below is a visual overview of the testing process – boxes representing weeks (from the Test Plan):



The transition from generative to evaluative testing was never intended to be sudden, but gradual; as the game became more refined and *playable*, the focus of tests would gradually change. I found this to work quite well; as the game gradually became more playable, participants would gradually be asked to play more and draw less. The transition seemed to come quite naturally.

I do think that the transition into evaluative testing started too late, or happened too slow, as our game designers were asking for more concrete feedback on the game's core mechanics and learnability as early as week 46. It probably is not particularly early, but it was not in accordance with my planning, which I take as a sign that my planning of tests was off.

The Protocols

For tests with participants I wrote protocols; I wrote two generative session protocols and four playtesting protocols in total. As mentioned the protocols consist of an overview section, which describes the purpose of the sessions, and a script to guide me in conducting the test. The protocols were very useful during the testing sessions, especially as I have very limited experience in interviewing and observing users; with the purpose clearly stated I always knew what to look for, and the script meant that I could consult the paper if ever in need of a question for the participant.

The playtesting protocols are based on *appreciation tests* and *usability tests* as discussed during our workshop with Canossa. The script contains questions concerned with both usability and with the users' experience, but most questions are not distinctly one or the other; the questions are intended to gauge the participants' immediate thoughts on particular aspects (appreciation/desirability) *and* their understanding of the game's mechanics and levels (usability/playability). Despite being based on existing types of tests, I wanted a more open-ended approach to my sessions, so as to not limit the participants in their ability to express their pleasures and frustrations. I would often ask specific questions about the participants' repeated actions or a particular area in which they spent a long time, and I found that it would easily evolve an interview into a narration of the participants' current experience with our prototype and their previous experience with other games. You might say that the protocols provided the backbone for the sessions, but responsiveness and curiosity made each session unique.

Player Profiles

I wanted to make use of Quantic Foundry's Gamer Motivation Model for comparing my participants, but I ran out of time during most sessions, so it never really became a part of the testing process. We discussed this model at the workshop with Canossa, and I thought it would be quite useful, as I chose to accept anyone willing to test. I imagine that knowing the motivational profile of my participants would have made it easier to compare their responses to my questions, as there was seemingly quite the difference between what spiked their interest.

Example of a test script

The script below is part of *Playtest Protocol 03*, from November 19th. Other protocols are accessible in the Google Drive folder linked in the overview section.

1. Participants play the game
2. Insights and reflections directly from playing the game
3. Discussion the character controls
 - About air control (e.g. changing directions) – too much or too little
 - What's the speed like?
 - What's the constant movement like?
 - Should levels contain small nodules for the character to "rest" on
4. How do the levels relate to the character's movement
 - Areas which are difficult due to controls?
 - Are there areas whose layout and ingredients make them difficult?
5. Thoughts on assets and background?
6. Thoughts on SFX and music?
7. Discussing and reflection upon visual design material with participants – *Visual Reflection*
 - Does the character's movement match her appearance?
 - Concept art and level design sketches
 - This may include sketching on materials, or generating new sketches

Benchmarking

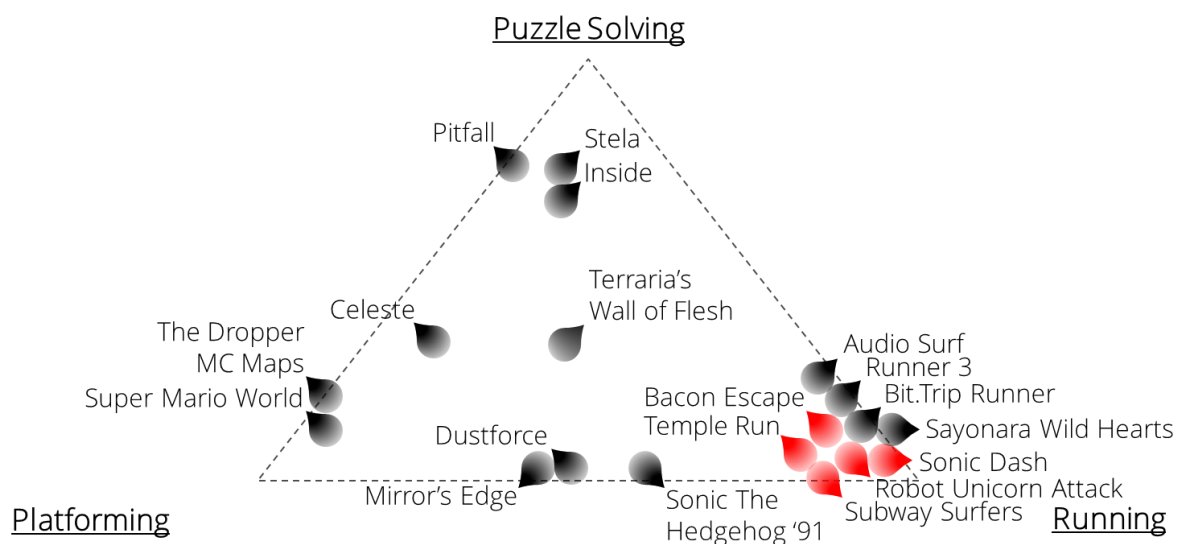
In the workshop with Canossa we discussed the various ways to benchmark a game. I did some benchmarking of my own as part of the early user research; this includes a single session of ZMET, and a study of existing games. The initial ZMET session provided me with a repertoire of aspects to explore in existing games, like freedom and qualities of the movement and the environments. This repertoire, as well as keywords from our early game concept, provided the foundation for an analysis of games that in some way are related to ours. This analysis would later serve the QA process indirectly, and has been useful in placing our game in the market.

These maps were also supposed to be used as part of playtesting, where participants could place the prototype that they played in one or several of the maps, but I frankly forgot.

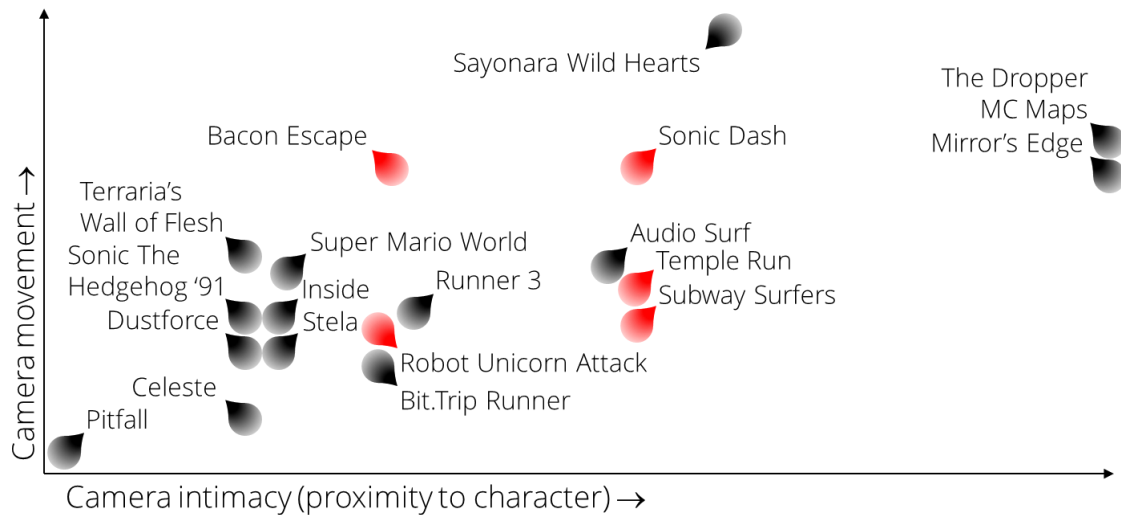
All the maps and a drawing of the games I looked at are available in the Google Drive folder:

https://drive.google.com/open?id=1reW7i-QfcncR0_F_BhBbPcoK1qEPHVhd.

Below is a map of three major genres that we looked at in developing the early concept:



The below map is based on two important aspects of our early concept; camera movement and intimacy:



Some Findings

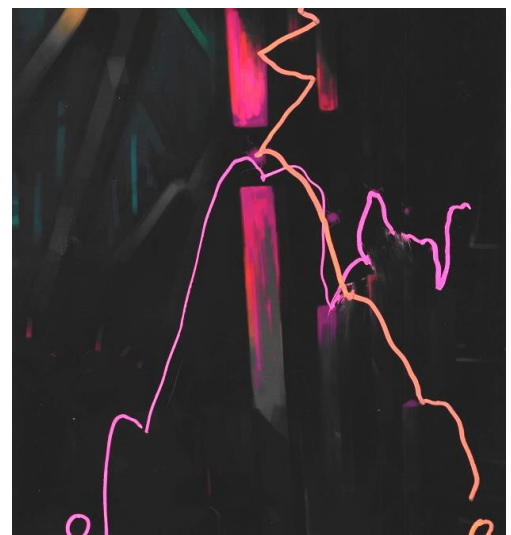
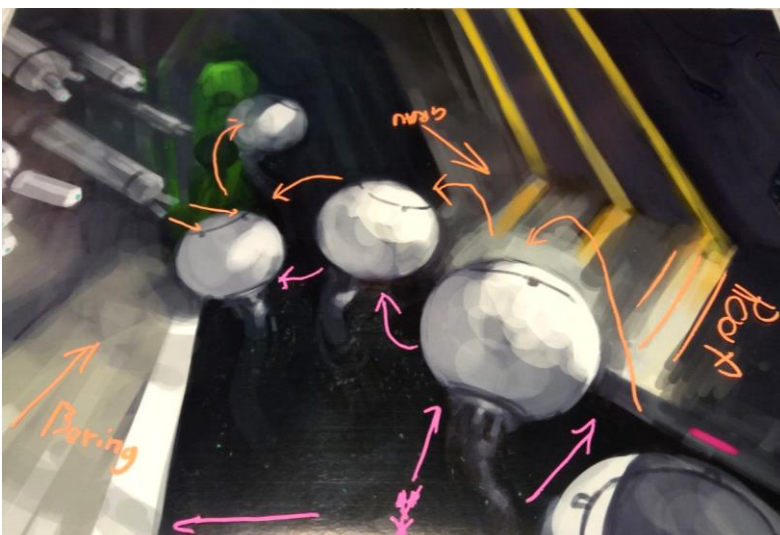
Before I finish off this report, I will go through some of my findings. The Google Drive folder linked in the Overview section contains all of the notes and most of the visual materials generated during my research and tests. There is also a log of my testing in the folder.

Early Generative Testing

The first batch of tests was completely generative, as there was no game to test. I engaged in discussions with participants, and armed with visual concept materials and markers we discussed how they expected characters and environments to behave, and participants marked the materials to indicate movement and possible routes for the character.

In these early tests I discovered that the expected path through the spaces in the concept art were very similar to the path envisioned by our art director. I discovered that the equipment that our main character wears in the primary concept has implications for how she's expected to move, particularly that the boots she is wearing would expectedly make her able to jump higher and longer than a "normal" human. I also found that sound would potentially play a large role in communicating some environmental assets, as participants had very different impressions of what they were.

Below are drawings made by participants on top of concept art, on November 3rd (left) and November 7th (right), and again there are more in the Drive folder linked in the *overview*.



I compiled the insights from the generative sessions into a tree-like model, separated into categories based on the design department for whom they would be most useful. I have my doubts about how useful they actually were, but I think that may have to do with the high detail. The insights should probably have been reduced even further.

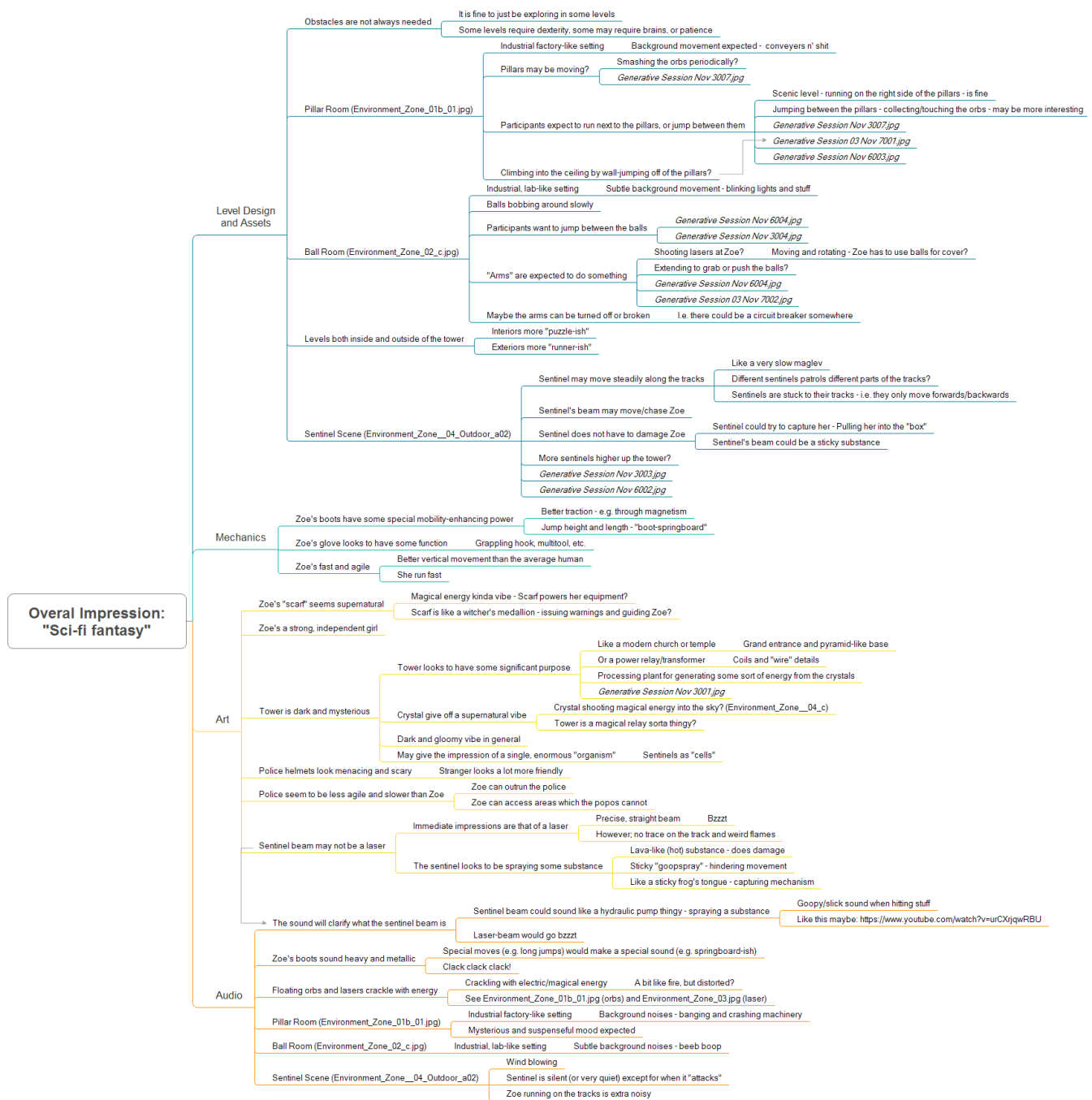
I wanted to compile a similar tree for later tests as well, but seeing as it was not that useful, I did not.

Early Concept Map

Here is the mentioned concept “tree” from the generative sessions.

A more easily readable version of this map is available on the Drive folder:

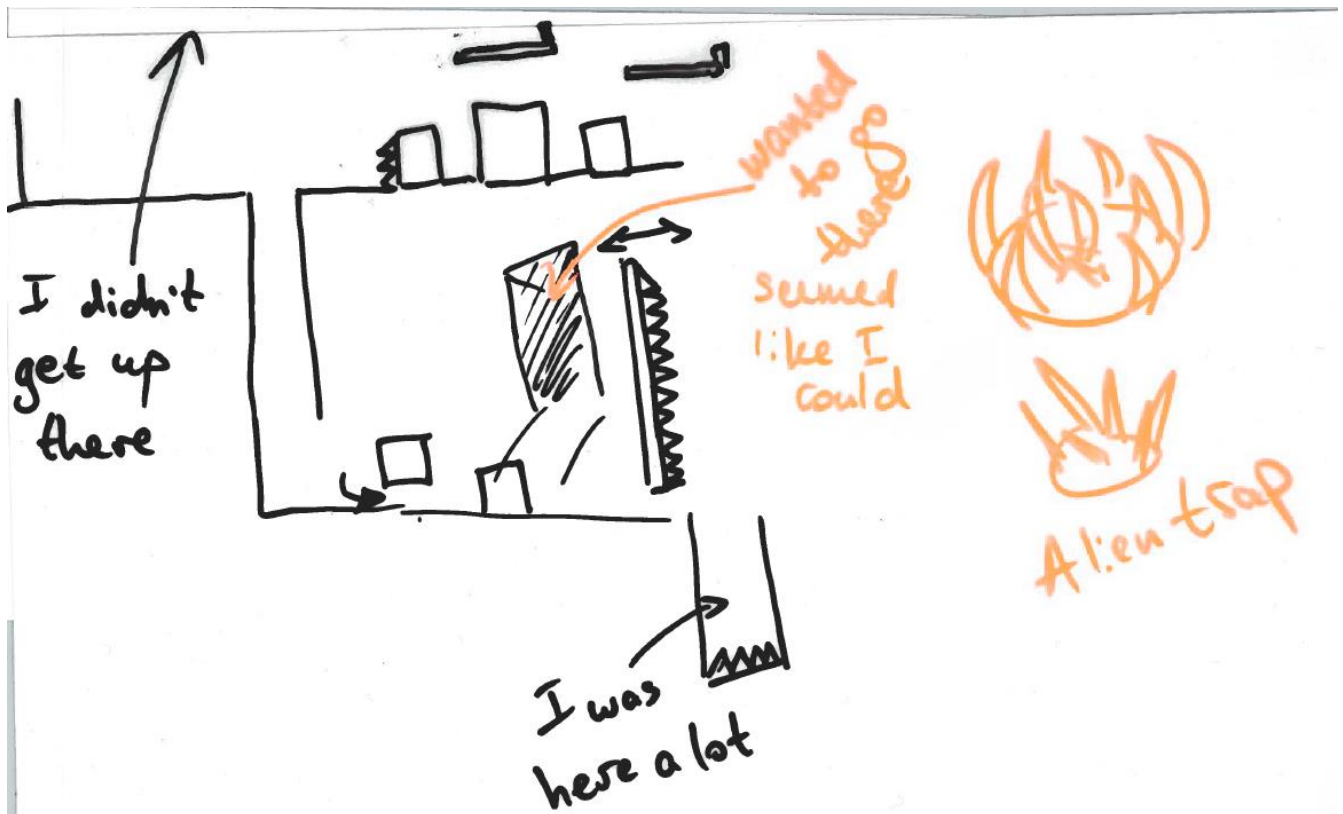
https://drive.google.com/open?id=14m00GcY6jvS2_KrVtEokLUTRNdRk6dQS.



The Transitional Test

A particular test I conducted just before the game entered beta (on November 12th) had the participant play an early prototype of the game, after which we discussed the character's movements over the primary concept art for that particular character. This was a unique test in that half of it consisted of evaluating the gameplay prototype, and the other half had the participant generate sketches; drawing on the character concept and drawing the level from memory to see which parts were most memorable.

From this test I learned that the most punishing areas were also most vivid in the participant's memory, and seeing the character concept after playing with the controls had them thinking that the boots were some sort of skates. Below is the sketch that the participant drew of the level.



Late Stage Appreciation

The last batch of tests was completely evaluative, and is the closest I got to the *appreciation* test and *usability* test that we discussed during the workshop. At this point I was still using my playtesting protocol, but revised to focus on usability questions. I asked participants to play the game without any assistance or instructions, to see if the tutorial served its purpose, and I would take care to ask participants about areas where they either spent a lot of time or areas they completed on their first try.

Generally the game functions well, and the controls are learned quickly in the first level; most participants express that they find the controls intuitive, though there are some that would prefer a more traditional, button-based control scheme. From a usability perspective the main issue is a special move that the character may perform by jumping while rolling; most participants do it by accident, and do not realize that it is a mechanic. If levels were not designed around the use of that mechanic it would not be an issue, but some segments require that move. The last iteration of the tutorial has arrows to better illustrate how to perform the special move, but it was implemented later than my last test, so I do not know if the issue still persists.

Other than the roll-jump I found no major issues with the game.

References

References for UR & UX

- Canossa, A. (2019, October 7-9), Three day lecture & workshop on user research at the National Film School of Denmark, Theodor Christensens Pl. 1, 1437 Copenhagen, Denmark
- Quantic Foundry (n.d.) Gamer Motivation Model [Web Page], Retrieved December 5, 2019 from <https://quanticfoundry.com/>
- Sanders, E. (2002). From user-centered to participatory design approaches. In J. Frascara (Ed.), *Design and the Social Sciences: Making Connections* (2005th ed., pp. 1–7). New York: Taylor & Francis.
- Sanders, E. B.-N. (2000). Generative Tools for Co-designing. In S. A. R. Scrivener, L. J. Ball, & A. Woodcock (Eds.), *Collaborative Design* (pp. 3–12). https://doi.org/10.1007/978-1-4471-0779-8_1
- Sanders, L., & Stappers, P. J. (2012). How it works: Generative tools and techniques. In *Convivial design toolbox: Generative research for the front end of design* (pp. 65–95). Amsterdam: BIS.
- Weber, S. (2008). Visual Images in Research. In J. Knowles & A. Cole, *Handbook of the Arts in Qualitative Research: Perspectives, Methodologies, Examples, and Issues* (pp. 42–54).