Metabeta test documentation

Gameplay test

This section will cover our gameplay testing. This includes information about how we conducted the tests, how data was gathered, and how data was processed.

The gameplay tests were conducted over a period of two weeks on two different builds of the game. The tests were done before any usability testing. It was a deliberate choice, to make sure in the first place that the game had the correct play experience that we aimed for, and that later we could adjust the accessibility of the game accordingly. Since the accessibility design was not in the focus during the gameplay test, the participants were informed about the controls and were given hints during the testing in case the players would get stuck, so as to make sure they could progress and playtest all the levels of the current game. A questionnaire was written to help interview the participants. The tests took about 10-20 minutes with a follow up interview. In order to document the outcomes of these tests and pass on the information to the other members of the team, summaries have been written. They were based upon the recordings of the playtests, the observations and the interviews.

We had five gameplay testers, four male and one female, in the age range of 25 to 31 years. Three of the testers playtested the first version of the build and two testers playtested the build, that was modified according to the feedback received from the first three testers. All of the testers were asked to verbalize their thoughts during the time of play and they were followed up with the observatory questions from the test facilitator.

Game tester 1 – Build 1

Gender: Male

Age: 25

Background: This game tester had a lot of gaming experience and his favorite game genre is puzzle games.

Feedback - More detail in summaries

Positive feedback	Suggestions
● Liked the reverse gravity mechanic ● Good challenge ● Fitting visuals ● Understood the game	● Found the panels controls hard to understand ● Trouble with orientation ● Wanted reverse gravity as a global control rather than fixed on panels

Game tester 2 – Build 1/Editor

Gender: Male

Age: 25

Background: This game tester was accompanying game tester 1. He overlooked how game tester 1 played the game and was therefore testing in Unreal Engine editor while game tester 1 was being interviewed. The editor contained some different textures, but was otherwise the same as the build. This tester also had a lot of prior gaming experience, but wasn't as excited about puzzle games as game tester 1.

Feedback - More detail in summaries

Positive feedback	Suggestions
● Liked translation mechanic ● Thought that the game was a good challenge and had a good learning experience ● Fitting visuals	● Scale mechanic was confusing ● Need for a feedback at winning stage (sound) ● Need for a better space overview for orientation ● Wanted all mechanics on all objects

Game tester 3 – Build 1

Gender: Male

Age: 31

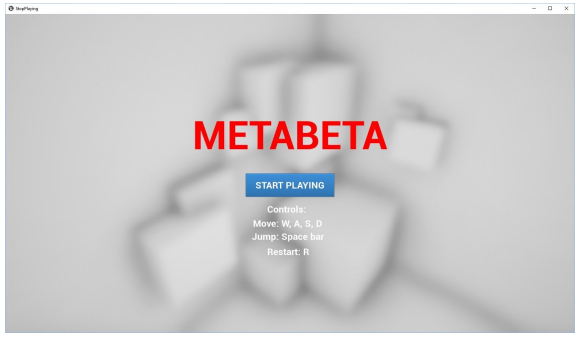
Background: The tester has the least gaming experience of the five gameplay testers. He hadn't played games for some years, but had tried Portal.

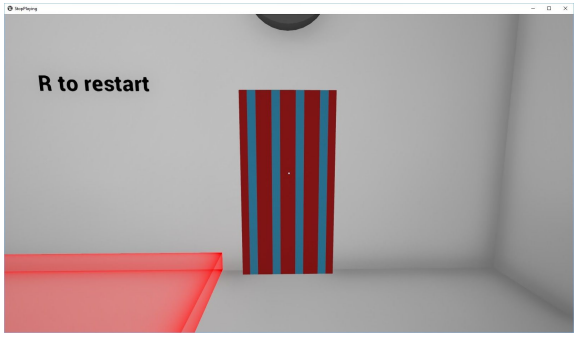
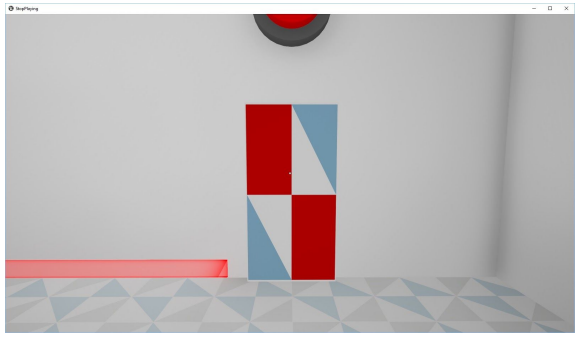
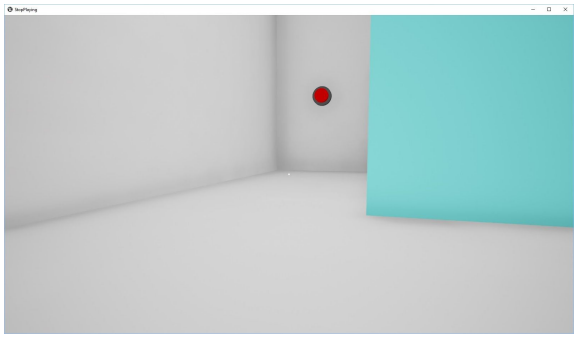
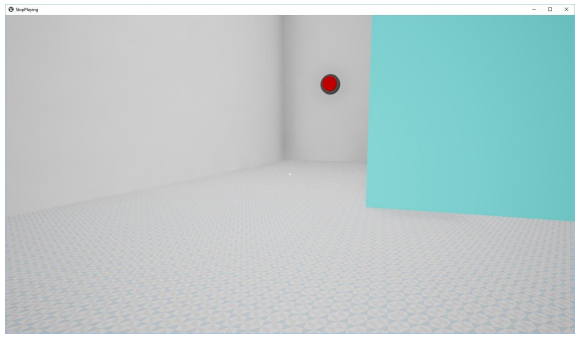
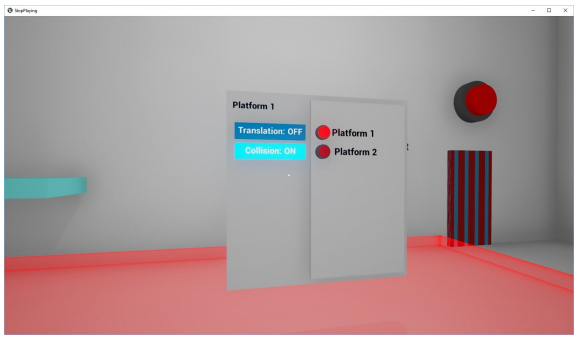
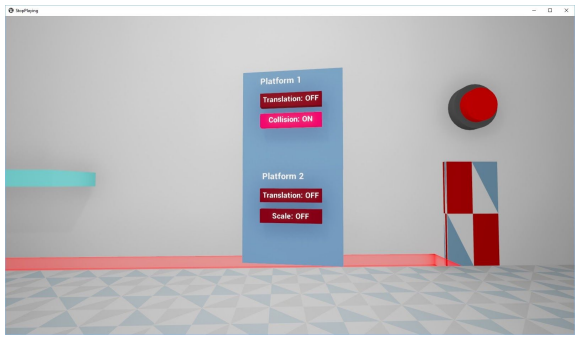
Feedback - More detail in summaries

Positive feedback	Suggestions
• Liked the pick-up ball mechanic • Visuals were fitting • Liked the uniqueness of the game	• Bad learning experience because of the panels controls • Wanted a win sound • Wanted better overview • Understood the game but quit out of frustration

Changes for build 2

After conducting the first three tests, changes to the game were made according to the received feedback. Some general bug fixing was made towards the ball and the last two levels. The colors of the panels were changed so they were easier to see, and some buttons got removed, to make every buttons function understandable and relevant. The floor in the game got a texture for a better depth perception and the control instructions were added to the start screen. The spawn point in level 6 was changed to give a better overview. Doors also got a new texture to align the art style of the game.

Build 1	**Build 2**
No control guide 	Control guide 

Door 	Door 
No floor texture 	Floor texture 
Panels 	Panels 

Game tester 4 – Build 2

Gender: Female

Age: 26

Background: This game tester is an experienced gamer and usually plays openworld games, but also likes the occasional puzzle games. She is a student at ITU and attends the digital media and design bachelor.

Feedback - More detail in summaries

Positive feedback	Suggestions
● Liked the pick-up ball mechanic ● Satisfying to shoot the ball ● The complexity made it satisfying ● Understood mechanics and how to solve the puzzles ● Liked the learning experience	● More feedback from rotation mechanic ● Panels were hard to see ● Didn't know if stuff were solid or not, because of the collision mechanic ● The walls gave bad depth perception

Game tester 5 – Build 2

Gender: Male

Age: 26

Background: This play tester usually doesn't play computer games, but had played a lot of computer when he was younger. He likes puzzle games as long as they aren't too hard.

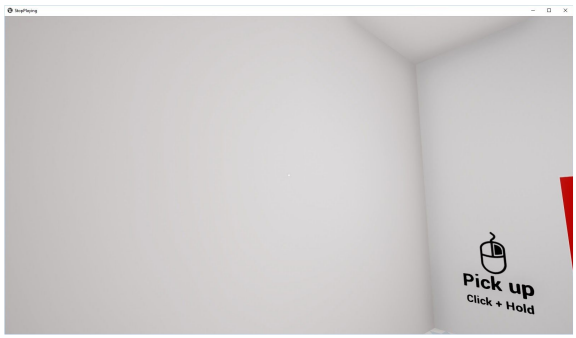
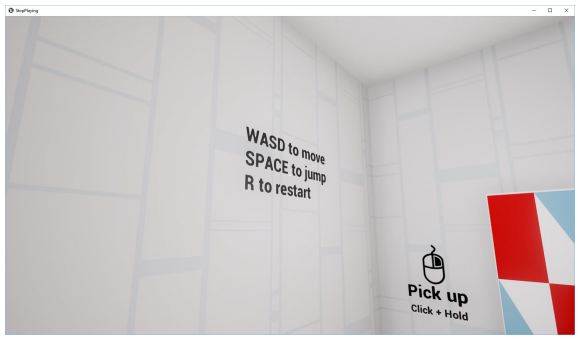
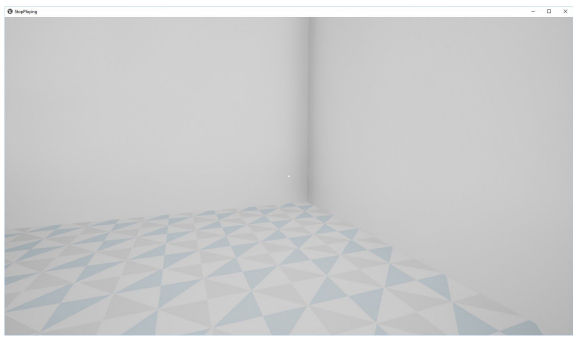
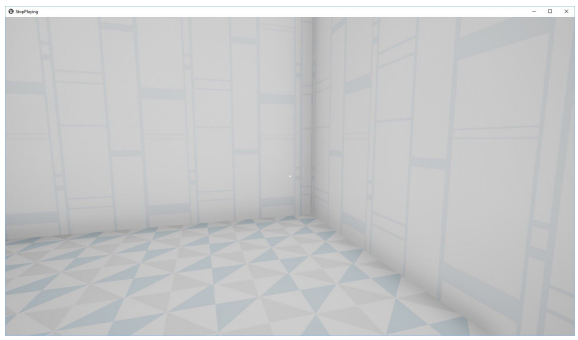
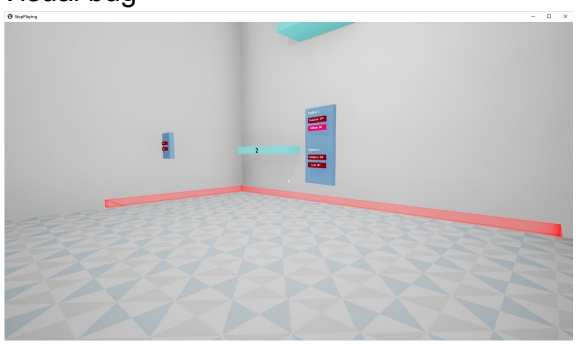
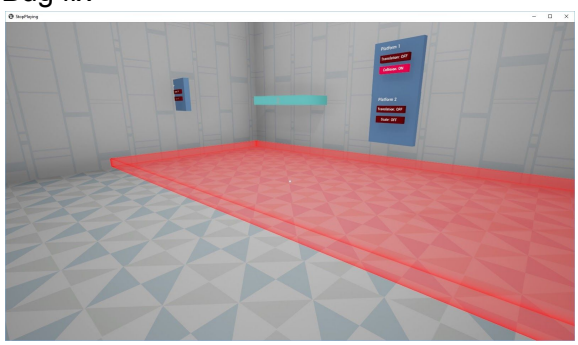
Feedback - More detail in summaries

Positive feedback	Suggestions
● Liked reverse gravity ● Liked gravity in general ● Levels were frustrating but that made them more satisfying to complete ● Understood how to solve the puzzles ● Good learning experience ● Fitting visuals ● Want more levels	● Orientation made mechanics too difficult sometimes ● If people got stuck they should be able to get a hint

Changes for build 3

The following changes are also based on feedback from our accessibility test since the tests were quite close and therefore went into the same build, which was the build we showed at demo night.

Wall textures were added to give better depth perception to the player. A visual control guide was added at the first tutorial level of the game, since most people didn't read the instructions at the 'start game' screen. Some minor bug fixes were made. The mechanic that removed collision has been removed from most levels since it mostly was confusing and you were forced to restart the level, because the ball didn't respawn.

Build 2	Build 3
No visual control 	Visual control 
Walls 	Walls 
Visual bug 	Bug fix 

Accessibility testing

This test was conducted with three testers, and was tested on ‘build 2’. We put the computer in front of them with the game at the start screen without an introduction to how they should play or what the goal of the game is. We would only give them hints if they asked for it and we asked them to verbalize their thoughts during the gameplay. We only made them test the four first levels of the game, because these are the tutorial levels. If the tutorial levels didn’t explain the mechanics and goal of the game, we would have to change them to make the game more accessible. By giving no directions of how to start and play the game, we expected to see if the game facilitates the required guidance itself. By not saying anything and staying as quiet as possible we could observe when and where the players would get stuck. The accessibility test had no questionnaire, the data was gathered by observing and recording the session.

The three testers were 19, 51, and 55 years old. They have either no or extremely little experience with gaming prior to this test. We figured that if they could handle and understand the game, it would be likely that most player could. The testers were recruited through the facilitators family.

Changes for build 3

The two elderly testers were so inexperienced with gaming that the main challenge for them was to understand the controls. They needed a lot of hints and guidance, and also had some of the same issues as the gameplay testers: the depth of perception(due to the lack of textures). They didn’t read the instructions to the controls on the start screen, therefore those were implemented within the game in the first tutorial level. Although the young tester didn’t need any guidance,he still needed clarification for the names of the buttons, that use game engine specific terms. By simply being more curious than the elderly testers, he managed to complete the four levels rather quickly, and he fully understood the mechanics and the controls. He did read the instructions of the controls on the start screen, but had the same problem with the depth perception.

The changes that were made in regards to the accessibility test feedback are included in the ‘changes for build 3’ in the gameplay testing section of this paper.

Discussion

Throughout our testing we have made a lot of clear mistakes. To reference our teachings we quote the book in what we did wrong in our tests.

“The right people to test for accessibility are:
· Part of your target market
· Objective (not friends or relatives)
· Those who have never played your game” [Game Design Workshop, Chapter 11]

The selection of our accessibility testers was not ideal as three of the testers were relatives and might not have been completely honest with what could have been wrong, since they might be afraid to hurt the feelings of the conductor or they could be too harsh to make sure we made the right choices. They were not part of our main target group either, which are players who have experience and interest in puzzle games. As they have never played computer games, the design focus for usability for this group might be very different from our target group due to differences in their perception of the game control and previous game knowledge. According to that data, if we were to design our usability for this group we should have made an extra five levels only to explain how to control the game via the mouse and keyboard. Also, people with more gaming experience might have an easier time understanding the game engine terms used in the mechanics like translation and scale. The only thing done right here is that they had never played our game. We tried to remedy these mistakes by filtering the feedbacks, focusing on those that could also affect our target group, and only make changes that do not damage the experience of our target group. Later on we did discover that after the changes, the game can be very user friendly towards a curious soul, and some of their troubles could confirm the same issues that the more experienced testers had. At the demo night the changes based upon the feedback from the accessibility testers had improved the game to the point, where no one had issues with understanding the controls of the game. So even though we didn’t do a textbook style usability test, we could find some feedback that was helpful and relevant.

“Write a script for a usability test in which you focus on critical tasks like starting a game, understanding objectives, making key choices, etc.” [SOURCE]

There was no written script for our accessibility test and the ‘script’ used for our gameplay testers was a questionnaire. We had prepared some assignments for our gameplay testers, e.g. that they should try to play the last two levels again, in order to see if they could find any bugs or if they could find another way to complete a level. In order to achieve this goal we should have made a special build of the game, where we could focus only on the part of usability we want to test. We could also design some tasks to test if the player really understand the control and the mechanics). At the moment of testing it was required to start the game from level 1, and therefore the tests would have taken longer than 20 mins, which was already a lot for some testers. The accessibility testers were too inexperienced to handle any specifics within the game, so potentially valuable data has been lost due to this.

Luckily the game testers were experienced enough to both give feedback based on their game experience, but also on some critical accessibility issues with our panels and controls.

After conducting these tests it's clear that more than one facilitator is obligatory. It's hard to take the observation notes, record, give hints, answer ongoing questions, with more than one tester at a time.With more facilitators more observations could be noted, more perspectives could be added to the feedback and more testers could test the game at the same time. We could have video recorded all gameplay testing and voice recorded all interviews, and reduced the loss of data when shared within the group.

However we have also done several things with the testing right. The scope of testers have been at a comfortable level, with not too many or too few. By having about eight testers the data that needed to be processed was perfectly fit for one designated person. This resulted in clear communication from testing data to group feedback, and the most essential bugs, annoyances, and other changes were caught and then improved/changed. By having one designated facilitator to make sure these tests took place, we were able to test while developing. Even though that some data has been lost due to few facilitators a lot of data is still preserved through the conducted recordings and several summary documents.

Overall we still got some incredible data and feedback. Bugs were discovered and fixed, design flaws were corrected, and the overall gaming experience have been improved a lot from the first test to the last. Even though the tests weren't optimal, the feedback was invaluable and absolutely beneficial.