Introduction

Young-adults with cerebral palsy (CP) often experience a decrease in physical activity as they reach adolescence and consequently an increase in social isolation from their peers. This decrease in physical activity negatively impacts both their physical and social well being. The project CP Fit ‘n’ Fun aims to improve the physical fitness and social immersion of youth with CP by developing multiplayer exergames that will allow youth to exercise in a fun social environment.

To achieve this goal, the game Liberi was developed. Liberi is a distributed, persistent, multiplayer exergame where players can engage in a variety of activities from building to fighting all while exercising on a stationary bike. The Liberi world consists of a set of mini-games that can be accessed from the main world. Because of the persistent nature of this world, where players can just come and go as they please, there is no guarantee of how many players will be playing at any given time. The mini-games can be played with any number of players so that, even if there were only one player in the world, he/she could still play the mini-games. This creates a problem with setting the difficulty of the mini-games though.

The goal when designing game difficulty is to balance the different elements of the game so that the player is neither board nor frustrated with the experience. If the game is too hard or too easy the player will not enjoy the game as much and be less inclined to play it. Because the Liberi mini-games can have any number of players, coming up with a difficulty level that will be fun for all amounts is a challenge. If just one difficulty level is set then this difficulty level will be ideal in only one situation and it will either be too hard or too easy in others. For example, if Liberi had a mini-game where players were trying to fill a bucket with water by transferring the
water from a far off lake, this task would be much easier to complete with four players carrying water than with just one player. The four players would have the capability to transfer water four times as fast as one player if all the game elements were the same. In this situation four players might find the mini-game too easy, but when the game is played with one player, this player might find it too hard. This scenario is clearly less than ideal; we would like players, regardless of their number, to have the same amount of fun when they play the mini-game. If the games are not fun and immersive for all different numbers of players then the players will be less likely to want to play the games and therefore will not receive the social or physical benefits of playing. The goal of this project is to come up with a way for various numbers of players to have the same amount of fun with the same mini-game regardless of the group’s size.

**Background Research**

Previous research by Bailey and Katchabaw in their paper, An Experimental Testbed To Enable Autodynamic Difficulty In Modern Video Games, listed four types of attributes that can be modified to make a game either easier or harder. These attributes are: player character attributes, non-player character attributes, game world and level attributes, and puzzle and obstacle attributes. Player character attributes are any qualities that relate to the playable character or characters in the game. These can be changed so that the player’s character is stronger, moves faster, jumps higher, has more health, or does more damage when they attack. These attributes can also be adjusted in the opposite direction to make the game harder. Another thing that can be adjusted to change the game’s overall difficulty is non-player character attributes. These attributes can be changed as before to make the non-players faster, stronger, or do more damage when they attack. A key difference though between adjusting player and non-player attributes is that in non-players the decision making processes can also be changed. For example, aiming can be adjusted so that attacks are less successful. The next
kind of attributes that can be changed are game world and level ones. This would include both the structure of the levels and their contents. These attributes are closely tied to the type of game being played. So for example, to make a platformer game like little big planet or super meat boy easier, you could reduce the amount of gap between jumps. Whereas in a more shooter type game like Halo or Uncharted, you might adjust the amount of cover available to the player. The last type of attributes used to adjust difficulty are puzzle and obstacle ones. These types of attributes can be adjusted to make the puzzle easier, by placing the solution closer to the puzzle itself, or harder, by placing the solution further away from the puzzle. This could be done by reducing the number of tiles on a sliding puzzle or in a game like sudoku, by providing more of the solution at the onset.

Research has also been done on qualities games should have to fully engage the player. These qualities are as follows: Games should provide challenges that match the players’ skill levels, games should provide different levels of challenge for different players, games should provide a level of challenge that increases as the player progresses through the game, and games should provide new challenges at an appropriate pace. These are the recommendations outlined in the GameFlow model created by Sweetser and Wyeth (Sweetser & Wyeth, 2005). The goal of adjusting all the various difficulty elements described above is to maintain this GameFlow model so that player enjoyment can be maintained. As stated in the paper “If the challenges are greater than the [player’s] skills, the result is anxiety; if the challenges are less than the [player’s] skills, the result is apathy”. The GameFlow model with its sets of goals and requirements seeks to keep the player engaged and immersed in the game and consequently enjoying it.

Method
In order to research difficulty scaling in variable player situations a mini-game was developed that had a balancing paradigm based on the perceived advantages and disadvantages of playing in each player situation. The developed game was then evaluated to determine the success of the balancing paradigm and any general player balancing principles that could be extrapolated from the scenario.

The Game

Turret Tumble is a variable player exergame which can be played with one, two, or four players. Players play as knights attacking a castle and their objective is to knock down the castle wall and to not be killed by the rocks dropped from above by the guards of the castle (fig. 1). Players attack the castle by collecting the rocks that are dropped by the guards and launching them back at the wall or guards using catapults placed on the ground. If a guard is hit with a rock launched from a catapult, it is killed and if the wall is hit by a rock it is damaged a bit. Players must hit the wall multiple times in order to win the game. If a player is hit by a rock dropped by the guards, he/she loses a life. If one of the catapults is hit by a falling rock then it is broken. Broken catapults can be repaired by players. Players can only hold a set amount of rocks in their inventory at a time and only have a set amount of lives before they die. The game is over when either the wall is knocked down or all the players are dead at the same time.
The Difficulty Paradigm

Adjusting the four different kinds of attributes to make a game easier or harder in a static player number situation is relatively straightforward; the easier games have easier adjustments and the harder games have harder ones. Figuring out when to adjust them in variable player situations is much more of a challenge. In order to determine a good difficulty scaling paradigm for Turret Tumble a list of perceived advantages and disadvantages of each player scenario was developed.

Single player has the most obvious disadvantages. In single player every game task must be completed by one person. In Turret Tumble, one player will have to manage catapult repairs, collect rocks, shoot the wall, and shoot guards. So in this situation the winning condition is much harder with only one person available to shoot rocks at any one time. The losing condition is also much easier with only one player needing to be killed by the guards before the
game is over. There is one advantage to single player though, which is complete control over
the actions of the game. In single player there is no need to coordinate with other players which
allows the player to know exactly what is going on at any given time.

When gameplay involves growing numbers of players you have to take into account
factors such as the division of labor among players and information complexity. Multiple players
have the advantage of being able to divide up game tasks into separate jobs and then assign
these jobs to individual players. This ability to strategize reduces the complexity of the game
by allowing each player to be focused on doing one task well, rather than many tasks. This
effectively ends up making the game easier the more players you have. In Turret Tumble this
could be done by having one player repairing catapults and another player shooting rocks. With
more players, though, the more complex the information in the game becomes. Players have
to adapt to more players actions making collaboration more difficult. Players may also have a
harder time identifying their avatar on the screen if the game displays character avatars as in
Turret Tumble. These factors make the generalization of more players make the game easier a
poor paradigm to use when designing game difficulty. From this information we can speculate
that a threshold on player number exists, where once we reach this threshold the confusion of
multiple players starts to take over and difficulty actually increases as more players are added.
In the situation of Turret Tumble, the highest amount of players the game goes up to is four so
this game is unlikely to see the threshold effect. Because of this, it has been assumed in our
balancing paradigm that playing the game with four players will be easier than playing the game
with two players. The last advantage of multiple players is that it is much harder for the losing
condition to occur. For each player added to the game another character is added to the list of
players the guards have to kill in order to win.

After evaluating the advantages and disadvantages of each player scenario a
balancing paradigm was created using this information. Because one player was seen to have
the most disadvantages of all the player modes it was given the easiest balancing values and because four player was seen as the easiest, it was given the hardest balancing values. Each player mode was given the same amount of guards and the same frequency of rocks being dropped. This was done to reduce the rocks available per player as the player number increased. Rocks, once they hit the ground would stay there for 15 seconds in one player mode, 10 seconds in two player mode, and 5 seconds in four player mode. This was also done to reduce the amount of resources per player as the player number increased. The last thing that was done to reduce the amount of resources as player number increased was the catapult number was kept at three. So in four player mode even though there were four players only three players could be shooting at a time. Also in one player mode if two catapults were broken the player would still have a catapult to use and wouldn't have to spend time repairing the broken ones. Four player was also made harder by requiring 20 shots at the wall to destroy it. Two player required 10 shots and one player only required 5. This was done to counteract the higher offensive capability of more players. Guards only required one shot to be killed in all modes of play because their presence was not a clear hindrance to gameplay as they provide rocks to shoot at the wall. Players were also given a rock inventory size of three which was the same across all player modes. This rock size was kept at three so as to not make one player mode too easy. If the rock inventory size was five or greater then a player in one player mode could simply collect five rocks, shoot them all from one catapult in quick succession, and win the game without ever having to worry about maintaining repaired catapults. The last difficulty factor that was changed was player life. In one player mode the player had to be hit five times before they died, in two player mode the player had to be hit three times, and in four player mode the player had to be hit two times.

After the paper prototype evaluation, this paradigm was modified to make one player mode slightly harder and four player mode slightly easier based on user feedback. Both one
player mode and four player mode were changed so that players would be allowed three
hits with a rock before they died. This changes the parameter so that it is the same as two
player mode. This was the only change that was made to the above paradigm before it was
implemented in the game. The final difficulty differences between one player, two player, and
four player are outlined in table 1.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>One Player</th>
<th>Two Player</th>
<th>Four Player</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guard Health</td>
<td>1 Shot</td>
<td>1 Shot</td>
<td>1 Shot</td>
</tr>
<tr>
<td>Wall Health</td>
<td>5 Shots</td>
<td>10 Shots</td>
<td>20 Shots</td>
</tr>
<tr>
<td>Catapult Number</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Player Life</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Rock Inventory Number</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Rock Time on Ground</td>
<td>15 Sec</td>
<td>10 Sec</td>
<td>5 Sec</td>
</tr>
<tr>
<td>Guard Spawn Time</td>
<td>20 Sec</td>
<td>20 Sec</td>
<td>20 Sec</td>
</tr>
<tr>
<td>Rock Drop Frequency</td>
<td>5-20 Secs</td>
<td>5-20 Secs</td>
<td>5-20 Secs</td>
</tr>
</tbody>
</table>

Table 1. An outline of the difficulty differences between player modes

Study

In order to test the effectiveness of the balancing paradigm a study was designed
to judge the perceivable difficulty differences between the one player, two player, and four
player versions. Two groups of four participants were recruited for the study for a total of eight
participants. Of these participants five were female and three were male. The average age of
a participant was 28 years old. The participants were prescreened for at least ten hours prior
experience using a game controller and ability to perform light exercise on a exercise bike. The
Physical Activity Readiness Questionnaire was used to screen participants whose health might
make exercise inadvisable. Participants were seated in the same room and each of the four
participants in each group had their own computer, xbox controller, and custom built cycling
ergometer.

After agreeing to participate in the study each person was given a background
questionnaire on their previous gaming and exercise experience. This was done to deduce patterns in the data or to account for possible statistical abnormalities if they occurred. Once the forms had been filled out participants received rule instruction and played practice versions of both one player and four player Turret Tumble. This was done to mitigate the learning effect and to allow players to start to develop strategies for playing with multiple players. The order that participants played the practice versions and the order that they played the actual trials were balanced for order.

After the practice version had been played and the participants were comfortable with the controls and objective of the game they played one of the three player situations. This player situation was played for a total of five minutes to allow the participants to get a good feel for the difficulty. If the player won or lost during that time the game would still continue but be reset so that they could keep playing (ie. If the player defeated the wall, the wall would go back to full health and the player could defeat it again).

At the end of the five minutes the player was asked to fill out a Likert scale questionnaire that gauged the difficulty, fun, and sense of satisfaction they had with that player version. The questions that were asked were as follows: ‘I had fun playing the game’, ‘I enjoyed the collaboration of the game’ (if multiple players), ‘I enjoyed playing alone’ (if single player), ‘I felt board’, ‘I was absorbed in the game’, ‘I found the game challenging’, ‘the game was physically tiring’, ‘The game required strategy’, ‘playing the game was gratifying’, and ‘Playing gave me a strong sense of accomplishment’. The first four questions of the questionnaire (including either the single player or multiplayer question) were used to gauge the level of fun of each version of the game. The next three questions were used to judge the difficulty of each game version. The last two questions were used to judge player satisfaction. Once participants filled out the questionnaire another trial would be run for a different player number and the questionnaire would be once again filled out for this trial. This would be again repeated until all three trials had
been run and the questionnaires filled out.

At the very end of the study an informal interview was run with each of the two groups of participants. In this interview general questions about the difficulty differences between games were asked and general comments were given by the participants. This informal interview allowed the participants to give any additional information they had that could not be captured on the Likert scale questionnaire. The equipment used in the experiment consisted of: four computers with four Xbox 360 wireless game controllers, three PC GamerBike minis, and one Tunturi bike. Besides the interview and Likert scale questionnaire participants score indicators were also recorded by the game.

Results

After running the studies the results of the Likert scale questions were tallied and averaged. In figure 2. you can see the results of the Likert scale questionnaires with difficulty, fun, and satisfaction charted against each player situation.
As indicated in this graph the difference between each of the player versions in difficulty, fun, and satisfaction was negligible. We can further see this when we run a two-way ANOVA analysis without replication with a significance level of .05. When this test is run our P value is 0.519 which is much greater than .05. The test statistic is also much less than the critical value (f = .776 and F-Critical = 6.944) so the null hypothesis is accepted and we conclude that there was no significant difference in difficulty, fun or satisfaction in each of the three player numbers. The difficulty balancing paradigm was not completely perfect, in which case there would be absolutely no slope in the lines, but it is pretty close to being completely unnoticeable.

The results from the informal interviews indicate that most participants did prefer two player to all the other player modes. It is suspected that this is because players enjoyed playing with others but that they found four player mode too confusing. As indicated by one participant “Four [player] was a little crowded”. It is interesting that the information complexity disadvantage of multiplayer starts to take place at such a low level as four players. The informal interviews also indicate that players did not perceive a difficulty difference between one player and two player mode but a few players were able to notice a difference between two player and four player mode. So to potentially improve the balancing paradigm four player mode could be adjusted in difficulty so that it is a little closer to that of two player mode. A surprising fact that rose from the data was that females reported more often that they cooperated together whereas almost all the males said that they had a hard time cooperating together. It would be interesting to see if these results are consistent with a larger sample size of males and females. It should also be noted that the results of the groups indicate that there was a learning effect while playing the trials so that whatever trial went last was the trial that the player most often found the easiest. This learning effect could possibly be mitigated in the future by giving participants
longer practice playing times.

From all the results we obtained from Turret Tumble’s balancing paradigm the following general player balancing principles could be extrapolated and be applied to the development of future variable player games. The more information you display about the various players the more complex the information becomes when more players are added to the game and consequently the harder the game becomes. This effect can take place at very low numbers of players as well. In Turret Tumble this effect was seen at only four players. Therefore when designing difficulty you should take into account how much information will be displayed to the user as player number increases and factor that into your estimation of difficulty. Another thing that should be kept in mind is the hidden advantage of single player, which is complete control. “When I was playing by myself I knew that I was going to get the stone” explained one participant. In single player mode the player does not have to work with anyone so can implement and modify his/her own strategies quickly and efficiently. No matter how well players work together as a team they will never be quite as fast at adapting strategies as a single player will be. This effect will be felt even more strongly if players are remote and not able to communicate with each other. So when designing a game you should take into account how easily players will be able to communicate. For example a game that has spoken word communication will be easier than a game that only has text communication, as most people can speak faster then they can type.

Conclusion

After evaluating the balancing paradigm used in Turret Tumble it was concluded that the chosen balancing method was near perfect. There were only a few modifications that could be made to improve the current balancing method. The statistical results showed no significant difference between each player version in terms of fun, difficulty, or satisfaction. From the
results of the evaluation we were able to come up with some general balancing principles to be used with variable player number.

The study though still leaves some unanswered questions. It would be valuable to learn if the results of the study would be consistent if a larger sample size was used. More groups would also make it easier to identify if there is in fact a learning effect going on in the data. It would also be useful to determine at what point information complexity starts to become a problem. It would be more practical to have a concrete boundary of where this occurs or at the very least a more definitive guide to use when trying to create a game difficulty paradigm. Further research could also be done to develop a guide that compared strategy development and modification in cooperative gameplay settings. This study could compare the success of strategies with various numbers of players and the success of strategies with various modes of communication (i.e., talking in the same room, talking over headsets, and text communication). An interesting statistic to collect from this study would be the amount of time it takes the group to switch from one strategy to another. This would give developers of cooperative games a better idea of the difficulty scaling that is required as their player numbers go up and the level of strategy skill and adaptiveness they can expect from remote players. This further research would clear up some of the unanswered questions of the study and give researchers a better idea of the difficulties and advantages of multiple players in a cooperative game. Variable game balancing, while a difficult problem, can be dealt with by carefully considering all the different advantages and disadvantages to each of the different player amounts. By better understanding these advantages and disadvantages we can create better balancing paradigms and ultimately make our games more fun.