

# External Playtesting

## ▼ Method

### ▼ Cooperative Performance Metrics (CPMs) - (Aghabeigi, 2006)

Definition of cooperative performance:

*"cooperative performance is composed to a set of social and game play behaviors and interactions that happen while people play cooperatively with other people in a both physical and virtual shared environment"* (Aghabeigi, 2006, p.53)

The following CPMs are labels to assign to events and behaviors observed during the playtests:

<b>Metric Name</b>	<b>Metric Description</b>	<b>Metric Application</b>
<b>Positive metric.</b> <i>Laughter and excitement together</i> (label events happening in the same space only once per cause)	<ul style="list-style-type: none"><li>• The playtesters laugh at the same time to a specific game event. (ignore instances were only one player laughs without the others)</li><li>• The playtesters express verbally that they are enjoying the game.</li><li>• The playtesters show facial expressions and other nonverbal behaviours that clearly express happiness or excitement.</li></ul>	This metric implies the explicit fun has been raised by participants while playing game in a cooperative mode, so having high values for this metric, represents the high degree of enjoyment.
<b>Positive metric.</b> <i>Worked out strategies</i>	<ul style="list-style-type: none"><li>• The playtesters talk aloud about solving a shared challenge.</li><li>• The playtesters divide</li></ul>	This parameter explains the degree which players have cooperated together for proceeding

	<p>the game zone to different parts in order to divide and conquer.</p> <ul style="list-style-type: none"> <li>• The playtesters navigate the world while consulting with each other.</li> <li>• The playtesters show pre-planned gameplay behaviour that emerges in similar cases.</li> </ul>	<p>in game play, and low values represents to less cooperation than cases which a high value for this parameter has been reported.</p>
<p><b>Positive or negative metric depending on the circumstance.</b> <i>Helping each other</i> (researchers should label events under the Helping CPM when one player is helping the other and not when both are helping each other, this metric is also directly tied to difficulty and can be used to tune the difficulty of the game)</p>	<ul style="list-style-type: none"> <li>• The playtesters talked about controllers, and how one can use the game mechanics.</li> <li>• The playtesters told each other the correct way of passing a shared obstacle.</li> <li>• The playtesters saved and rescued the other player while they were failing.</li> </ul>	<p>This parameter shows the positive social atmosphere which a cooperative game can create, and mostly dependent on what players do in physical space, for example helping each other using controllers, guiding each other, and etc.</p>
<p><b>Positive metric.</b> <i>Global strategies</i></p>	<ul style="list-style-type: none"> <li>• Playtesters take different roles during gameplay that complement each others' responsibilities and abilities</li> </ul>	<p>This parameter measures both social and game play aspect of role taking in cooperative games, as people play indifferent roles, they can have a different experience, so this parameter will try to consider the number of times the people changes their roles, which high values shows players are interested in different dimensions of game-play, and low values represents the</p>

		<p>minimum effort for looking at game from other perspective, anyway, this parameter depends on this fact that game offer different roles or not.</p>
<p><b>Negative metric.</b> <i>Waited for each other</i></p>	<ul style="list-style-type: none"> <li>• One playtester waits for the other to catch up</li> </ul>	<p>This metric can show both engagement and frustration, and it really depends on the skill gap between players who are playing the games, if the players are in similar level, then this value represents a positive social and game-play atmosphere which encourage players to support each other, otherwise it depicts a high frustration for one player who need to sacrifice all his time for another one to catch up.</p>
<p><b>Negative metric.</b> <i>Got in each other's way</i></p>	<ul style="list-style-type: none"> <li>• One playtester leads and the other lags behind.</li> <li>• One playtester wants to do an action and the other wants to take a different actions, and by taking these actions they interfere or hinder each other's goals</li> </ul>	<p>This value mostly presents a negative situation which can lead to frustration moments, having a high degree of conflicts can lead to frustration while having an average value can be acceptable and in some cases challenging that encourage players to decide with each other.</p>

▼ **Running user tests with limited resources and experience (Huguenin, 2018)**

“The key factors are, first, to properly identify what would benefit your game the most with the amount of resources you can dedicate to playtesting, and, second, to take the most efficient course of actions to prepare, run, and report on your test.” - (Huguenin, 2018, p.418)

Preparation:

- Decide what to focus on with the test
  - “determine strict goals for your test session” (p.418)
  - “By restricting the size and the number of focuses, you will be able to produce feedback of greater quality (it is much easier to analyse data with a clear question in mind), and you will spend fewer resources (i.e., time) to set up, run, and analyse the test.” (p.418)
  - “each objective should be reflected in a portion of your test plan”
  - Main typed of test goals: usability (sample: no more than 5 users, observation), how satisfying the experience is for your players (sample: around 20), taking general measures of what your players do in the game (observation)
- Decide when to test
- Decide where to test
- Decide how to run the test
- Determine player profiles: have them being familiar with the genre and the support used to play the game (ideally they should not make or talk about games for a living but bro we are limited here let me live)

Managing resources “small and manageable tests will need to have a very limited focus” - p.420

organize the space: again, I suggest we use one of the smaller rooms at DR

Running the session:

- welcome the players, have snacks and water, explain what you need to explain
- [see the previous two sections of notes ]
- be welcoming but not too friendly (lol good luck, no but seriously if testing with our friends keep them focused)

## ▼ Observing the player (Sangin, 2018)

“When preparing for observing the player experience, you need to consider two main elements: (a) have a clear idea of what you need to observe, and (b) have a clear idea of how you will document your observations.

To do so, you need to consider three main components:

- Collect and clarify research objectives and business needs.
- Develop a thorough understanding of the game content and the player experience.
- Identify the key events where the content is not experienced as intended.” - p.176

Clearly stating the usability goals should enable you to identify elements of the player experience you need to cover during your observation - p.177

Once you have listed the study objectives, it always helps to rephrase them as questions. Whenever possible, break down each research question into sub-questions. For each sub-question, assess if it relates to clarity (can players work out what to do and how to do it?), way-finding (can players work out where to go?),

and difficulty (are players able to perform what is required of them?) - p.177

How much detail do I need to answer the main objectives? - p.178

Start with the categorisation of the objectives in terms of 'knows what to do', 'knows how to do it', 'is able to do'.

'Knowing what to do' is subject to game instructions and clarity of objectives. In your game, identify the instructions and prompts that are designed to help players understand the game objectives and rules.

'Know how to do' is related to on-screen clues on how to control the game. You need to compare the inputs used by the player and the inputs required by the game (e.g., Is the player pressing the correct button to jump? Is the player tapping the appropriate element on-screen?).

'Is the player able to do' is related to game difficulty. This usually implies looking out for failure states such as deaths, crashes, or failing to complete a level or mission in time. Repeated failures need to be captured as well as signs of frustration (e.g., sighs, swearing) as symptoms of an imbalanced game. - p.178

The structured approach requires careful preparation of a note-taking protocol and template. The purpose of a structured approach is to establish a clear observation grid on which all relevant events are laid out in advance. The researcher only needs to tag a specific event on the grid when it occurs (often with specifically assigned codes and shortcuts). A structured approach frees up a lot of mental bandwidth for observing. The drawback is that it can be often constraining and does not enable the freedom to capture events that fall outside the established structure.

A semi-structured approach is a happy middle ground where you prepare a global structure for your notes but still allow enough freedom to adapt on the fly. Typically, you should have an idea of how to group events in meaningful buckets. However, the form and level of details can vary from one note to another.

Typically, the analysis of observations follows these main steps:

- Collating data: This is where you group observations from all players that may have been scattered across multiple sheets or files into a single consolidated one.
- Organising data: In this phase, you group events thematically. If you have been writing observations on individual sticky notes, a good technique is to apply a closed card-sorting method to group issues into categories defined by your research objectives or business needs.

- Aggregating data: Aggregating observation events boils down to transforming all single usability events observed into general usability problems. A general usability problem is a rewrite of your events where you describe the what, the why, and the how of the problem, and, if relevant, shedding light on the 'when' and 'to whom' they happened.

Plan and prepare for observing and capturing the usability events that your players may encounter. Make sure that your observation and notetaking guides take into account the research questions and the intended player experience.

- Build a thorough understanding of the player experience in your game to list key usability events where the players' experience may not match the intended experience.
- Build a clear template to structure and organise usability events that you capture in terms of time, space, categories, or hierarchies.
- Make sure to optimise the time you spend observing by reducing the time you spend taking notes. Use key words or phrases when taking notes and avoid capturing events that are irrelevant to

your research questions.

- Whenever possible, have a sidekick who can take notes when you moderate sessions.
- Use the appropriate tools that allow you to save time both during capturing and analysing. Consider using a mind-mapping tool as it gives you flexibility to organise and structure your data in various appropriate ways.

## ▼ Retrospective think-aloud

Retrospective think-aloud protocols, also known as 'retrospective testing' (Nielsen 1993) or 'aided subsequent verbal protocol' (Henderson et al. 1995), differ in one respect from concurrent think-aloud protocols: rather than thinking aloud while working, participants initially carry out their tasks working silently, and only verbalise their thoughts afterwards on the basis of a video recording of their task performance. - Maaïke van den Haak , Menno De Jong & Peter Jan Schellens (2003)

Video-recording-aided retrospective think-aloud protocol: ask a playtester to verbalize their thought as they watch a video recording of their playthrough.

Why retrospective think-aloud instead of concurrent think-aloud?

- We are already taking observation data during the playtest with
- it does not interfere with the performance and with the data gathering of

- while still allowing us to have the playtesters verbalize about their experience
- we don't really care about checking for usability, since it was already covered in the previous playtests, we want them to verbalize about their experience of playing together
- according to van den Haak , Menno De Jong & Peter Jan Schellens (2003), retrospective think-aloud is more effective in revealing problems that are not observable, and the participants are more likely to give explanations and suggestions, which is what we want from the playtests

Who should conduct the sessions (Knoll, 2018, p.192):

- Ideally it should be someone not involved in the production of the game, because our presence can influence the playtesters. But we can't really do it in another way so we'll just have to be aware of this fact and take it into account while analyzing the results. Make sure to let people know that we won't be offended etc. we are developing the game and we want to improve it and change it based on their experience and feedback.

How to prepare (Knoll, 2018, p.193):

- Recruit representative participants
- Play not at play lab but in one of the more chill rooms at DR
- Prepare questions for after the the playtesting sessions

Procedure:

1. Record the screen while the playtesters are playing
2. Let the participant know that they can and should be honest, what they say will not cause offence.
3. When showing the video tell participants that we would like them to think out loud and give us an in-retrospective insight on their thought process. Give them examples of the type of things that we are interested in, but don't be too specific
4. *"Once the session is underway, it is common for participants to forget that they should be thinking aloud, especially if they start to*

*become immersed in the game"* → this will not be a problem for us because when playing co-op games players talk to each other

Capturing think-aloud data:

- Take notes on what the playtesters are saying

Limitations:

- usually thinking out loud in a simulated gaming environment is not a natural experience, in our case however, it should come natural as long as it's about the two players communicating with each other
- playtesters might play the game in a way that is easier to explain than how they would normally play
- similarly, while it is usually more difficult for a player to explain real time what they are doing in difficult sections of single player games, in a co-op experience they do have to communicate verbally to each other as part of the experience
- it takes a long time to analyze the results

## ▼ Shorter versions of the playtesting protocols

Is the prototype working as intended? Is the prototype usable and accessible?

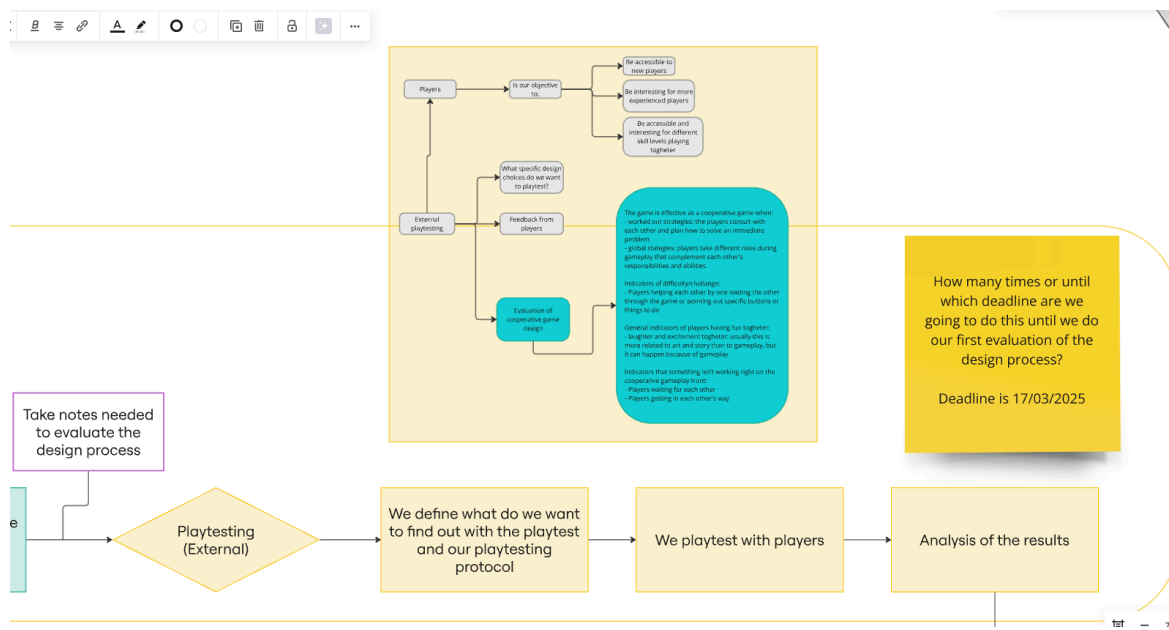
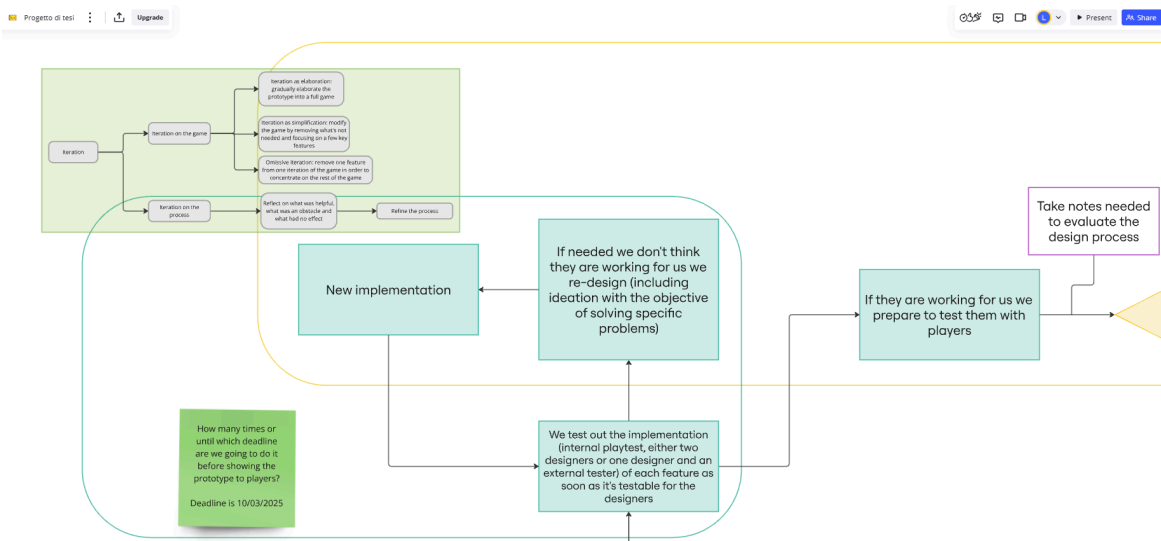
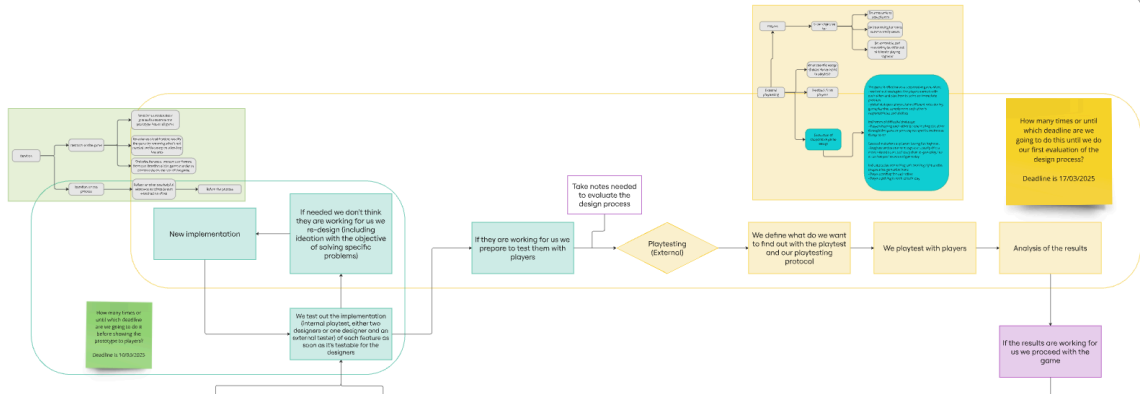
Is our prototype enabling cooperative gameplay? And is the game enjoyable by players of different skill levels?

## ▼ Playtesting Protocol

When writing our playtesting protocol, we have to keep in mind that our thesis' research question is:

*How can insights from an initial game prototype be used to refine not just the final game outcome, but also the design process and development methodology?*

In this context, our reports will have to include how we used playtesting as part of our design process and development methodology.



We will use playtesting for three main goals:

- Verify that our prototype works as intended

- Verify that each feature works as intended
- Verify the usability and accessibility of our prototype
  - Report any bugs
  - Report any needed improvements with regard to usability and accessibility (this is where player feedback is key) → We will use them to iterate on our prototype's design.
- Verify if our prototype is enabling cooperative gameplay
  - Use Aghabeigi (2006)'s cooperative performance metrics to verify how well our game enables cooperative gameplay → Identify areas of improvement and make suggestions to iterate on our prototype's design.

The information gathered through the playtests, once the prototyping phase is finished, will be used for:

- Writing our reflections and report on the design and development process.
- Inform our design of the final game.

### ▼ When do we playtest?

- Week 10: one designer playtests by playing with one external players while another team member observes
- March 6th and Week 11: one designer observes two external players playing

### ▼ Where do we playtest?

- At ITU/DR
- In one of the smaller rooms, except for the 6th when we playtest with everyone else

### ▼ What do we need to playtest?

- A build
- Snacks and water
- A computer
- Two controllers

- For week 10 we need two team members, for week 11 one is enough
- External playtesters
- Something to record the screen
- Something to take notes on

## ▼ Who do we playtest with aka who are the external playtesters?

- Since it's a local co-op game: players who would seek out co-op games with someone they know (friends, family, partners, etc.)
- For the question "Is the prototype working as intended?" it can be played with one team member and one external playtester
- For the question "Is our prototype enabling cooperative gameplay? And is the game enjoyable by players of different skill levels?" it should be played by two external playtesters
- For the question "Is the prototype usable and accessible?" it can be played with one team member and one external playtester, but better with two external playtesters

## ▼ What are the objectives of our playtests?

Our objectives for the playtesting are:

- Is the game enjoyable by players of different skill levels?

	Beginner	Medium experience	Expert
Beginner	two beginners	one beginner and one medium experience	one beginner and one expert
Medium experience	-	two medium experience	one medium experience and one expert
Expert	-	-	two experts

Each playtesting session should focus only on one of these three areas, because otherwise the quality of the notes and feedback will decrease (Huguenin, 2018). We should not conflate "working as intended" with "usable and accessible", because design mistakes can decrease usability

and accessibility even if technically they are working as intended, and we might have to consider whether to correct them or not (in the vast majority of cases we should).

Objectives	Ideal number of playtesters required
<p><b>Is the prototype working as intended?</b></p> <ul style="list-style-type: none"> <li>- Can each player use the Targeting for the Target Mode of their ability?</li> <li>- Can each player use the Target mode of their ability?</li> <li>- Can each player use the AOE mode of their ability?</li> <li>- Can each player change the force strength of the Target mode?</li> <li>- Can each player change the force strength of the AOE mode?</li> <li>- When a Movable Object is placed on a Pressure Plate does linked Unlockable Door open? Does this happen in every instance of this interaction?</li> <li>- Can the Sliding Door be Pushed or Pulled open by the players? Does this happen in every instance of this interaction?</li> <li>- Can the Sliding Door be blocked open with a Movable Object? Does this happen in every instance of this interaction?</li> <li>- Is it impossible for both players to get past the Sliding Door without blocking it open? Does this happen in every instance of this interaction?</li> <li>- Can the players Push and Pull the Movable Platforms? Does this happen in every instance of this interaction?</li> <li>- Is it possible for the players to Push or Pull the Movable Platform? Does this happen in every instance of this interaction?</li> <li>- Is it impossible for the players to Push or Pull the Movable Platform they are standing on? Does this happen in every instance of this interaction?</li> <li>- Can the players make a Movable Object bounce on a boucy surface so that the other player can Push/Pull it in mid-air to get to the other side?</li> <li>- Can the players jump? Is it possible for the players to reach with the jump the areas that are supposed to be able to reach with the jump and not in areas they are not supposed to?</li> </ul>	<p>5 external playtesters (Huguenin, 2018)</p>

<p><b>Is the prototype usable and accessible?</b></p> <ul style="list-style-type: none"> <li>- Are the controls intuitive to understand?</li> <li>- Are the controls easy to use?</li> <li>- Ask players to give feedback on the the usability of each feature: Targeting, Target Push, Target Pull, AOE Push, AOE Pull, jump, movement</li> <li>- Ask players to to give feedback on how easy or difficult it was to collaborate to do what they wanted to do</li> <li>- Check for the ease of understanding of each level's sections</li> </ul>	<p>5 external playtesters (Huguenin, 2018)</p>
<p><b>Is our prototype enabling cooperative gameplay? (see )</b></p> <ul style="list-style-type: none"> <li>• Positive: The playtesters laugh at the same time to a specific game event. (ignore instances were only one player laughs without the others)</li> <li>• Positive: The playtesters express verbally that they are enjoying the game.</li> <li>• Positive: The playtesters show facial expressions and other nonverbal behaviours that clearly express happiness or excitement.</li> <li>• Positive: The playtesters talk aloud about solving a shared challenge.</li> <li>• Positive: The playtesters navigate the world while consulting with each other.</li> <li>• Positive: The playtesters show pre-planned gameplay behaviour that emerges in similar cases.</li> <li>• Negative: The playtesters talked about controllers, and how one can use the game mechanics.</li> <li>• Positive or Negative (contextual): The playtesters told each other the correct way of passing a shared obstacle.</li> <li>• Positive or Negative (contextual): The playtesters saved and rescued the other player while they were failing.</li> <li>• Positive: Playtesters take different roles during gameplay that complement each others' responsibilities and abilities</li> <li>• Negative: One playtester waits for the other to catch up</li> <li>• Negative: One playtester leads and the other lags behind.</li> <li>• Negative: One playtester wants to do an action and the other wants to take a different actions, and by taking these actions they interfere or hinder each other's goals</li> </ul> <p>→ for each of these we want to also want to know which feature or level section has caused the event</p>	<p>20 external playtesters (Huguenin, 2018)</p>

**Is the game enjoyable by players of different skill levels?**

- for the positive metrics of the cooperative gameplay performance, also note down the skill level of the players

## ▼ What methods do we use for our playtests?

[write here later a summary of )

## ▼ Metrics for the observations

### ▼ Is the prototype working as intended?

[https://docs.google.com/spreadsheets/d/1hPPmC4-IL8bKR6cG6Nh\\_TEZciAnIjyj26IW29hH3fJk/preview?usp=sharing](https://docs.google.com/spreadsheets/d/1hPPmC4-IL8bKR6cG6Nh_TEZciAnIjyj26IW29hH3fJk/preview?usp=sharing)


### ▼ Is our prototype enabling cooperative gameplay? And is the game enjoyable by players of different skill levels?

<https://docs.google.com/spreadsheets/d/1jR7CwrvFUB1eL8KfkEb27G3-FCCJshvZVvsJ9sAwPVk/preview?usp=sharing>

## ▼ Questionnaires

### ▼ Is the prototype usable and accessible?

Push&Pull Co-Op Game - Usability & Accessibility

 [https://docs.google.com/forms/d/1QsygUrGDKpIX-WZWDLN7pfckw9\\_6ktUgZZWaYw\\_pU/edit](https://docs.google.com/forms/d/1QsygUrGDKpIX-WZWDLN7pfckw9_6ktUgZZWaYw_pU/edit)

## ▼ Questions for the interviews

### ▼ Is our prototype enabling cooperative gameplay? And is the game enjoyable by players of different skill levels?

Ask:

- Where there any specific moments where you remember working together to get past an obstacle?
- Where there any specific moments where you could have overcome an obstacle just as easily on your own?
- Where there any specific moments where it felt like playing with someone else was in the way of getting past an obstacle?
- [Also ask any follow up question you might have to what you observed
- Is there any particular part of the game that you liked?
- Is there any particular part of the game that you disliked?
- Do you have any feedback about how you think the game could be improved to be more fun to play together with another player (for you)?

## ▼ Playtesting Reports (Mirza-Babel, 2018)

### ▼ Report #1 - Example template

#### ▼ Objective(s) of the playtest

[which ones of the main areas of objectives did you focus on? ]

#### ▼ Study design and methods

(use the notes from and explain wich ones of did you use)

#### ▼ Summary (and severity of the findings if relevant)

(very short recap of the answers to the questions, additionally report any bugs or issues)

#### ▼ Description of the findings (what did you find out, analysis, if you have suggestions for changing, adding, removing, or fixing something make them and explain why)

(go as much into detail as you want about your findings)

#### ▼ Next steps (recommendations for future sessions)

## ▼ Bibliography

[https://drive.google.com/file/d/1NaCLtyiQ1yO-Mbyb75xt8eZPD99QKD Gh/view?usp=drive\\_link](https://drive.google.com/file/d/1NaCLtyiQ1yO-Mbyb75xt8eZPD99QKD Gh/view?usp=drive_link)

[https://drive.google.com/file/d/1RV929kEGQcm\\_MdQFKTKQGn-kuwgeJUj0/view?usp=sharing](https://drive.google.com/file/d/1RV929kEGQcm_MdQFKTKQGn-kuwgeJUj0/view?usp=sharing)

[https://drive.google.com/file/d/1u3Jw-Q2SeeRpY1G\\_iek2ffMa7TxNGOBt/view?usp=drive\\_link](https://drive.google.com/file/d/1u3Jw-Q2SeeRpY1G_iek2ffMa7TxNGOBt/view?usp=drive_link)

[https://drive.google.com/file/d/1dkeNmcdp87cSL1Ra6R9rbOPF6OO06LP-/view?usp=drive\\_link](https://drive.google.com/file/d/1dkeNmcdp87cSL1Ra6R9rbOPF6OO06LP-/view?usp=drive_link)

<https://drive.google.com/file/d/1Nd4INB5wQWkcltZS2jDsSpcRFgE0MIqO/view?usp=sharing>

<https://drive.google.com/file/d/12Jm1LDCXgAzTAnZPZVrXuLcgpT4564ih/view?usp=sharing>