A Study of Time Perception in Video Games through the Use of Timed Puzzles (May 2013)

Morgan Davis Author, Myque Ouellette

I. Abstract

This thesis explores how timers affect a player's perception of time while playing a game, and whether the timers add or detract from the fun and/or difficulty. Designers often use timers in puzzles or levels, but they do not always understand the implications of specific timers, and potentially fail in effectively communicating to the player.

To explore some practical methods when using timers, the test uses the board game, *Rush Hour* converted into a top-down level (made with the Unreal Development Kit) consisting of a tutorial and three puzzles. The test takes advantage of three timer types: a Normal Clock, a Seconds Only Clock, and a Draining Bar. The thesis consists of four tests: a control (no timers), and three experimental tests. In order to prevent order bias, each experimental test presents timer types in different orders. This allows each of the timer types to apply to each puzzle at least once.

The results of the study provided an interesting look into how timers affect a player's perception of time. Players, on average, consistently finished puzzles faster with timers, and more often than not, believed they had spent less time than they actually did. Whereas, the testers without timers believed time to move much slower. However, both groups found the puzzles fun and moderately difficult.

Index Terms—

Unit – Seconds, Minutes, Hours, Bars
Puzzle – Timer, Games, Pressure, Blocks, Top-down, Psychology, Shapes, Movement, Sounds

II. Introduction

All video games, in some way or another, attempt to convey information to the player. Through a Heads-Up Display (HUD), video games communicate various statistics (health, armor, ammunition). The statistics generally take the form of units, which could be feet, inches, seconds, minutes, bars, icons, and so on. These units help players measure their status, time, and progress. The concrete knowledge from a HUD provides designers a plethora of communication techniques. Ultimately, designers want to communicate to the player in a way that is clear yet still intriguing, immersive, and exciting. Otherwise, players grow bored or confused. Players perceive this information in unique ways, and unfortunately, designers do not always understand how to improve the communication.

Psychologically, units can trick players. One such phenomenon, the Unit Effect (proposed by M. Pandalaere, and B. Briers in the *Journal of Consumer Research* in 2011), occurs when players see two identical measurements portrayed in different units (e.g. 300 seconds versus 5 minutes), yet, perceive the measurements as different.

This study used the Unit Effect phenomenon to investigate player perception, in regards to time. One cannot compare two durations of time unless the durations are the same. If players know the durations are the same, then they logically believe their perceptions are the same (not the desired result). Using the Unit Effect as a way to control results, the test specifically looked into the following question:

"Do players perceive time as faster or slower while solving timed puzzles?"

The test also investigated the amount of fun a player has while doing a timed puzzle in contrast to a puzzle with no time limit. In the timed scenario, the differences in timer types (e.g. minutes, seconds, or an animating icon) could show changes in fun or pressure felt.

When people compare numbers (minutes versus seconds), they rarely do the hard math to calculate the significant differences, especially when given a task such as a puzzle. While occupied with a puzzle, a timer pressures the player into making quick, nervous actions, potentially shifting their perception of time. This occurs due to the impending failure that could result. Players constantly run into challenges in video games, and these challenges present tension that may cause time perception changes. Even with a timer in front of testers, which tells them how much time remains, their perception of time may move faster or slower. In contrast, results with no timers rely on the player's preference of video games or experience with the genre. A person who enjoys puzzle games might find the puzzles fun, while, a person who

Manuscript received September 27, 2012. **The Author is a student at Southern Methodist University's Guildhall, in Plano, TX (phone: 214-478-8138; e-mail: mpdavis@smu.edu).**

Myque Ouellette, is a design professor, teaching at Southern Methodist University's Guildhall in Plano, TX. (e-mail: myque@mail.smu.edu).

hates puzzle games but enjoys action, might prefer a timed experience. The study investigated the following questions:

"Do players have more fun during a timed challenge than a non-timed challenge?"

"Do players perceive time accurately while solving timed puzzles?"

"Do specific timer presentations change a player's perception of time more or less than another?" (Discovered through the Unit Effect)

In order to answer these questions, the study presented a test subject with a tutorial puzzle followed by three full puzzles. After learning the game's mechanics, the player began a top-down puzzle (An Unreal Developer Kit version of *Rush Hour* the board game, published by ThinkFun in 1996) with a timer at the top. Testers had 240 seconds for each puzzle (4 minutes). The different timer types displayed the 240 seconds in their prospective units (e.g. 4:00, 240 (seconds only), and a draining bar with no numbers). To avoid order bias, there were three versions of the experimental test, each with a different timer type on a given puzzle. In order to give testers incentive to succeed, the level contained a scoring system. Successfully completing a puzzle in the allotted time rewarded the player 100 points. If the tester ran out of time, then the timer reset. During the second attempt, the tester played for 50 points. If the tester failed a third time, the player received no points.

Before the test, testers took a pre-test survey to determine their playing styles (hardcore, casual), demographic information, and gaming preferences. After players finished all three puzzles, they filled out a survey, which asked questions concerning their perception of time while playing, how much fun they had, and the level of challenge they experienced.

III. Research Review

In order to understand the unit effect's use in video games, the study observed a number of recent games and articles.

1. *The Unit Effect and Player Perceptions,* Jamie Madigan, psychologyofgames.com, March 9, 2011 [1]

This article covers the central idea behind the Unit Effect. Numbers can look bigger or smaller than they really are, depending on the units used. Many business firms, marketing agencies, and advertisements use this technique to convince and deceive their viewer.

Consider the following numbers:

- 7 yards
- 21 feet
- 252 inches

The numbers measure up exactly, yet the similarities become less and less obvious. The study behind this article showed that people saw a "smaller difference between warranties lasting 7 and 9 years, than between warranties lasting 84 and 108 months." Even when the math is the same amount of time (84 months equals 7 years), the user does not always recognize this.

Additional studies concluded that larger numbers tend to exaggerate the unit effect. When comparing numbers, people tend to think the differences are larger if the numbers are large (10,000 to 20,000 versus 1 to 2). If people cannot quickly calculate even simple arithmetic when comparing units, then the effect ties in with pre-conceived assumptions about units. 21 feet might sound longer than 7 yards because the difference between feet and yards does not differ greatly, yet the difference between 7 and 21 does. Thus, going from 21 feet to 252 inches, the differences grew even more drastically. In speculation, it might have to do with people *visualizing* an inch versus a foot, instead of calculating it as 1/12 of a foot.

2. *Brain, Time Perception and Synesthesia,* PBS YouTube Kostas Partsinevelos September 2011 [2]

This YouTube video follows a study on Human Perception and other brain related functions. The interesting segment follows Dr. David Eagleman, a Neuroscientist researching if people truly see life in slow motion when they shift their perception of time. In order to achieve this, he drops people from a tower into a net. They fall backwards, which scares them so much that believe it took a long time before they hit the net.

In addition, he has them hold a device that flickers numbers on its screen at extremely fast speeds. If they truly slow down time, they could see the numbers. He found that they could not distinguish the numbers on the device's screen, concluding that perception shifts do not actually slow down visual movements. Instead, people merely "feel" as though time moves slower.

If time does not slow down visually during a time shift, then the same idea applies when a player looks at a HUD during a time shift. Instead, the test should show that players perceive time as moving faster during a time shift due to pressure of failure.

3. *Psychology Time Perception Experiment in Teenagers,* YouTube Daniel Hopkins, Monroe Technology Center, November 2010 [3]

Two students at Monroe Technology Center performed a study on time perception through the comparison of two tasks: watching Al Gore's Global Warming video and playing *Portal*. Students would play *Portal* and watch the Global warming video back to back. Afterwards, they answered the question, "How long did the video last?" and "How long did you play the game *Portal*?" The students accurately predicted the time allotted while playing *Portal*, but had trouble specifying the time after watching the Al Gore film. Many video games have their boring parts (walking across long stretches of land), which, might distort time perception and overall experience a player feels. Consequently, players could lose their patience when faced with boring tasks in a game, even if the game has interesting gameplay overall.

With this in mind, the control group of this study should replicate the results of the *Portal* responses. Players should perceive time accurately. However, the timed players should notice a shift in their perception of time due to the tension. Unlike the Monroe Technology Center, this study seeks to analyze direct effects of the timer and unit effect on time perception, rather than intentionally creating something less interesting, which would slow down the perception of time.

4. *Time Perception, Immersion and Music in Videogames*, Timothy Sanders, Paul Cairns, September 2010 [4]

This study looks at player immersion while players solve a maze with music in the background. The experiment's control group uses the same maze without music. They measured immersion using an immersion questionnaire with 31 questions answered on Likert scales. The primary question before the questionnaire being, "How much time do you estimate this maze to have taken?" Unfortunately, they found that music did not alter immersion or time perception while playing. Instead, immersion relied more on how much the player enjoyed the specific genre of the experiment.

The study proves useful because it suggests players who like puzzle games could perceive time differently (immersion) even without timers in the level.

5. *Mass Effect 3* (Multiplayer), Bioware, Edmonton Canada, March 2012 [5]

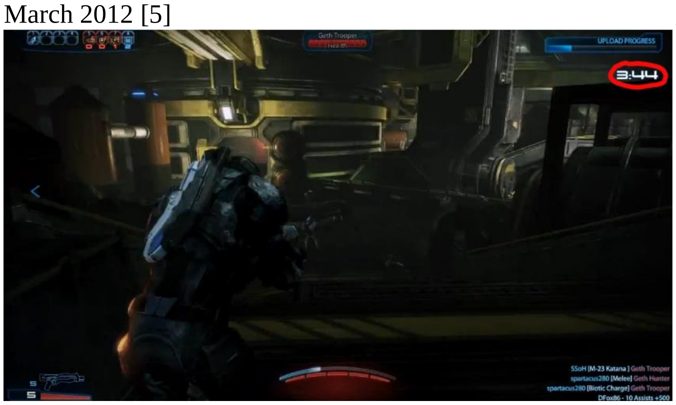

Figure 1 - Mass Effect 3 Multiplayer uses a traditional timer (top right)

Mass Effect 3 is a third person shooter set in a science fiction universe. In the multiplayer mode, players unite to fight waves of enemies until the players evacuate the area. Most waves consist of the player fighting and defeating all enemies to continue. However, some waves include timed challenges.

The challenges include players taking out various targets, or disabling devices before the timer runs out. *Mass Effect 3* uses a typical timer (e.g. 3:42 minutes/seconds) to convey the time remaining. The timer blinks red when it gets low, therefore, the communication is clearly given to the player that if they run out of time, they fail and must start over. The non-timed waves are just as fun because ultimately, players must survive ten total waves (both timed waves and regular waves). If the players fail to meet the criteria for any particular wave, they lose.

Knowing these details, one notices when playing the game that the perception of time moves quicker during the timed waves. Because of the tension, other players often encourage the team through the voice chat system, saying, "We need to complete the objectives now!" During the non-timed waves, players take their time to regroup and prepare for the harder waves. The experience changes based on the timers alone (presented in a traditional timer e.g. 3:20), which makes this study believe it is the most direct emotional connection to time perception and immersion.

6. *Shift*, ArmorGames by Tony [6]

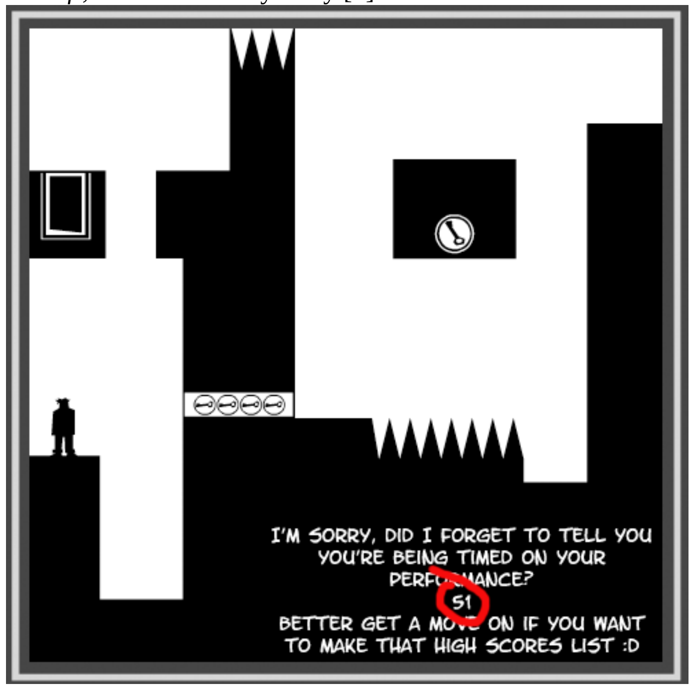

Figure 2 - Shift uses a seconds timer that goes up (higher means worse).

Shift is a simple 2-D platformer puzzle game in which the player platforms through levels that are strictly black and white. The black is solid to the player until the player "shifts" the level and it essentially flips the map upside down, making the whites solid to the player. Each time the player shifts the level, it creates new paths or platforms so the player can reach the end door (objective). The game features a Score that directly measures the amount of time the player took to beat the level. The longer the player spends trying to solve the puzzle, the higher their 'second count' is. The higher the number, the longer it took the player to navigate the level. Therefore, a low number constitutes a good score.

This same concept could help in this study, as far as incentive and punishment, with a few adjustments. For example, a timer displays to the player with approximately 3 minutes to solve the puzzle, with a reward of 100 points. If they do not finish the puzzle within the first 3 minutes of the puzzle, they now have the chance to play for 50 points, then 0 points. Each new chance nets the player a new timer reset back to 3 minutes.

7. *Super Mario World 2: Yoshi's Island*, SNES by Nintendo 1995 [7]

Figure 3 - This mini-game uses a draining bar timer.

Super Mario World 2: Yoshi's Island contains a mini-game puzzle accessible in the main world. In this mini-game realm, the player competes against a bandit. One of the mini-games called, "Tossing Balloons" challenges the player to hit the corresponding controls in order before a "bar timer" drains. The timer resets when the player hits a wrong button, but they only get three chances to do it correctly. After the player finishes their sequence, the computer/bandit attempts the same thing, and if successful tosses it back to the player. This continues until either the player or the bandit fails.

The "Bar timer" in this scenario is convenient because it allows the player to focus on their task while still knowing the amount of "time" left. Out of the corner of the player's eye, a draining bar remains legible. On the contrary, a textual reference of time such as "0:05" looks the same to the player as "0:01" when the player has to focus on the controls (see the picture above). Therefore, it is harder to pick out specifics.

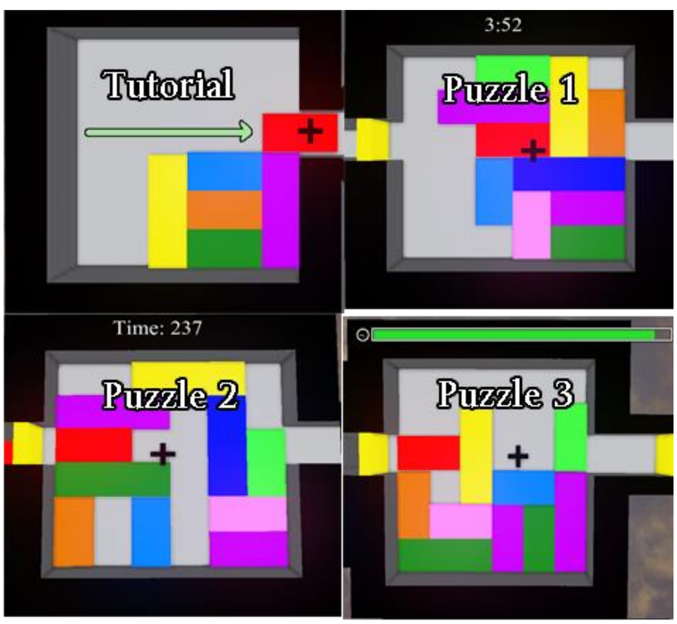

Figure 4 - Overview of Map Layout

	Timer Type:		
Test #:	A	B	C
1	Normal	Seconds Only	Bar
2	Bar	Normal	Seconds Only
3	Seconds Only	Bar	Normal
4	Control Group	Control Group	Control Group

Figure 5 - Table of Timer Organization

IV. Methodology

Immersion, time perception, and fun all tie in together. Players do not visually see time move slower, rather, emotions such as fear or stress affect time perception. Ultimately, the purpose of the thesis is to help designers improve the fun factor of their game by understanding changes in a player's perception of time and to utilize timers in more effective ways.

Therefore, this study investigated the following questions:

"Do players have more fun during a timed challenge than a non-timed challenge?"

"Do players perceive time as faster or slower while solving timed puzzles?"

If so,

"Do specific timer presentations change a player's perception of time more or less than another?" (Discovered using the Unit Effect)

The study relied on the use of varied timer presentations within a series of puzzles. To avoid order bias, there were three versions of the experimental test, each with a different timer type on a given puzzle. Thus, each puzzle gets all of the timer types when combining the tests together. The challenges remain the same in each puzzle. Players may perceive the challenge to be greater or lesser based on the timers. However, players who do not normally like logic puzzles may find that the puzzles' challenges create tension without the timers.

Before the test, testers took a pre-test that collected demographic and playing preferences information. In addition, the researcher instructed the tester and gave them an instructional sheet as they worked their way through the Tutorial puzzle. This ensured the tester did not jump ahead at any point, thereby giving the researcher full control of the test and minimizing any potential confusion of the interface and controls. During the test, subjects worked through the artifact, which consisted of a series of logic puzzles based on the board game *Rush Hour* (published by ThinkFun in 1996).

For the map of level, see Figure 3 above. Testers moved left to right, with each section containing a puzzle (T for tutorial, A, B,C for puzzles 1, 2, 3). Note that the game does not contain gravity, and blocks can only move in the direction that they point towards (e.g. a block stretched left to right can only move left to right, never up and down). If the tester failed to solve the puzzle, the timer reset, however, the blocks did not. In contrast, when the player moved the red block (objective) to the next section they successfully solved the puzzle. However, a wall prevented testers from continuing to the next puzzle. While, the tester then took a short questionnaire about that particular puzzle, the instructor prepared the tester's cursor for the next puzzle. The most important question on the questionnaire was:

"How long do you think it took to complete the puzzle?" (Note that "think" is a key word because this urges testers to perceive time rather than over-thinking it.)

Afterwards, the tester gained permission to start the next puzzle.

After completing the test, the testers took a survey exploring their perception of time, challenge, and fun factor overall. In order to compare the tester's thoughts to actual time spent playing, the researcher recorded the player's allotted time solving the level with the use of a stopwatch, pen, and paper. Examples of some of the survey's questions included:

- "In minutes and seconds, how long do you think it took to complete the entire level?" (this compares with the control group's total time)
- "Which puzzle was the hardest?"
- "Which puzzle was the easiest?"
- "Which puzzle provided you with the most time to solve?"
- "On a scale of 1-5, how fun would you rate the puzzle?"

The researcher hypothesized the following:

1. Perception of Time moves faster in the experimental groups than in the control groups (higher rate of immersion and fun)

2. The time estimates in the experimental group should consistently differ from those in the control group.

3. The prospective timers vary in time estimation (unit effect). Expected time estimations include (in minutes):

a. "Normal" Clock. Player accurately estimates the amount of time they spent on the puzzle with no difference between actual times and estimated times.

b. "Seconds Only" Clock. Player inaccurately under-estimates the amount of time they spent on the puzzle by one minute (on average).

c. "Draining Bar" Clock. Player inaccurately over-estimates the amount of time they spent on the puzzle by two minutes (on average).

The researcher selected testers on a convenience basis from the Dallas area. Testing took place in an empty test room on the Guildhall campus, using the researcher's laptop, mouse, and keyboard.

Using the Unreal Developer Kit provided a perfect platform for the study and ensured a clean test. The engine used simple BSP construction for quick construction, and Kismet scripting for smooth sequences. In order to record written responses, the researcher utilized Microsoft Excel.

In summary, time perception and immersion play a huge role in video games, and designers could use this study as helpful research regarding the perception of time and its effects on player experience.

V. Results

During the course of the study, 40 students from the Guildhall at SMU participated in the testing process (each of the 4 maps had 10 randomly chosen testers). All testers filled out a pre-test demographic survey before playing the level. Between the puzzles, each tester filled out a survey about the puzzle that they just solved (three total surveys) and after the entire level, testers filled out a post-test survey. Testing time took between 10-20 minutes, depending on the speed in which the tester completed each puzzle.

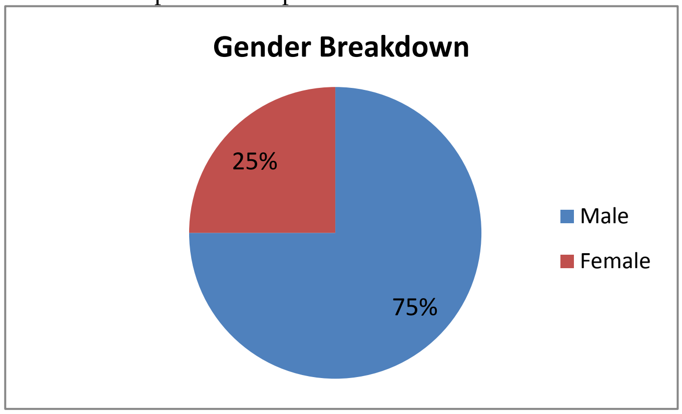

Figure 6 - Overall the Gender ratio was 1 to 4

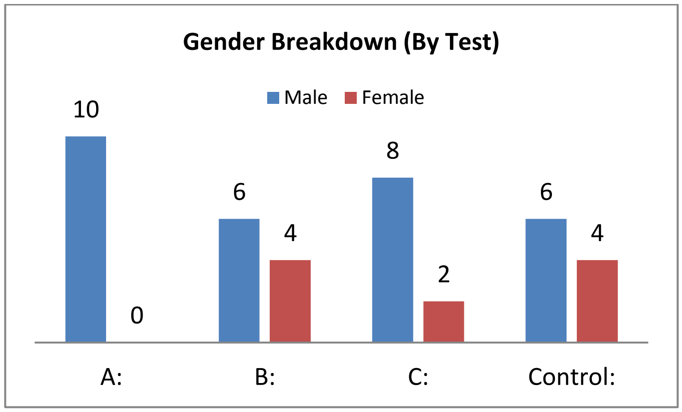

Figure 7 - Gender Breakdown (per test)

Although the test results split males/females 1 to 4 overall, Test A had only males. See Table 3 for a comparison between male and female times.

Fastest Time		
	Male	Female
P1:	1:24	1:37
P2:	1:01	1:50
P3:	0:53	1:08
Avg:	1:06	1:31
Slowest Time		
	Male	Female
P1:	5:44	7:33
P2:	11:16	7:25
P3:	4:45	2:44
Avg:	7:15	5:54
*Gave up at this point, could be longer		

Table 3 - Fastest/Slowest Times of Genders

The chart above shows the fastest and slowest time per puzzle for each gender. One female gave up on puzzle #2 at

7:25. Males had the fastest times, the biggest difference being 49 seconds on Puzzle #2. Males also had the slowest time at 11:16. On average, males had both the fastest times and slowest times (fastest average of 1:06, slowest being 7:15). This most likely occurred because the data compared 30 males to 10 females. Had the data included an equal distribution of genders, the numbers could change.

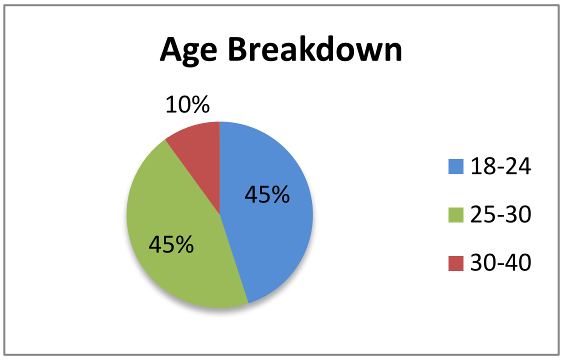

Figure 8 - The Age ranged between 18-40

The age of testers was mostly between 18-24 and 25-30. No testers over the age of 32 or under the age of 20 participated in the study. The results (actual times, average estimations) showed no major differences between the age groups.

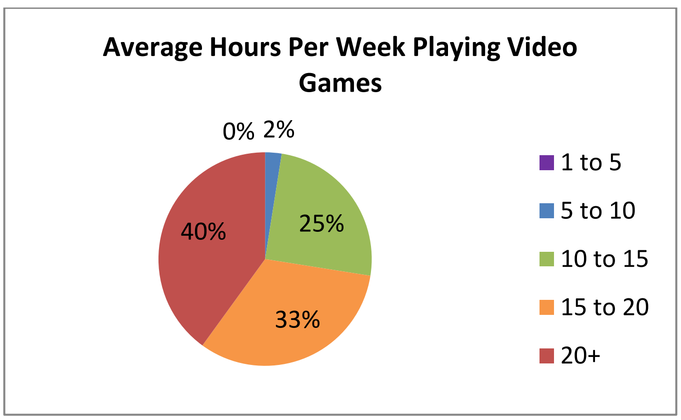

Figure 9 - The amount of time testers put into gaming

All testers claimed to have played video games prior to the test. Most of the testers identified themselves as having spent 20+ hours a week playing games. However, none of the testers played as little as 1-5 hours a week. Therefore, the test consists of persons familiar with video games, which means they should understand WASD controls and simple interactivity with objects.

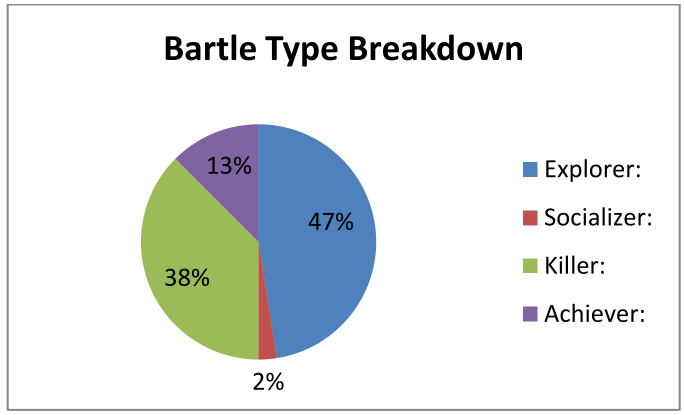

Figure 10 - Overall Bartle Breakdown

Bartle Types:	A	B	C	Control
Explorer:	5	4	4	6
Socializer:	0	0	1	0
Killer:	5	5	4	1
Achiever:	0	1	1	3

Testers identified with one of four Bartle types by circling a specific option to the question, "What do you like to do most when you play games?" The options included: Explore, Socialize, Fight, Get Achievements. Explorer types formed the majority of the test pool, while only one person of the forty claimed their type as Socializer. Only Test C consisted of all four types, and contained the one Socializer. Every test had one or more killers and explorers. Perhaps a connection is possible between the Bartle Types and the test results, but the researcher did not see it factor into the hypothesis given the small sample size.

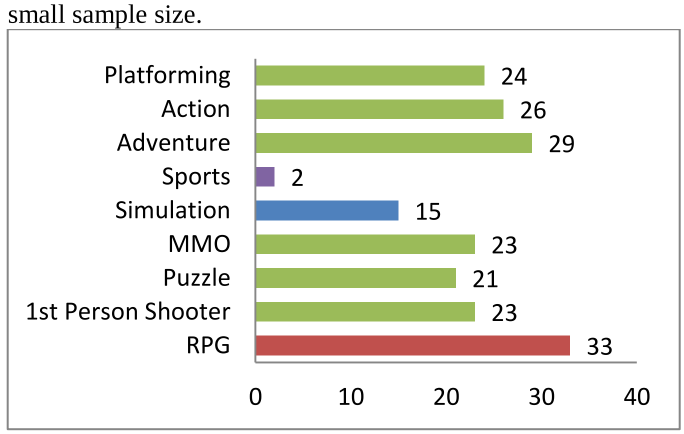

Figure 11 - Favorite Genres (multiple options allowed).

Players reported on their favorite genres by circling multiple selections in the demographic survey. Testers preferred Role-playing games (RPG) the most (33), with Sports ranking the lowest (only 2 people). Every test contained at least 4 testers who selected "Puzzles" as a favorite genre. The researcher looked into a difference between those who selected "Puzzles" versus not choosing puzzles, but discovered no connection. However, it does show that the testing group has diverse interests.

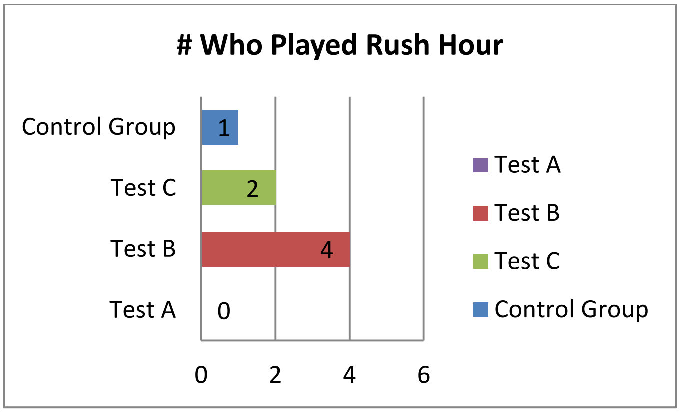

Figure 12 - Those that played Rush Hour prior to the test

Only seven people said they played the board game *Rush Hour* prior to playing the level. Their opinion of the game remained positive throughout the test, however, their knowledge of *Rush Hour* prior to playing did not appear to affect time perception, nor did it change the Unit Effect. The majority of testers who played *Rush Hour* took Test B (Bar, Normal Clock, then Seconds Only). No one from Test A knew of *Rush Hour*.

	Test A:	Test B:	Test C:
P1:	Normal Clock	Bar	Seconds Only
P2:	Seconds Only	Normal Clock	Bar
P3:	Bar	Seconds Only	Normal Clock

Table 1 - Test Reference Table

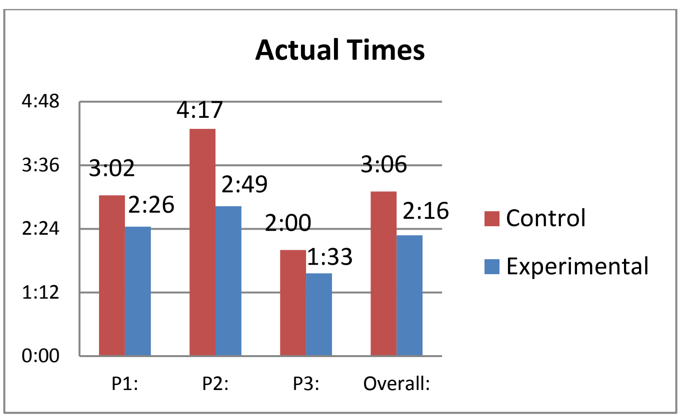

Figure 13 - The Amount of Time Actually Spent

Average/Total of Experimental (Test A, B, and C):

Actual Times:	Average Time	MAX / MIN / CONTROL COMPARISON
P1:	2:26	3:02
P2:	2:49	4:17
P3:	1:33	2:00
Overall:	2:16	3:06

Estimated Times:		
P1:	1:46	4:14
P2:	2:40	4:56
P3:	1:16	2:46
Overall:	1:54	3:58

Table 1 - The Table of the Average Actual Times

Overall, testers in the control group spent more time on the puzzles than the experimental group (e.g. all three puzzle times averaged together: 2:16 experimental, versus the 3:06 control). In addition, each specific puzzle took longer in the control group than in the experimental group. The longest amount of time for someone to solve a puzzle in the control group was on puzzle #2, in which a tester took 7:25 seconds before giving up. All questions still applied to this tester, however, instead of "how long did it take to complete the puzzle?" the question became "how long did it take before you gave up?" The fastest time in the control group occurred in the same puzzle (#2) at 1:01. The fastest time in the experimental occurred in Test A, puzzle #3 at 0:53 (draining bar clock). The slowest time in the experimental happened in Test B on puzzle #2 at 8:30 (Normal clock). This data suggests that the control group consistently spent more time on the puzzles than the experimental and that the timers might have had a role in pushing players to finish sooner.

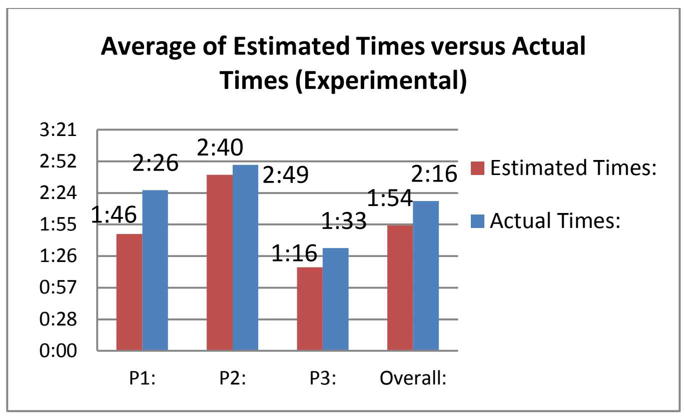

Figure 14 - Time estimations of Testers

After each puzzle, testers estimated the amount of time spent on the puzzle. The chart above shows their estimation, as well as the actual time spent by the tester on a puzzle. It also shows an overall average of all of the puzzles' times put together. In the experimental group, testers consistently estimated their times lower than the actual time spent. The first puzzle had the largest difference between estimation and actuality, with 40 seconds discrepancy. On the last puzzle, on average, testers estimated within 17 seconds of their actual times (1:16 actual, 1:33 estimated). The significance of this chart comes from the fact that testers consistently underestimated their times, regardless of the amount. If testers spent more time on the puzzles, it is possible that the same consistency would carry over.

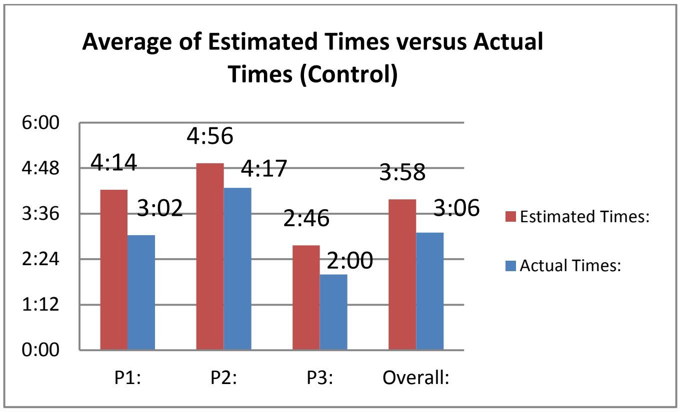

Figure 15 - Estimations in the Control Group

Contrary to the experimental group, testers in the control group consistently estimated the amount of time spent on each puzzle to be higher than their actual times. The difference in time in the control group was also larger, with the smallest difference between actual and estimated time being 39 seconds, and the highest being 1:12 (72 seconds).

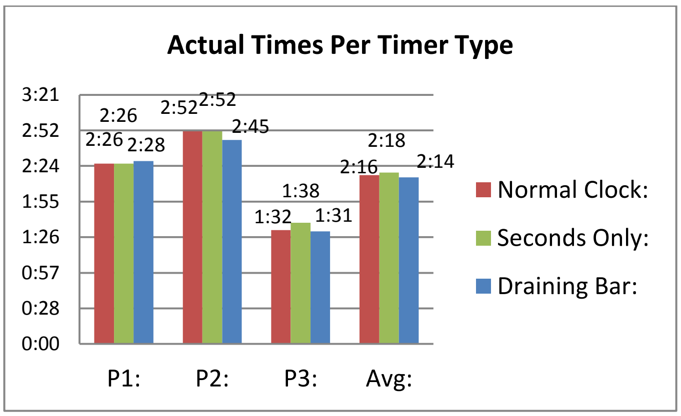

Figure 16 - Actual Times per Timer Type

No matter which timer type the tester had, testers spent relatively the same amount of time on each puzzle. For example, puzzle #1 took testers around 2:24, regardless of the timer type. Therefore, the results suggest that the different timer types did not have a significant effect on delaying or speeding up the player. Puzzle #3 usually went the quickest, and puzzle #2 usually took the longest. Eight testers attributed the draining bar as adding the most pressure; however, it made no difference in the resulting times.

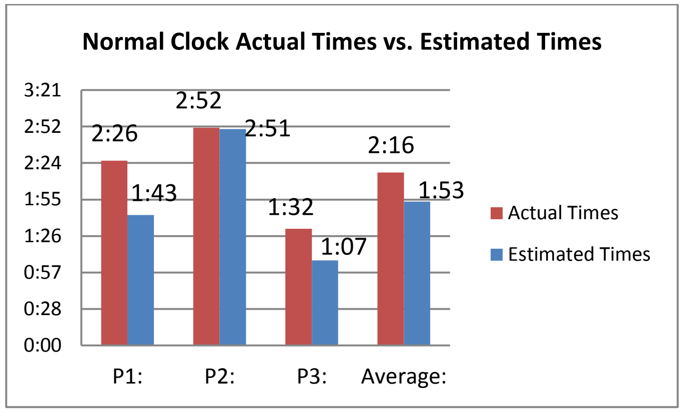

Figure 17 - Normal Clock Actual Times vs. Estimated Times

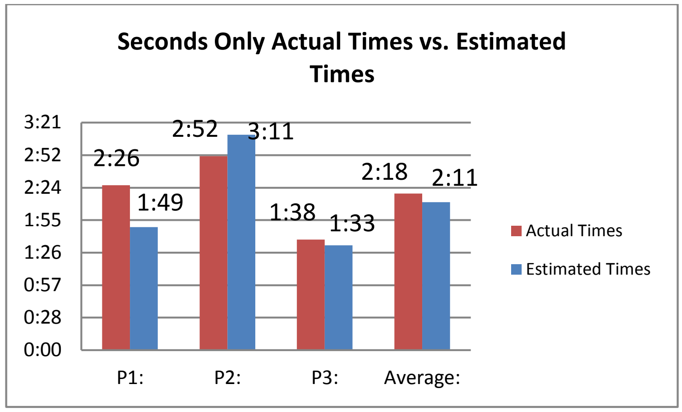

Figure 18 - Seconds Only Actual Times vs. Estimated Times

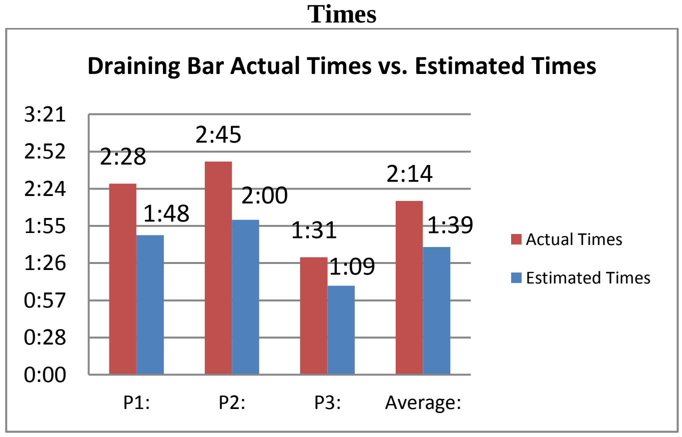

Figure 19 - Draining Bar Actual Times vs. Estimated Times

After each puzzle, the corresponding survey asked testers to calculate the amount of time they believe to have spent while solving the puzzle. If the tester accurately estimated their time, then nothing affected their perception of time.

On average, testers estimated the time spent on any of the puzzles to be lower than the actual time spent. The Normal clock had an occurrence on puzzle #2 in which players accurately measured time spent (2:52 actual, 2:51 estimated). Overall, the Seconds Only Clock showed the smallest difference between actual/estimation (2:18 actual, 2:11 estimated). However, with the Seconds Only Clock on puzzle #2, testers estimated their time as higher (2:52 actual, 3:11 estimated). The Draining Bar showed the greatest difference between actual times and estimated times (2:14 actual, 1:39

estimated). The longer time spent on a particular puzzle, the greater the difference. For example, the chart directly above shows that on puzzle #2, players spent, on average, around 2 minutes 45 seconds. Yet, they estimated 45 seconds shorter (27% off the actual time of 2:45). On puzzle #3, testers spent 1 minute and 31 seconds, but estimated to take 1:09 (24% difference of 1:31).

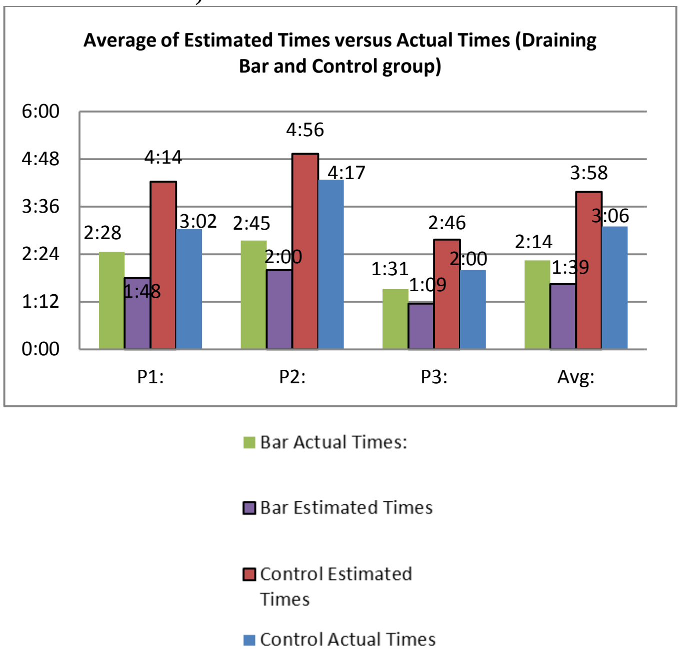

Figure 20 - Average Estimated Times vs. Actual Times (Draining Bar and Control group)

The Draining Bar accounted for the highest difference between actual times and estimated times, however, the control group still surpassed the bar's values in over-estimations. Thus, the two make a great comparison. Notice in the two charts above that the actual times of the control group (blue) are consistently lower than the estimated times (red). The opposite is true about the Draining bar, in which the actual times (green) are higher than the estimated times (purple). For quick comparison, the outlined bars are estimated times. The highest difference between actual times and estimated times in the Draining bar is 45 seconds (puzzle #2, 2:00 is 72% of the actual time of 2:45, which translates into a 28% under-estimation). Meanwhile, the highest difference for the control group is 72 seconds (puzzle #1, 3:02 is 72% of than the estimated time of 4:14, 28% over-estimation). Both show a similar percent difference of 28%. The average difference between the actual times and estimated times in the Draining bar is 35 seconds (1:39 is 73% of the actual time of 2:14). The average difference for the control group is 52 seconds (3:06 is 78% of the estimated time of 3:58). The results in the charts could suggest that the less information given to a player, the more likely they are to over-estimate or under-estimate the amount of time spent on the puzzle. It is worth noting that the Normal clock is the most detailed timer, the Draining bar the least, and the control group provides no details on time at all. Interestingly, one tester pointed out after the test that he felt more pressure from the Draining bar because he could see the relative progress he had made on the puzzle. For example, when he saw he had drained 50% of the bar, yet felt as though he was not close to finishing the puzzle, he felt more stress. The Normal clock and the Seconds Only clock do not show the player how much time has progressed unless they do the math of initial time minus current time.

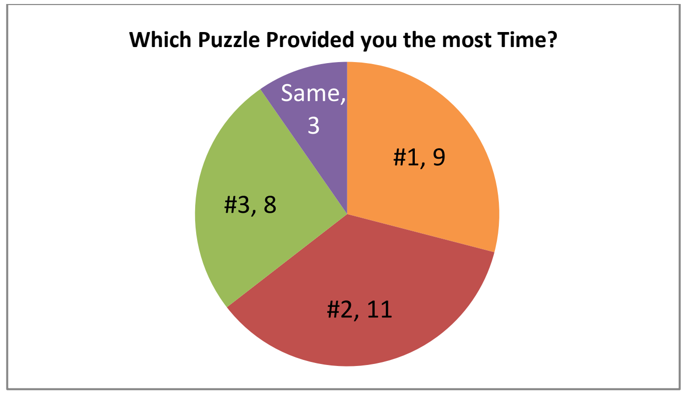

Figure 21 - The Unit Effect's Effectiveness by Puzzle

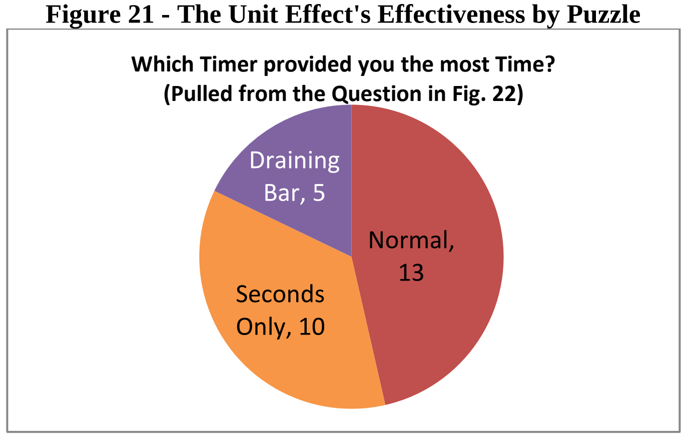

Figure 22 - Which Timer Provided you with the most Time? (Pulled from the question in Fig. 22)

After the test, testers filled out the question on the post-survey, "Which puzzle provided you with the most time?" The chart above is a tally of the number of times experimental testers chose a particular puzzle/timer, or stated an equal distribution. Only three testers of the forty recognized that each puzzle gave the same amount of time. One tester said that two of the three puzzles gave the most time (hence the 31 total in the chart). Puzzle #2 received the most attention, perhaps because 4 testers spent 5:00 or more to solve puzzle #2. The researcher also looked to see if the respective timer types had any effect on the distribution. The Normal Clock accounted for 13 of the tallies. The Seconds Only Clock accounted for 10. Finally, the Draining Bar only accounted for 5 of the tallies. The data suggests that the Unit Effect was relatively successful and that players might perceive the Draining Bar as moving so fast that they believe it contains less time.

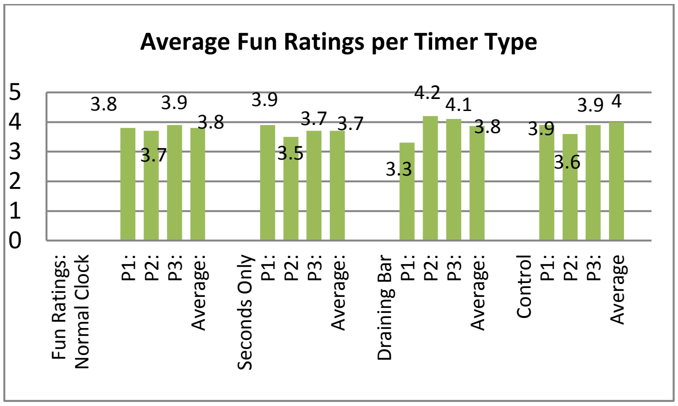

Figure 23 - Fun Ratings by Timer Type

	Normal Clock	Seconds Only	Draining Bar	Control
P1:	3.8	3.9	3.3	3.9
P2:	3.7	3.5	4.2	3.6
P3:	3.9	3.7	4.1	3.9
Average per Timer/NA:	3.8	3.7	3.8	4

	P1:	P2:	P3:
Normal Clock	3.8	3.7	3.9
SO	3.9	3.5	3.7
Bar	3.3	4.2	4.1
Control	3.9	3.6	3.9
Average per Puzzle:	3.725	3.75	3.9

Figure 24 - Tables of Fun Ratings

In order to compare and measure player satisfaction based on timer types, testers rated how much fun they had on a scale of 1-5 after each puzzle. Testers found the Draining Bar on Puzzle #1 the least fun (average 3.3). However, on puzzle #2 they found the draining bar the most fun (average 4.2). Overall, the control group's fun averaged out slightly higher (average 4) than the experimental group (average 3.75) most likely because it was consistent (no timers). Interestingly, however, one tester on map C commented after the test, saying he enjoyed the puzzles more because of the timers. Thus, it is difficult to measure what testers attribute to their fun ratings with the current scale system.

Hardest Puzzle	A	B	C	Total:
P1:	3	4	3	10
P2:	5	5	7	17
P3:	2	1	0	3

Easiest Puzzle	A	B	C	Total:
P1:	4	3	3	10
P2:	4	2	1	7
P3:	2	5	6	13

	Totals Per Timer:	
Key for Quick Use:	Easiest	Hardest
Normal Clock	12	8
Seconds Only	12	9
Draining Bar	6	13

Table 2 - Difficulty for Experimental Group

In order to see if timers had an effect on difficulty, the post-survey asked the question, "Which puzzle was the hardest/easiest?" The table above shows that a majority of testers picked Puzzle #2 as the hardest (selected by 17 testers, the most from test C), and Puzzle #3 as the easiest (selected by 17 testers, the most from test C). Based on observations of the testers, puzzle #2 proved the most difficult because if a tester did not solve it early, they continued to second-guess their moves. In addition, because the puzzles did not come with a reset mechanic, these testers did not know if they were making progress or going back to the beginning layout.

It is also interesting to pull out the timer types associated with the selected puzzles. For example, if a tester picked Puzzle #3 as the hardest puzzle on Test A, then the puzzle used the Draining Bar. Interestingly, the Draining Bar accounted for 13 of the "Hardest puzzles" and only 6 for the "Easiest." In contrast, the Normal clock and Seconds Only clock had almost the same tally count, accounting for more "Easy" than "Hard."

Hardest Puzzle	Control group
P1:	2
P2:	5
P3:	3
Easiest Puzzle	
P1:	1
P2:	5
P3:	4

Figure 25 - Easiest/Hardest Puzzle within the Control Group

It is interesting to see the control group's response to the same question because timers do not have an effect. One must remember that the puzzles in the test were the same difficulty level (Beginner *Rush Hour* layouts). Half of the testers said the easiest puzzle was puzzle #2. The other half said puzzle #2 was the hardest. 40% of the testers said Puzzle #3 was the easiest, however, three testers said it was the hardest (33%). A larger pool of testers could show a better comparison to the experimental group; yet, the current data shows that puzzle #2 had an inconsistent measurement of difficulty. This inconsistency most likely stems from the layout of puzzle #2. The observer noticed that if testers moved the orange block up into the corner, the rest of the puzzle was simple for the tester. However, if testers failed to notice the orange block's significance, they found themselves stuck and over time grew frustrated.

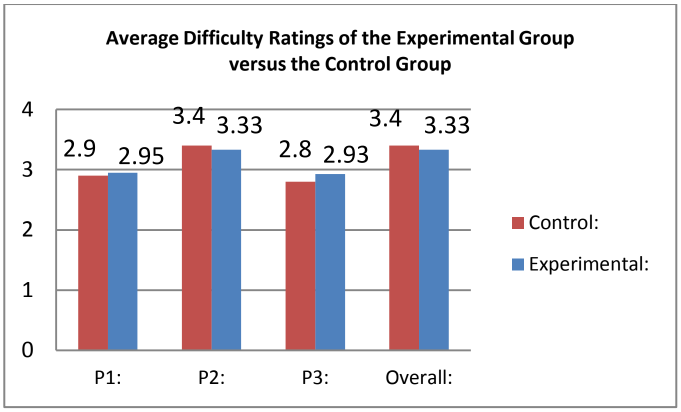

Figure 26 - Difficulty Ratings of the Entire Level

By asking testers about the difficulty of the puzzles, the researcher seeks to know if timers have an effect on the difficulty. When looking at the experimental test as a whole versus the control test, the difference in difficulty is not apparent (3.4 average, to the 3.33 average). Perhaps this occurred because timer or no timer, the objective remained the same (get the red block to the other side). Therefore, when a player measures "difficulty" they might see the puzzle as simple after beating it. In addition, a difficulty of "2" in one person's mind might be the same as a difficulty of "1" or "3" to another person. Four testers admitted that the timers added to the difficulty, and used words like "anxious", "panic," "stress" when describing the experience. Overall, the control group and the experimental group are effectively the same in difficulty ratings.

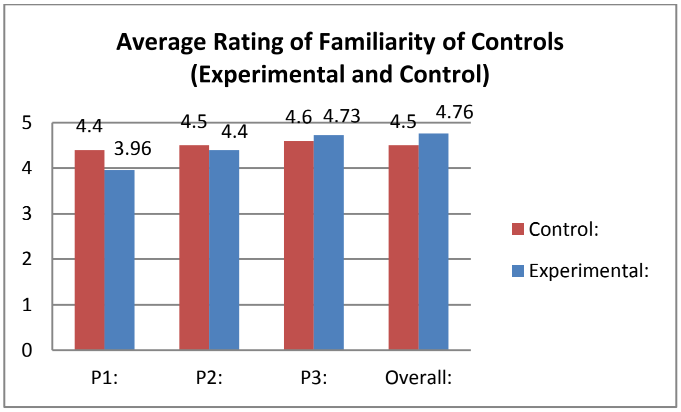

Figure 27 - Average ratings of the Controls

Controls play a huge role in whether or not players have fun or become frustrated by the puzzle. Therefore, without knowing how players feel about the controls, the test would not know if controls affected fun/difficulty or even time perception. After each level, the surveys gave testers a chance to admit how familiar or comfortable they felt with the controls on a scale of 1-5. On average, testers rated a 3.9 or higher from puzzle 1-3. In the experimental group, testers gradually went from a 3.96 to a 4.76 average. Meanwhile, the control group steadily stayed between 4.4 and 4.6 on average. As expected, the ratings increased from puzzle #1 up to puzzle #3. Players should then see puzzle #3 as an easier puzzle since they know the controls. However, as mentioned earlier, the difficulty of puzzle #3 differed across test groups. It is possible that the controls were a factor, but the conclusions are not clear.

VI. Conclusions

Did players really have more fun during a timed challenge than a non-timed challenge? Did they perceive time as faster or slower while solving the timed puzzles? Did the different timer presentations change a player's perception of time more or less than another? The researcher discovered that fun was hard to measure, the timers made players finish quicker than without timers, the timer types differed from one another, and the Unit Effect was useful. This data should help designers better understand the implications of timers, and their effect on the player's perception of time.

In order to consider the differences of fun between a timed challenge and non-timed challenge, the test used a scale of 1-5. However, measuring fun in this way made comparing the two difficult because of several factors. When measuring fun, the researcher did not expect the control group's fun rating to match (or beat) the experimental group ratings (overall 4 rating, versus the overall rating of 3.8 or 3.7 in the experimental). The scale only went from 1 to 5. A 1 to 9 scale might show better comparisons because a tester who put 2 out of 5 might actually have meant 5 on a scale of 9.

Testers also did not attribute what made the puzzle fun unless they commented afterwards in the question asking about "how timers made them feel." Thus, they did not have the chance to say that a particular element (the timer, the objective, etc.) made the puzzle more or less fun.

In order to answer the second question of the hypothesis (whether players perceived time as faster or slower with timers), the observer measured the testers' actual times and compared them with their estimated times. Timers generally made players finish the puzzles sooner than if there were no timers. The times from the control group were consistently slower than the times in the experimental group (average 2:16 experimental, versus 3:06 control). The researcher expected the timers to add pressure, but did not expect 9 out of 10 testers from the control group to take their time in solving the puzzles. Not only did testers finish faster with timers, they also believed that they spent less time on a puzzle when they had a timer (actual time of 2:16, under-estimated time 1:54). This underestimation was obvious with the Draining Bar because testers did not know how much time remained, but they could see the progression of time (the amount of time drained versus amount of time they used to have) more clearly than with the other clocks.

Thus, the test showed that the timer types differentiated from one another. However, the timers did not differ exactly as the researcher expected. The researcher expected testers to only accurately estimate their times on the Normal Clock, and only under-estimate their times on the Seconds Only Clock. However, testers did everything possible (they over-estimated, under-estimated, and accurately estimated their actual times with the Normal Clock and Seconds Only clocks). On average, testers under-estimated their actual times (actual time of 2:16, estimated time 1:54). The hypothesis also stated that the Draining Bar would result in a consistent over-estimation by 2:00 on average. Yet, testers, on average, under-estimated their times by 0:35, most likely because the Draining Bar did the best job at constantly pressuring the player (it accounted for the most "Hardest puzzle" tallies). The hypothesis foresaw a consistent pattern per timer type, but the behavior of the testers was inconsistent. This most likely depended on whether or not testers looked at the timers for reference. Not every tester admitted to ignoring the timers, but it happened multiple times. Unfortunately, the researcher did not record the amount of times testers said that they ignored the timers. Moreover, it is possible that other factors contributed to the inconsistency. The test had an open-ended question, "Briefly describe how the timers made you feel." Nine testers pointed out the difference of the Draining Bar to the other timers. The following comments come from six of those nine testers:

"I felt most comfortable with min + sec, the bar was hard to keep track of and made me feel worried."

"Draining Bar was aggressive."

"Bar motivated me most."

"Liked bar, had better visual of time left."

"Felt rushed, but not in a negative way."

None of the testers admitted that the timer made it less fun, nor did their respective fun ratings say that they disliked the puzzle using the Draining bar.

The Unit Effect successfully "tricked" 27 out of 30 experimental testers, helping the researcher obtain unbiased test results. A successful "trick" was when a tester had no idea that the times of the timers were the same (4 minutes). It is because of the Unit Effect that the researcher is able to compare the various timers, and it is highly recommended for future research. For example, a designer who wants to compare health bars could use the same measurements within different units of health (e.g. "Heath Points: 200/200", a draining sphere container, or hearts that split into quarter hearts and half hearts).

The test had several limitations. As mentioned earlier, the test did not explicitly ask testers to define what made the test/puzzle fun. Only one tester said in the open-ended question, "Timers added to the fun." All of the remaining 29 testers defined the amount of fun in a general sense. It would have been desirable to add the question, "Which puzzle was the most fun?" In addition, controls made the test run longer than it could have. For example, a separate mouse cursor would have sped up tester movement because a tester could move a mouse on top of blocks faster than the slow moving cursor with WASD. In the actual board game, *Rush Hour*, players can physically grab the car pieces, allowing two hands to grab two pieces at once for plenty of quick movement. Thus, the current WASD controls added to the tension/pressure because players felt sluggish. Another limitation was that the Guildhall student tester pool proved uneven. Most of the students are male. Perhaps, not having as many females left the test with a small hole of data, but the researcher is unsure based on the results. Nothing in the data suggests that gender could change the results, but the researcher would have preferred to cover it, just in case. In addition, a reset button could have helped testers who grew frustrated with their progress. It is not always helpful, but when playing a puzzle game like *Rush Hour*, it is not obvious what the starting layout was. If testers know they are returning to the beginning layout with their moves, then they understand the lack of progress made on the puzzle. If a reset button were not possible, another solution would be putting transparent blocks underneath the starting layout. That way, testers would know the initial spots of the blocks, and get an idea of the progress they have or have not made. It would also be worthwhile to investigate more deeply how players felt per timer type. Adding questions such as, "how much pressure did the timer add to the puzzle?" and "Did you look at the timer while solving the puzzle?" Both questions could help the researcher understand how testers felt about the timers and if the timers actually affected their performance.

For future research, the observer suggests finding an equal pool of testers (50/50, male/female. It is possible with a larger pool of testers that a particular female could outperform the fastest time set by a male. In addition to a 50/50 pool, the observer discovered that progression of time made the Draining Bar stand out. The length of the bar could also factor into communication progression. The longer the bar, the faster the perceived drain could occur. With a faster drain, the player might assume time as moving faster. Finally, adjusting the Normal clock or Seconds Only clock to have a "time spent" number could prove interesting (e.g. 1:59 on the left going up and 2:01 on the right going down). By doing this, the other clocks show progression like the Draining Bar currently does.

In closing, the data above suggests that timers can make time seem to move faster and have the potential to add difficulty to puzzles. Designers should look into the components of each timer type mentioned in the study, so that they can better understand each one's impact on their levels and games. The Normal Clock, Seconds Only Clock, and the Draining Bar each have their own effect on games, and it could prove crucial if a designer implements one in the wrong way.

VII. References

\\ghsrv\users\C17\MIT Repository Folders\Morgan Davis

[1]: The Unit Effect and Player Perceptions, Jamie Madigan, psychologyofgames.com, March 9, 2011
http://www.psychologyofgames.com/2011/03/the-unit-effect-and-player-perceptions/

[2]: "Brain, Time Perception and Synesthesia", PBS YouTube Kostas Partsinevelos September 2011:
http://www.youtube.com/watch?v=iOVtGTPsFSI

[3]: "Psychology Time Perception Experiment in Teenagers", YouTube Daniel Hopkins, Monroe Technology Center, November 2010
http://www.youtube.com/watch?v=Ssh6ZxGl1j4

[4]: Time Perception, Immersion and Music in Videogames, Timothy Sanders, Paul Cairns, September 2010:
http://ewic.bcs.org/upload/pdf/ewic_hci10_paper18.pdf

[5]: Mass Effect 3 (Multiplayer), Bioware, Edmonton Canada, March 2012:
Screenshot taken from YouTube video:
https://www.youtube.com/watch?v=zXIkhtChwHk

[6]: Picture from *Shift*, ArmorGames by Tony 2008:
Screenshot taken while playing the game on ArmorGames

[7]: Picture from *Super Mario World 2: Yoshi's Island*, SNES by Nintendo 1995:
Screenshot taken from YouTube video:
http://www.youtube.com/watch?v=9bOpsQKHLGw

This author became a student at Abilene Christian University from 2007-2011, earning his Bachelor's degree in Information Technology. Later he enrolled at Southern Methodist University's Guildhall in May 2011 and pursues a Master degree in Level Design. The author lived in Abilene, TX for most of his life, but currently lives in McKinney, TX.

VIII. APPENDIX

For an in-depth look at all of the survey questions, level abstract, and results, see the following documents located in the repository folder (\\ghsrv\users\C17\MIT Repository Folders\Morgan Davis)

Davis_AppendixA_Survey.docx
Davis_AppendixB_LevelAbstract.docx
Davis_AppendixC_Results.xlsx