STUDENTS' FRAMING OF A DIGITAL PHYSICS VIDEOGAME AS SEEN IN STUDENT DISCOURSE AND POST-GAMEPLAY DATA

By

James Michael Hughes

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

MASTER OF SCIENCE

in

Learning, Teaching, and Diversity

August, 2013

Nashville, Tennessee

Approved:

Dr. Douglas Clark

Dr. Pratim Sengupta

To my parents, Jimmy and Jackie Hughes for all they have given

and

To the other friends and family for building my parts

ACKNOWLEDGMENTS

This work would not have been possible without the financial support of the National Science Foundation (#1119290 Enhancing Games with Assessment and Metacognitive Emphases – EGAME). I am especially indebted to Dr. Douglas Clark, professor of the Department of Teaching and Learning, and Dr. Pratim Sengupta, Professor of the Department of Teaching and Learning, who have supported my work and career goals whole-heartedly while providing meaningful feedback.

I am grateful to the teachers and students that allowed me to work with them, and participate in their classroom experience as I gathered the data necessary for the completion of the studies represented in this thesis.

I am perhaps most grateful to my loving family and network of supportive friends, who have always pushed me to succeed, reach my goals, and never stop learning. These folks have been there through all of the times, both good and challenging.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO BOTH STUDIES

This thesis details two studies that take quantitative and qualitative approaches in exploring middle school students’ perspectives in engaging with digital video games that are designed to teach Newtonian Mechanics. The qualitative study explores a pilot study and a predecessor study that utilize the constant comparative method to categorize students frames about gameplay into useful discourse categories that can be refined in future research and that also guide design improvements.

The quantitative study looks at students’ post-gameplay data from various class-room drawing activities and compares the complexity of these drawings to in-game data about the frequency and length of gameplay. Results from this study support future designs of larger hierarchical linear modeling studies that could contain populations as large as school districts.

CHAPTER II

STATEMENT OF PROBLEM AND SIGNIFICANE FOR QUALITATIVE STUDY #1

Focus of Inquiry – Epistemologies and Beliefs in Surge

Existing research on the use of digital games in education has shown that games can be productive educational tools (Lepper & Malone, 1987; Loftus & Loftus, 1983, Csikszentmihalyi, 1975; Deci, 1975; Lepper & Greene, 1975; Ferster & Skinner, 1957; Aldrich, 2004; Gee, 2003; Shaffer, Squire, Halverson, & Gee, 2005). This work takes a deeper analysis at understanding how students shape their beliefs in one such gaming environment called Surge. Surge is a game designed to teach Newtonian Mechanics to middle school students. Surge does this by situating the content knowledge within an exciting videogame that contains elements which attempt to activate students intuitions and drive conceptual change to modify and align faulty intuitions with the appropriate reality defined by physics content knowledge.

The goal of this research is to eventually foster new design elements to modify students' beliefs in more productive ways while they engage with Surge. A simple example is to help students form beliefs of the activity as a tool for learning physics, instead of having students simply think of Surge as a fun game that they are meant to obtain no knowledge from. What are the affordances that students' beliefs bring to the interaction between games and player, and what can we hope to identify as potential new design elements that can foster more productive beliefs prior to, and during interaction?

Unfortunately, little research has been conducted that observes the variety of beliefs that students bring to the interaction between learner with educational games. Little work has been done within the research on educational physics video games that attempts to identify student's beliefs about the environment during play, and to observe how these beliefs emerge and change. Even less research has been done that identifies key design elements that generate beneficial beliefs for students that foster improved physics learning.

To support the improvement of physics learning environments, a deeper understanding of student's beliefs about those environments is needed. An analysis of student interactions with such an environment that uses a lens of framing (Goffman 1974, Tannen 1993) to show evidence of students' beliefs is a productive step towards understanding the nuances of how designed environments and beliefs about those environments interact productively, and non-productively.

Why is this important? As games become a more widely used and widely researched educational tool, it is important to seek meaningful data that fosters support for improved design. A qualitative analysis of students' beliefs and interactions with such learning environments will generate more specific hypotheses about the types of out-of-the-box game design that can increase the appeal of educational games, and improve the medium as a learning tool.

Fit of Paradigm to Focus and Substantive Theory

In attempting to understand the beliefs that students form about the learning environment in question, a naturalistic research approach has been chosen because qualitative analysis of students interactions with the learning context will bring forth richness to understanding their beliefs that a broad scale survey could not achieve. The use of the constant comparative method to identify what kinds of beliefs students take on regarding the Surge gaming environment will lead to grounded theory, at least about this group of students. Such a grounded theory can then be systematically tested and enhanced in future studies. In addition, this research can stand as a small-scale example of future long-term microgenetic analyses which could map students changes in beliefs over time. Iterative microgenetic design experiments would be the most beneficial approach to understanding the important "how" questions regarding ways to change beliefs. In some ways this work can be viewed as the first step towards a path of such design research.

A naturalistic approach is also vital to studying beliefs and epistemologies because these kinds of construct and phenomena must be studied as close to the context that they emerge in. Doing reflective interviewing or large scale survey analysis forces participants to attempt to recall their epistemologies and beliefs, and the simple act of performing this reflective recall has the potential of muddying the true phenomenological nature of the epistemologies and beliefs that would have emerged in their original context (Elby, 2011).

A Closer Look at Surge

Surge is an online interactive two dimensional game that takes place in outer space. Students control a red ball (called surge) from a birds-eye view by applying forces to the ball. The forces are implemented by dragging elements that look like red arrows beneath surge on the playing screen. When the icon that denotes the red ball overlaps the icon that denotes an arrow, a force is applied. The game engine functions in such a way that mimics the physical laws of nature. Thus a right arrow will trigger the ball to move right from rest, and an upward arrow will trigger the ball to move upward from rest. However, the addition of both a rightward arrow and an upward arrow will trigger the ball to travel diagonally at a perfect 45 degrees from the starting point upwards AND to the right.

Besides arrows that direct motion, the game also contains other objects that act to interfere and enhance the motion of the red ball. For the purposes of understanding this research, all that is important is to know that all motion within the surge realm acts in a frictionless outer space environment and all movements follow Newton's three laws of motion.

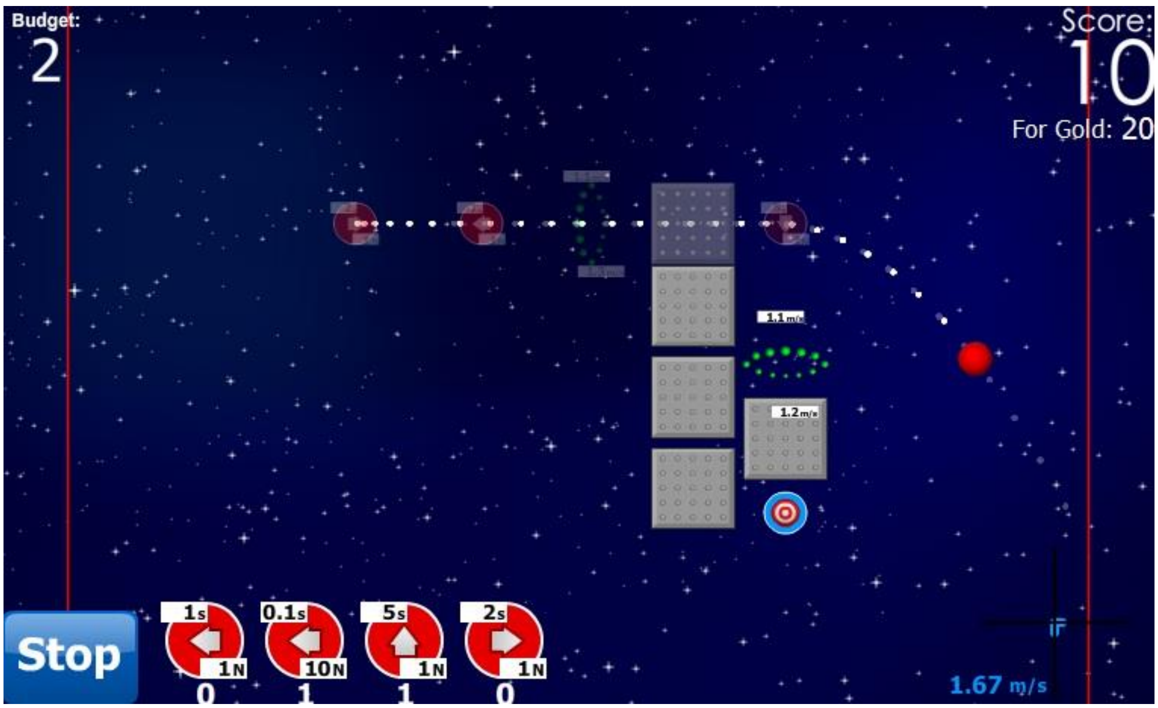

Figure 1: A screen-capture of a typical level in the Surge videogame

CHAPTER III

NATURAL HISTORY OF THE RESEARCH – PILOT STUDY – ALPHA MIDDLE SCHOOL

In this section I will discuss the selection of the research location for the qualitative study, field entry procedures, the methods of the data collection, analysis of the data, and strengths and limitations of the research. The focus of this part of the manuscript is on the original pilot study that took place in Alpha Middle School. The following chapter will focus on the follow up study in Beta Middle School. Both parts are necessary for understanding the evolution of the researcher's methods. Only the data collected in chapter IV is discussed in Chapter V of this manuscript (the section detailing the findings and conclusions). I have made my best effort to provide this information in chronological order, without leaving out any details of the experience. When in doubt, a liberal inclusion of details has been provided because I believe that tiny pieces of information that I might deem unworthy could still be highly important to the reader, bringing about thoughts that I may not have the resources to think of on my own.

The methods section will be further broken down into sections on sampling, successive phases of inquiry, the various kinds of data that I collected, and how those data were recorded and processed.

The data analysis section will discuss a priori thoughts regarding how the data was going to be analyzed in the second study, to follow.

Site Selection: Alpha Middle School

Alpha Middle School, a low socioeconomic status middle school in suburban Tennessee in the United States was selected for several reasons for the pilot study, which was performed in the early summer of 2012. The primary reason for selection was that the teacher of the classroom where the study was conducted was motivated to participate in the research project and was enthusiastic to help out the research team and provide classroom time for our study. In addition, the context provided the team with a diverse group of students across six separate periods ranging in various levels of socioeconomic status and science ability. Students were also racially diverse and had various familiarities with English as a primary language. Some students were English language learners at the time of the study. Alpha Middle School was also easily accessible by the research team since the home-base for the group was Vanderbilt University. This provided short commuting times and ease-of-access.

The study took place over 6 separate class periods of students. All students participated in all elements of the study except for the *Epistemological Beliefs Assessment for Physical Sciences (EBAPS)* survey, which required a secondary permission slip to participate in the survey. This permission slip was necessary based upon the IRB that had previously been approved by Vanderbilt University.

Class 1A contained ten students who completed secondary permission slips, and thus could be included in a final pool of data. Class 1B contained eight students who could be included. Class 2A contained ten students. Class 2B contained ten students.

Class 3A contained fourteen students. Class 3B contained eight students. Effort was made to ensure that an equal number of students were from each gender, but this was difficult considering it was up to students and their families regarding the completion of the permission slip.

Of the students who brought in permission slips, six students per class section were chose to be interviewed on camera. The details of the interview procedures will be expanded in later sections of this paper, and copies of the interviews have been attached in the appendix.

Field Entry Procedures: Alpha Middle School

Access to Alpha Middle School was obtained due to previous correspondences with the specific teacher that we worked with, Ms. Abbett. These correspondences were previously made by my research advisor, Dr. Douglas Clark. Ms. Abbett worked with the middle school to ensure that our access to the classroom was permissible and that our intervention was in line with the Common Core Standards required for seventh grade physics. Thus, no ethical dilemma was raised in terms of students not receiving the appropriate content knowledge for the physics curriculum. Students were simply being exposed to a new method of teaching Newtonian Mechanics that has been grounded extensively in the research on conceptual change.

In order to ensure the safety of the students in the school, the research team was required to sign in at the principal's office every day of the study, and obtain name tags. The research team was also required to sign out every day just prior to leaving the school.

The parents of all students had the opportunity to consent to the study via a permission slip (See the IRB in Appendix A) and all students were informed that at any point they could opt to not participate in the study, at which point Ms. Abbett would give them an alternate assignment to work on. Because the game was an individual assignment that presented no harm to students, those who did not bring in permission slips still engaged with the designed environment, but their data was never collected. Once again, students could fully opt-out, however we never saw this occur. Students who brought in their permission slips were given the EBAPS survey to complete in addition to playing the game. The EBAPS was administered on paper and was passed out to students as soon as they completed the gameplay task.

All student data was tied to the students by a username that they created in the gaming environment. Students who took the paper-based EBAPS survey were informed to write this username at the tops of their papers. All students had an index card on their desks with their username written at the top. This index card was the first thing the video camera was focused on when the record button was pressed. This improved the ease of sorting video data because the thumbnails of the mp4 files on the computer screen showed the relevant screen name for each student. The data was stored locally on an intranet created by the research team during the study, and thus the potential for information leaking outside of the encrypted network onto the internet was very low. In addition, only one central hard drive contained all the participant data. This hard drive was stored in a locked room in a locked office immediately upon returning to Vanderbilt University each day. The data from the server was processed by other researchers on the team besides me. Those researchers then disseminated the encrypted, de-identified data to the rest of the team for their own data analysis needs during our weekly research lab meetings. In all of these data files, the use of the students' usernames was dropped, and their unique identification number was assigned. One single master file existed that correlated the students' usernames with their identification numbers. This file was encrypted and did not need to be accessed during any data analysis.

Methods: Alpha Middle School

The attempt at this research at Alpha Middle School was a pilot study, and thus the sampling, plans for successive phases of inquiry, data collection/data recording, and data analysis were messy and full of trial-and-error. Once again, details about this pilot study are only being included in section B1 of this paper to show the contrast of the juxtaposed study that follows in section B2 (at Beta Middle School).

Sampling. The data from Alpha Middle School was never analyzed in great depth. A systematic design for what to collect and who to collect data from was never achieved. The researcher did obtain valuable insights from the practice of implementing the pilot study, which was the initial intent of the process. The students that were selected for interviews were simply selected based upon the first-come first-serve completion of the current level of the video game that they were playing when researchers passed by them in the classroom. No strict methodology was discussed in terms of the interview protocol, even though a standardized set of questions existed. On some days, varying individuals conducted the interviews introducing error to the data collection process. In addition the sampling methodology was never discussed prior to video interviews being conducted.

Plans for Successive Phases of the Inquiry. The researcher began the Alpha Middle School study knowing that several mistakes would be made and that an opportunity for new data collection with a new population of students would arise roughly two weeks after the completion at Alpha Middle School, thus the researcher expected to obtain valuable experiential learning lessons that would translate into improvements for the subsequent study design. The overall collection plan for the Alpha Middle School study began with the researcher knowing that three video recorded interviews would happen per student. These numbers were dropped to two recordings per student after the first day, because the researcher noticed obvious distress from the questioning and use of cameras on the students. In addition, the researcher had not yet invented follow up probes to make three uses of the students' time for video interviews justifiable in nature. Students would be better served by spending the additional time focused on gameplay and acquiring physics content knowledge.

Data Collection/Data Recording. There were nine kinds of data that were collected for this study:

1. Pre-test answers on a physics assessment
2. Post-test answers on a physics assessment
3. Qualitative gameplay performance data measured as bronze, silver, and gold stars
4. Quantitative gameplay data such as the number of levels and trials completed
5. Students' responses to the Epistemological Beliefs Assessment for Physicals Sciences survey
6. A post-gameplay drawing activity asking students to design their own level of the Surge Game.
7. Daily field notes written by the researcher
8. Video recordings of student interviews during their gameplay on the first day of the study that were transcribed
9. Video recordings of the student interviews on the last day of their study that were transcribed

Items one through four were automatically collected by the computers as students engaged with the videogame. Items five and six were administered by the researcher as soon as students completed the last item within the video game. Students were directed to raise their hands at this point so the researcher knew to administer these follow-up items. Items seven through nine were collected directly by the researcher throughout the span of the study. Note that for this pilot study, field notes were not taken in a systematic way, and were often taken in several forms, including electronic and paper based methods concurrently. Although the collection of field notes was not organized, the processing of those field notes was completed each day after the study. Also, along with the researcher, two other Master's Degree students performed interviews, and sometimes the research team's post-doctoral researcher would stand in for absent Master's Degree students. Nonetheless, this data collection process was not thought out very critically, and was messy at best. This was a vital learning lesson moving forward for the second study at Beta Middle School. Corrections will be discussed in the next chapter.

Data Analysis: Alpha Middle School

The data from Alpha Middle School was never analyzed in a structured fashion. The researcher instead reviewed video clips after each day of data collection and thought critically about the methods of improvement for future study. It was at this point that the researcher did begin to see a pattern emerge in the data about the kinds of knowledge strands students might hold regarding the game. The researcher noticed at this point very obvious differences in student discourse regarding the game for games-sake, and the game as a physics learning tool, but this pattern was not explored deeply at this point-in-time during the study, because the researcher felt very uneasy with disarray of unorganized data and poor methodology he had initially conducted. A simple note was made of the pattern for further analysis once new data had been collected at the follow-up study at Beta Middle School. The researcher also never gave thought to continue data analysis throughout the study, and only field notes were reviewed daily. At this point the researcher didn’t have the appropriate knowledge to identify a unit of analysis for any coding purposes, and data was evaluated holistically, at best.

Strengths and Limitations: Alpha Middle School

Alpha Middle School was intended to be a messy pilot study for a beginning graduate student interested in getting “his feet wet” as a live researcher. The process allowed for deep experiential learning and prepared the researcher for his second attempt at data collection. Thus the major strength to this study was simply improving the researchers approach at refining his working style, with an attention to best-practices moving forward. Obviously, many weaknesses in the data collection process, and lack of data analysis process are evident.

CHAPTER IV

NATURAL HISTORY OF THE RESEARCH – BETA MIDDLE SCHOOL

In this section I will discuss the selection of the second research location, field entry procedures, the methods of the data collection, analysis of the data, strengths and limitations of the research, and how the research aligns with trustworthiness criteria. The focus of this part of the paper is *now* on the second study which took place in Beta Middle School. The previous part of the paper (Part B1) focused on the pilot study in Alpha Middle School with Ms. Abbett's classes of students. Both parts are necessary for understanding the evolution of the researcher's methods. Only the data collected in this chapter are discussed in the following chapter of the manuscript, V, (the section detailing the findings and conclusions). I have made my best effort to provide this information in chronological order, without leaving out any details of the experience. When in doubt, a liberal inclusion of details has been provided because I believe that tiny pieces of information that I might deem unworthy could still be highly important to the reader, bringing about thoughts that I may not have the resources to think of on my own.

The methods section will be further broken down into sections on sampling, successive phases of inquiry, the various kinds of data that I collected, and how those data were recorded and processed.

The data analysis section will discuss a priori thoughts regarding how the data was going to be analyzed, how data analysis methods changed, and the renegotiation of the sampling unit mid-study.

The section on trustworthiness will discuss the credibility, transferability, dependability, and confirmability of the study.

Site Selection: Beta Middle School

Beta Middle School, also a low socioeconomic status middle school in suburban Tennessee in the United States was selected for many similar reasons as the pilot study school. The study was performed in the early summer of 2012, two weeks after the pilot study that was previously discussed. Primarily, the teacher of the classroom where the study was conducted was motivated to participate in the research project and was enthusiastic to help out the research team and provide classroom time for our study. It should be noted that this teacher, Mr. Barnes, was even more enthusiastic about the implementation of the Surge videogame for his students. Unlike Ms. Abbett, Mr. Barnes began every single class with a short PowerPoint lesson on Newtonian Mechanics, thus situating the video game activity more directly in a frame of physics education. Mr. Barnes's lecture style was collaborative in nature, and he frequently selected students to answer his questions about Newtonian Mechanics, making conscious effort to tie the terms and concepts of Newtonian mechanics directly to the objects in the Surge videogame.

In addition, Mr. Barnes's classroom also provided the research team with a diverse group of students across four separate periods ranging in various levels of socioeconomic status and science ability. Students were also racially diverse and had various familiarities with English as a primary language. Some students were English language learners at the time of the study. Beta Middle School was also easily accessible by the research team. This provided short commuting times and ease-of-access just like the pilot study provided.

The study took place over four separate class periods of students. All students participated in all elements of the study except for the EBAPS survey. The periods for this middle school were unique because the third period timeslot was divided in half. Students left their third class period with Mr. Barnes halfway through the allotted time to eat lunch in the cafeteria. Then they returned. While this didn't influence the amount of time that students engaged the Surge videogame, it certainly changed the nature of the learning activity by introducing a large interruption, and perhaps time to communicate with other students in a unique fashion that periods one, two, and four were not granted.

Class Period 1 contained eleven students who completed secondary permission slips, and thus could be included in a final pool of data. Class Period 2 contained eleven students who could be included. Class Period 3 contained sixteen students. Class Period 4 contained sixteen students. Effort was made to ensure that an equal number of students were from each gender, but this was difficult considering it was up to students and their families regarding the completion of the permission slip.

Of the students who brought in permission slips, six students per class section were chosen to be interviewed on camera, similar to the pilot study. However the selection process of the students was improved. The three individuals performing the video interviews had a paper-based semi-structured interview protocol with built in probes (Appendix H). The researcher and two other interviewers discussed beforehand the many various examples that could emerge from student discourse, and the kinds of probes that should be used to follow up. A spreadsheet was generated that provided each interviewer with the two students in each class section that they would interview, split for gender equality and equality based on the version of the Surge video game that the students were playing. In addition the interviewers were told to inquire with the student about any known upcoming absences so as to avoid loss of data during the timespan of the study. Some students were absent frequently for extracurricular activities like orchestra and sports.

Field Entry Procedures: Beta Middle School

Access to Beta Middle School was obtained due to previous correspondences with the specific teacher that we worked with, Mr. Barnes. These correspondences were previously made by my research advisor, Dr. Douglas Clark. Mr. Barnes worked with the middle school to ensure that our access to the classroom was permissible and that our intervention was in line with the Common Core standards required for seventh grade physics. Once again, no ethical dilemma was raised in terms of students not receiving the appropriate content knowledge for the physics curriculum. Students were simply being exposed to a new method of teaching Newtonian Mechanics that has been grounded extensively in the research on conceptual change.

In order to ensure the safety of the students in the school, the research team was required to sign in at the principal's office every day of the study, and obtain name tags. The research team was also required to sign out every day just prior to leaving the school.

The parents of all students had the opportunity to consent to the study via a permission slip and all students were informed that at any point they could opt to not participate in the study, at which point Mr. Barnes would give them an alternate assignment to work on. Because the game was an individual assignment that presented no harm to students, those who did not bring in permission slips still engaged with the designed environment, but their data was never collected. Once again, students could fully opt-out, however we never saw this occur. However, we did have one small incident were the researcher incorrectly approached a student and politely asked to record her. This student declined, and then the researcher realized that she was not on the list of students who had provided a permission slip. The researcher then apologized, explaining his confusion and the student said the situation was OK. No data was collected. The researcher continued to monitor the student for signs of distress related to being approached incorrectly.

Students who brought in their permission slips were given the EBAPS survey to complete in addition to playing the game. The EBAPS was administered on paper and was passed out to students as soon as they completed the gameplay task. The EBAPS survey was redesigned this time to allow students to simply circle numbers along a Likert scale from zero to five, instead of handwriting the numbers to agree with the provided statements. The researcher felt it was easier for students to visualize the spectrum of decisions from “agree” to “disagree” when the words existed along an arrow on the paper. The EBAPS can be seen as part of the IRB in Appendix A. Note that the IRB itself contains its own appendices.

All student data was tied to the students by a username that they created in the gaming environment. Students who took the paper-based EBAPS survey were informed to write this username at the tops of their papers. All students had an index card on their desks with their username written at the top. This index card was the first thing the video camera was focused on when the record button was pressed. This improved the ease of sorting video data because the thumbnails of the mp4 files on the computer screen showed the relevant screen name for each student. The data was stored locally on an intranet created by the research team during the study, and thus the potential for information leaking outside of the encrypted network onto the internet was very low. In addition, only one central hard drive contained all the participant data. This hard drive was stored in a locked room in a locked office immediately upon returning to Vanderbilt University each day. Once again, the data from the server was processed by other researchers on the team besides me. Those researchers then disseminated the encrypted, de-identified data to the rest of the team for their own data analysis needs during our weekly research lab meetings. In all of these data files, the use of the students' usernames was dropped, and their unique identification number was assigned. One single master file existed that correlated the students' usernames with their identification numbers. This file was encrypted and did not need to be accessed during any data analysis.

Methods: Beta Middle School

The attempt at this research at Beta Middle School was much more structured than the pilot study at Alpha Middle School, and thus the sampling, plans for successive phases of inquiry, data collection/data recording, and data analysis were more coherent and produced usable data that could be analyzed.

Sampling. The data from Beta Middle School was gathered on a daily basis in a systematic fashion. Part of this systematicity was ensuring that the same exact interviewers were recording each day that interviewing took place, and that they were interviewing the same individual students. The interviewers also communicated each day before and after the interview process to ensure they had a similar vision of how the interviews should take place. It was encouraged to first approach the students and to introduce themselves as people who were interested in helping the students learn about Surge. Interviewers were encouraged to sit on a chair at the same height as the student. The interviewers were encouraged to watch the students' gameplay and ask a few comforting questions prior to opening the video camera. Once students had completed a level of the Surge game, the interviewer was then directed to ask politely if they could record the screen while asking some questions. The interviewers reiterated the fact that the screen would be recorded, and several times throughout the study students were reassured that their faces would not be recorded. This was told to the students by their teacher, the researchers, and the primary investigator.

As previously discussed, interviewers already had a predefined spreadsheet of the students that they were required to video tape, balanced for gender and version of the Surge videogame. Because the interview protocol, selection of interviewers, follow up probes, and selection of students had all been streamlined and made as constant as possible, the researcher felt much more confident that the data from Beta Middle School was comparable across students.

After each day in the field, the researcher took all video data files and labeled them according to a standard protocol. Files were then transcribed by an outside company that was local to the Vanderbilt University area. The company worked with the researcher to ensure that each file was appropriately formatted for direct import into the NVivo 10 Software program that would be used for constant comparative coding data analysis.

This very streamlined system of the data collection allowed for the researcher to select a sampling unit for the video data at the level of single lines of student speech. Each line of student speech was thus coded when the data was processed. More of this procedure will follow in the data analysis section.

Plans for Successive Phases of the Inquiry. The researcher began the Beta Middle School study having learned from several mistakes that arose during the pilot study.

Thus, the entire data collection process was planned out. The researcher decided that day one of the study would be used to acquaint himself with the classroom and students as they began to immerse themselves in the videogame playing task. The researcher simply floated around the room, taking notes of students' interactions, but not interviewing anyone on camera. The interviewers helped students understand the game a bit more clearly, without providing obvious answers to students' questions to help foster learning.

Day two of the study was when the interviewers began recording video data. At this point the students were entering class and going straight into gameplay. As stated in the theoretical overview above, it is very important when studying epistemologies and beliefs to ensure that the data collection process is tied as closely as possible to the exact context being studied. Post gameplay interviews would not be as accurate at identifying epistemologies as the interviews that our team conducted, live, during students gameplay.

Day three and Day four of the study provided the researchers more opportunities to interact with students without the burden of the cameras. This was planned out in such a fashion to ensure that all data collection happened simultaneously (unlike the pilot study) and also to ensure that students had time to advance in the game without significant interruptions. Because the study took place on six nonconsecutive days across two weeks, the theoretical hope was that students' minds had sufficient time to integrate learned knowledge about both the game and Newtonian Mechanics during these periods without interviews.

Day five was selected for the researchers to once again complete video recorded interviews, and these were conducted identically to the interviews on day two. More or less the exact same questions were asked, because the researcher was interested in understanding changes that might have occurred in student thinking from the day two interviews. Of course, additional probes and follow up questions were asked where necessary, to ensure interesting discourse and anecdotes by the students were not lost. During day five, the EBAPS assignment was also administered and so was the post gameplay level design task. This task was administered this time on a worksheet with guided instructions and a solid-line boundary depicting where students should generate their ideas for a level of the Surge videogame. The hope was that a standardized worksheet would frame the activity and increase its value more than a blank sheet of paper (as was handed out in the pilot study). These two assignments were administered on day five because the researcher had learned from previous mistakes in the pilot study not to leave anything until the last day, since students could be absent and their data unattainable.

Day six acted as a wrap up day where students could continue playing the game simply for fun, and could complete anything that they previously had not accomplished. This day also consisted of a thank you pizza party provided by the research team to thank the Beta Middle School students and Mr. Barnes for their help in the study.

Data Collection/Data Recording. Similar to the pilot study, there were 9 kinds of data that were collected for this study in Beta Middle School:

1. Pre-test answers on a physics assessment

2. Post-test answers on a physics assessment
3. Qualitative gameplay performance data measured as bronze, silver, and gold stars
4. Quantitative gameplay data such as the number of levels and trials completed
5. Students' responses to the Epistemological Beliefs Assessment for Physicals Sciences survey
6. A post-gameplay drawing activity asking students to design their own level of the Surge Game.
7. Daily field notes written by the researcher
8. Video recordings of student interviews during their gameplay on the first day of the study that were transcribed
9. Video recordings of the student interviews on the last day of their study that were transcribed

Items one through four were automatically collected by the computers as students engaged with the video game. Just like in the pilot study, nothing had to be done by the researcher to obtain this data, as the task of cleaning, encrypting, and passing on the data was performed by other members of the research team.

Items five and six were administered by the researcher as soon as students completed the last item within the video game, or during the last 40 minutes of their day five class, whichever came sooner. This was to ensure that students had plenty of time to complete both tasks, and also to ensure that absent students could be identified and asked to complete these tasks on the last day of the study, day six.

Items seven through nine were collected directly by the researcher throughout the span of the study as previously discussed. Field notes were captured daily in a notebook and later applied to transcripts since it had previously been difficult to manage the collection of random scraps of paper and electronic notes. Field notes were processed each day for key items that would be necessary to consider for the following days field work experience. The video recorded data from day two and day five were instantly copied from the video cameras, appropriately renamed, and delivered for transcription to the local company that was working with the research team. Within one week of providing video data to the company, Microsoft Word files were returned, one per student, per day of interview. These were the essence of the qualitative analysis that took place using line-by-line coding that will be discussed in the data analysis section to follow.

Compared to the mistakes made in the pilot study, the data collection at Beta Middle School was much cleaner, more streamlined, systematized, and highly balanced.

Data Analysis: Beta Middle School

Let us refer one more time to the nine types of data that were collected:

1. Pre-test answers on a physics assessment
2. Post-test answers on a physics assessment
3. Qualitative gameplay performance data measured as bronze, silver, and gold stars

4. Quantitative gameplay data such as the number of levels and trials completed
5. Students' responses to the Epistemological Beliefs Assessment for Physicals Sciences survey
6. A post-gameplay drawing activity asking students to design their own level of the Surge Game.
7. Daily field notes written by the researcher
8. Video recordings of student interviews during their gameplay on the first day of the study that were transcribed
9. Video recordings of the student interviews on the last day of their study that were transcribed

The first major decision in the analysis of this data for Beta Middle School was as follows. I made the decision for this research to focus solely on one individual video camera worth of data. Although three individuals conducted video recordings, and a total of 24 students were recorded across two different time periods, the decision to reduce the data pool to the eight students who were all interviewed directly by me felt like a sizeable chunk to begin the data analysis. I did not want to over commit to too much analysis considering the iterative nature of design research. More time could be spent on even better data acquisition and analysis down the road. Also, despite the continued conferencing and training of the interviewers, the quality of interviews across all three cameras was starkly different during a preliminary pass at the data. Thus, focusing on eight students seemed once again to be the best decision.

During the preliminary pass coupled with my memory and observation of the transcripts, the analysis of one student out of the eight students was also dropped due to little response from her to the interview questions. It would have been impossible to analyze this student's beliefs about the videogame context because she said little to give insight into her beliefs.

Items one through four were all automatically collected by the students' computers and sent directly via the intranet that was created by our research team within the school to a central server. At the end of the entire study, a portion of our research team from the Vanderbilt Computer Science department cleaned the data and removed identifiers. This team then shared the data, in the form of an extensive Microsoft Excel Workbook with the rest of the researchers at our weekly team meeting. At this point, I found the pre-test, post-test, qualitative gameplay and quantitative gameplay data for each of the seven students that I had also been responsible for interviewing. Those students' data were analyzed further in excel, turning them into bar charts that could provide meaningful comparisons.

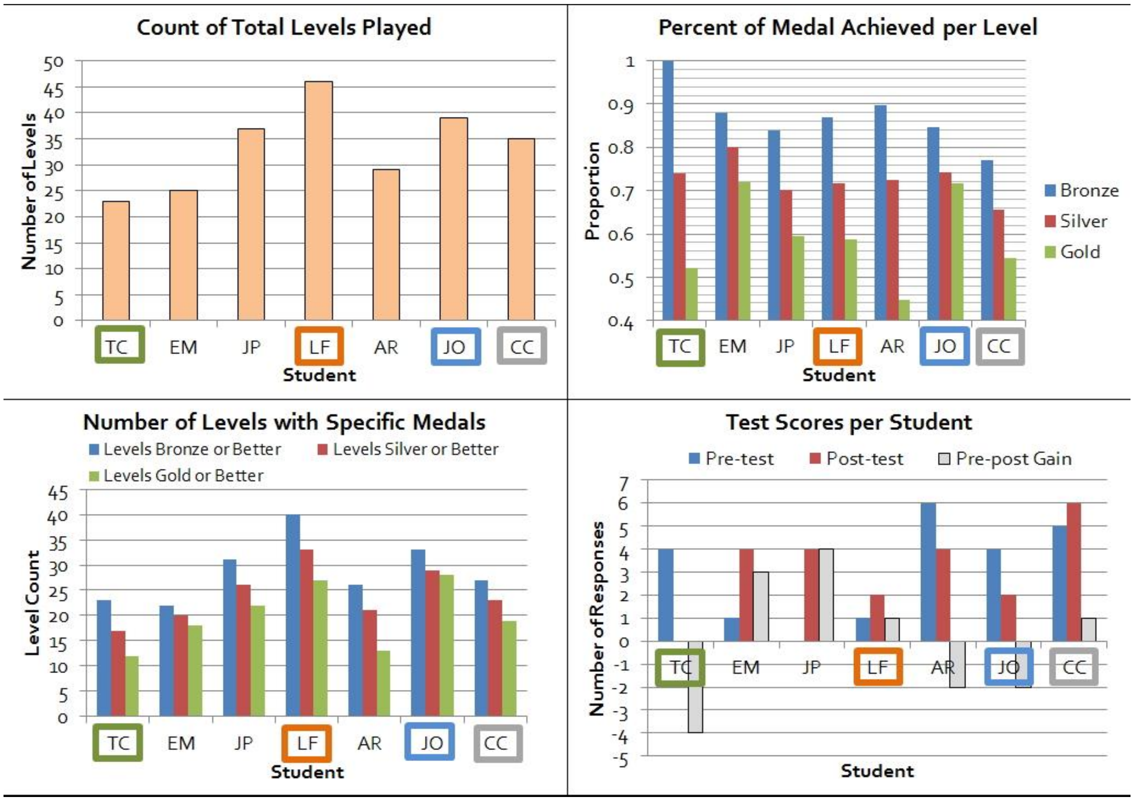

Figure 2 – Figure 5, Clockwise from the top left.

Item five, the EBAPS data, was analyzed as follows. Upon return from the middle school every day, the paper EBAPS surveys were given to a Master's Degree student worker who was responsible for entering students' responses into a large Microsoft Excel spreadsheet, thus the specific pieces of EBAPS data could be identified for each of the seven students that I was responsible for interviewing. One bar chart was created for all thirty questions of the EBAPS survey. These bar charts consisted of the *entire* population of students. This was done so that individual comparisons could be made from the seven students being studied compared to the entire populations of their class.

No quantitative statistical analysis was performed to find significant differences; instead a qualitative lens of looking for interesting trends was adopted.

All individual samples of student work from the post-gameplay activity was also gathered at the end of the study period and brought back to the university where it could be analyzed alongside the quantitative and qualitative data that was collected. Student work was reviewed after the constant-comparative-coding that was performed in NVivo 10 software in order to identify patterns and trends that matched the grounded theory that emerged from line-by-line analysis of student discourse during the interviews.

Once transcriptions were returned by the transcription company they were immediately entered into NVivo 10 software for qualitative data analysis using constant comparative methodology. As previously mentioned, the researcher watched all 16 videos for each student – two videos per student – and performed line-by-line coding for all student responses to the interviewers questions. The line-by-line coding was performed rapidly and at this level no effort was made to group lines of speech by the students. During a second pass, the interviewer looked for key aspects of the data that could then be grouped as a grounded theory began to form amongst the student talk. This emergent data began to reinforce what the researcher had made a mental note of the weeks prior during the pilot study. Student's epistemological beliefs about the Surge gameplay activity acted along two strands, gameplay-knowledge, and physics-knowledge, and the students were typically segregated one way or the other. However some students existed who could span both dimensions of beliefs about the activity.

Strengths and Limitations: Beta Middle School

Beta Middle School was a cleaner and more streamlined data collection process thanks to the experiential knowledge gained from the mistakes of the previous pilot study. The planning and training were vital to proper data collection, as was the visualization by the team of the day-to-day events that would take place. In addition, the interview protocol from the pilot study was a great tool for generating a much improved protocol for the Beta Middle School data collection. The protocol can be viewed in appendix H. These were the many strengths of the data collection.

However, many limitations still exist and there is much room for improvement in the future. It would be beneficial if the computer software could collect additional demographic information from the students including gender, age, and socioeconomic status. Some of these questions might even need to be administered to families in a separate survey since students might not have the knowledge to correctly answer the questions. It would also be beneficial if the research team could be in the classrooms for much longer than two weeks, because this would ensure that we could integrate with the students and overcome any kind of bias due to our presence to the learning environment. However, because our research involves curriculum-dominating interventions, obtaining this much time is rather difficult.

It would also be helpful if we could spend more time teaching the game, and embedding the game inside of the standard Newtonian Physics Curriculum that students are expected to know. We were lucky in Beta Middle School that Mr. Barnes did this naturally, but it was not required for our study and we know from our pilot experience that not all teachers will do this.

In the future, renegotiating the use of handheld video cameras should be pondered because during data analysis it was often difficult to see what the camera was supposed to be observing. In addition, the presence of the camera makes students apprehensive. It could be beneficial to utilize a simple audio microphone and desktop recording software like Camtasia Studio. If both of these were time-synced then students would feel much less intrusion, and the resolution of the recordings would be much higher. Finally, having consent to record student's whole bodies and faces with a future cohort, and implementing several tripod cameras that capture whole parts of the classroom could be useful in recording gestures that were not seen during this study. Gesture analysis might lead to improved understanding about the beliefs students take during this activity, since the few gestures that were caught in this study were vital in detailing our findings.

Trustworthiness: Beta Middle School

Credibility. This work could be more credible if longer engagement could be achieved with students, however due to the demands of the intervention, achieving extended classroom time is difficult. Ideally, in future work, extended interventions could be achieved with a cohort of volunteer students in a volunteer after school program designed to help improve their physics learning. In addition persistent observation could be obtained.

Negative case analysis in terms of this specific research is difficult to obtain since the grounded theory followed students along two differing strands of belief systems, and thus finding a student who is "supposed" to fall on one strand but really falls along another is difficult. What could be performed in the future is to keep our eyes open for students who score very high on physics achievement tests and conceptual understanding tests, yet for some reason still describe the Surge videogame with lots of game-based discourse and little physics discourse. Long-term micro-genetic analysis of such a student might point at specific mechanisms as to what exactly causes game based discourse to occur.

Transferability. The nature of this study is very unique, and the context is quite a unique design. However, breaking down students' discourse and attaching their discourse to the specific elements of gameplay could provide, in the future, types of gameplay elements that could be introduced into many other kinds of physics games to generate similar discourse and thus similar beliefs by students. For example, students typically think of the red arrows in the Surge game as force producers, and so red arrows in other gaming contexts, if situated similarly to Surge, might show transferability from this study.

Dependability. We did our best in this short study to ensure dependability by systematizing and streamlining the interview questions and approach towards students while they were engaged in the gaming context. Of course, a long-term after school implementation of surge, where systematized surveys could be administered once every week would greatly enhance the dependability of the data, and might even show systematic changes/growth in students beliefs and understandings. In addition, new types of questions, and the introduction of whole classroom cameras and gesture analysis could be a way of triangulating data this increasing the dependability.

Confirmability. Although the researcher did have some interesting views about what the students were seeing while they played Surge (due to observations during the pilot study) the researcher still followed a line-by-line coding protocol designed to ensure that the students were given as much voice as possible. Had the researcher coded larger chunks of student discourse, then the intent and meaning behind the students voices could have easily been lost. The use of student transcripts also contains pure raw data of many examples of the grounded theory that emerged, and the counting of percentages of student talk that fit the two emergent categories (game and physics discourse) further confirms the results of the analysis.

These trustworthiness criteria are adopted from Erlandson, Harris, Skipper, and Allen (1993).

CHAPTER V

FINDINGS FOR QUALITATIVE STUDY #1

Initial analysis of the data via line-by-line coding of all student transcripts resulted in lines being coded for many different entities. All of the codes for each line were attached to the student that they belonged to by including the students name in the code. For example, a student named Timothy who had a line mention the force is strong would have been coded “Timothy Force.” Upon completion of this initial round of line-by-line coding, I had obtained roughly 40 base level codes that spanned physics terms, stuttering utterances, fantasy a fairytale terms, and more.

Subsequent passes at the data were performed systematically to identify patterns in various ways. I first watch all the videos from the first day of interviews in alphabetical order by students’ first name, and then I watched all the videos from the second day of interviews in the same alphabetical order. As this process occurred, I considered potential groupings for the line-by-line codes that I had originally determined.

On a third pass at the data, I watched the first interview, and then the second interview for each student. I worked in reverse alphabetical order to maintain fresh eyes on different students. The majority of these observation sessions took place in five to seven hour chunks of time to fully immerse myself in the continuity of the experience.

After several iterations of observing the data it became evident that all of the students had examples of talking about the videogame and answering my questions along several strands, but the two major strands were the heavy use of physics terminology, or alternately the heavy use of game-based fantasy type language. Here is one such example:

"Interviewing T"

Interviewer: Can you tell me what this game is all about?

T: Well it's about this, well it's about this guy and his name is Surge and hes is um, he us um trying to save his friends!

"Interviewing A"

Interviewer: Can you tell me what this game is all about?

L: Newton's Laws of Motion. Like when you put the arrows, it makes the ball move where you want it to go…

Now that a pattern had developed in my iterations through watching the video I began to group these nodes in Nvivo into higher-order nodes with titles like "Canonical Physics Terms" and "Gameplay Speak." Nvivo subsequently let me view the percentage of student talk that was devoted to one strand of beliefs about the learning context, or the other. Students were marked qualitatively as a "physics speaker" or a "gameplay speaker" or "both" depending on their discourse. The following chart shows these results:

Types of Student Discourse		
Student	**Canonical Physics**	**Gameplay**
J	X	
T		X
E		X
J		X
L	X	
A		X
C	X	X

Table 1: Types of Student Discourse

CHAPTER V

CONCLUSIONS FOR QUALITATIVE STUDY #1

While this research is still preliminary in nature, the findings are significant because they substantially highlight two key strands of beliefs that students naturally adopt when exposed to a designed environment like the Surge videogame. This naturally leads to follow up questions like the following: 1) What causes students to adopt the kinds of beleifs they do about the nature of the game? 2) What design elements can be implemented to attempt to modify students beliefs prior to them engaging with the design, or during their engagement with the design. 3) What a priori experiences mold students to adopt a game orientation versus a physics orientation when interacting with such a learning tool. 4) How do students beliefs about the nature of the learning context change over time with repeated exposure?

CHAPTER VII

STATEMENT OF PROBLEM AND SIGNIFICANCE FOR QUANTITATIVE HLM STUDY #2

Focus of Inquiry – Are Students' Post-Gameplay Drawn Representations Influenced by Differences in Game Engagement?

While the previous qualitative studies attempted to observe emergent frames across the individual students that were interviewed, this study hoped to quantify links between specific gameplay behaviors that students engaged in with the hope that certain behaviors would be predictors for the complexity of drawings that students made about the gameplay environment after having engaged with Surge.

This quantitative work is, in essence, the beginnings of a potential approach to quantifying frames. Students were directed post-gameplay to draw representations of Surge levels, and rubrics were used to make scalable, quantified variables that could then be entered into a hierarchical linear model analysis that would account for the natural variation of students clustered into different types of classroom sections. This has even more important implications for future work where studies are being conducted across several schools and/or school districts, while trying to observe the frames that many students take. As the number of subjects increases, the ethnographic/qualitative work observed in Study #1 is not efficiently achievable, and thus smaller populations that are studies ethnographically should have emergent data be triangulated with larger populations in HLM studies such as this one described here. The question that this HLM study attempted to answer was: **“Does the game version, number of completed levels, or number of completed trials effect students drawing complexity and/or plausibility of their proposed force producer?”**

More Specifics of Surge

Besides the details of the gameplay environment that were provided in Chapter II of this manuscript, several other features are important to mention in terms of the mechanics of the gameplay environment. Most importantly, this quantitative HLM study took place with a more recent version of the Surge software that contained two versions of the game. Once version allowed students to control the game in real-time, by clicking on representations of force (rocket boosts/impulses) at the bottom of the gameplay screen, and these vectors of force would instantly be added to the Surge icon (Red ball), thus adding to the current state of the Surge icon in real-time. The other version of the game—the predictive version—was similar to the version of surge used in the quantitative study. In this version students planned out an entire course of vectors (rocket boosts/impulses) prior to running the level and witnessing how their predicted course influenced the Surge icon. After running the level, students could stop, review, and correct their previously planned course. Then they could run the level again, and again, until they reached the target.

This version of the surge software had also been streamlined to better collect data on the back-end of the software, and thus, obtaining the number of levels that each student completed and the number of trials that each student completed was simple.

CHAPTER VIII

METHODOLOGY FOR QUANTITATIVE HLM STUDY #2

This section discusses key aspects of the methodology that allow for a hierarchical linear model comparison of clustered groups of students in the single school where the study was performed. Portions of this section will be referenced in the conclusion when I discuss potential future scale-up studies.

Forms of Data and Collection Methods

The top level of clustering for the study involved observing that the students were naturally grouped (or clustered) into several class periods throughout the day. Qualitatively, the variation of the different class periods was quite large. One section of students were mainly non-English speakers, and other sections consisted of honors students, while some sections had a higher proportion of academically challenged students. Thus, it was assumed that some of the variation in the data could be accounted for at this higher level of clustering.

For all students, the variables of interest that were analyzed involved which version of the game they were playing, the number of levels that the students had completed overall, the number of trials that the students had completed overall, and two variables that were determined from a post gameplay drawing task. These variables were called “TotalDrawingComplexityScore” and “Plausibility17”

The post gameplay drawing task involved having the students create their own representation of a gameplay level in the surge game using crayons and pencils on a blank piece of computer paper. The students were given minimal direction, besides being told to make a challenging level for their friends, and also to show how their friends would solve the level if they were playing the game. These drawings were then observed by several individuals on the research team to determine the two variables previously mentioned.

Explanation of Variables

“TotalDrawingComplexityScore” (TDCS) was a numerical variable that could range from 0 to 5 points. Students were given one point for each of the following assets. If there drawing showed evidence of the 4 potential directions of motion, students received a point. They also received a point for evidence of a force generator (like a rocket booster), evidence of a means of thrust (like drawing fire or air-blasts), evidence of an animated character (like the anthropomorphic form of the Surge character), and evidence of a person or cockpit where the ship would be controlled from.

“Plausibility17” (P17) was a variable that ranged from 1 to 7 points, and students artwork was given more points if their representations of force producers was very plausible, usable, and showed evidence of moving the Surge icon. Scores of 1 point had little to no resemblance of the gameplay environment students had engaged with.

CHAPTER IX

RESULTS FOR QUANTITATIVE HLM STUDY #2

The final number of students that were included in the data analysis from the original pool of students at the middle school was 83. Students were removed for a lack of complete data due to computer error, and/or no post-gameplay drawing. There were a total of six different class periods of students. This was used to determine the Intraclass Correlation for the two variables of interest: TotalDrawingComplexityScore, and Plausibility 17.

The ICC for the TDCS null model was 15.73 percent, while the ICC for P17 was 6.107 percent. These numbers represent the percent of variability that exists between the 6 different classroom groups. The value for both is rather low, but statistics like this generally increase with much larger sample sizes. Considering that the sample for this beginning study was so small (at N=83, with only six groups total) there is some evidence to believe that these numbers will increase in future studies.

Data Tables and Explanations

After computing ICC for the null models for the variables of interest and noticing that some variability exists between groups, it was then justifiable to compute fixed effects for the variables of interest. Table 2 below shows estimates of fixed effects that the TDCS had on the version of gameplay (In other words, could TDCS be determined in any way by the version of the game that the students were playing.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.634164	.451229	42.781	3.622	.001	.724040	2.544288
Version	-.213395	.270310	80.398	-.789	.432	-.751289	.324499

a. Dependent Variable: TotalDrawingComplexityScore.

Table 2

Table 3 below shows whether or not the number of levels played by the students overall was a predictor of the TDCS.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.179401	.458036	44.893	2.575	.013	.256809	2.101994
level_count	.003940	.009417	79.604	.418	.677	-.014803	.022682

a. Dependent Variable: TotalDrawingComplexityScore.

Table 3

Table 4 below shows whether or not the number of trials played by the student had any impact on the TDCS.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.596646	.466086	45.978	3.426	.001	.658451	2.534841
trial_count	-.001844	.002867	81.142	-.643	.522	-.007548	.003859

a. Dependent Variable: TotalDrawingComplexityScore.

Table 4

Table 5 below shows whether or not the version of the game that students played predicted the plausibility score on their force producers.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	2.018761	.464262	57.252	4.348	.000	1.089181	2.948340
Version	-.275218	.295612	67.989	-.931	.355	-.865105	.314668

a. Dependent Variable: Plausibility17.

Table 5

Table 6 below showed whether or not the number of levels completed by students was a predictor of their plausibility score on the force producers.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.628473	.463791	53.672	3.511	.001	.698498	2.558447
level_count	.000245	.010398	68.358	.024	.981	-.020502	.020992

a. Dependent Variable: Plausibility17.

Table 6

Table 7 below shows whether or not the number of trials performed by students influenced the plausibility of their force producers.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.879362	.460412	48.099	4.082	.000	.953691	2.805032
trial_count	-.001741	.003122	68.951	-.558	.579	-.007969	.004486

a. Dependent Variable: Plausibility17.

Table 7

CHAPTER X

CONCLUSIONS FOR QUANTIATIVE HLM STUDY #2

Based on all six tables above, there was no significant relationship showing that the version of gameplay, number of levels completed by the students, or number of trials that were completed by the students in any way functioned as predictors of their TDCS or their P17.

More than likely, the major reason for this lack of significance has much to do with the low level of students who participated in the final pool of data, and this a very low power to the data analysis. Future improvements to the study would attempt to single out even more reliable predictors for the 6 groups of students, and would increase the sample size of the total population and the number of groups.

CHAPTER XI

CONCLUSIONS FOR BOTH STUDIES

Both the qualitative and quantitative studies presented here were introductory approaches to a research trajectory geared at understanding how students frame their perceptions of gaming environments. The qualitative studies followed students over two sets of interviews before and after gameplay in order to report detailed changes in their conceptual understandings of the game, and also to identify emergent frames that clustered amongst several students. These frames centered around discourse related to the game, the physics, and a combination of the two.

The quantitative study took a much larger sample of students and accounted for the variation in their clustering into class-periods of different education levels and language capabilities. The goal of this analysis was to see if certain in game behaviors, such as the gameplay type, the number of levels played, or the number of trials played, could act as predictors of how complex students would be in their generation of new levels, and how plausible their representations of force-producers would be outside of the computer screen.

Both studies showed fruitful basis for improving the study designs, increasing sample sizes, adding more predicting variables, and refining the approaches and methodologies for improved analysis. It is the hope that this manuscript will act as a building block for future researchers.

APPENDIX A

INTERVIEW GUIDES / OBSERVATION GUIDES

Version 001 – May 6th

1. What are you doing right now?
2. What is this level all about?
3. What is this activity of the game all about?
4. What is this game all about?
5. Describe what these things do (interviewer points to each)
    a. Impulses in palette
    b. Surge
    c. Fuzzy
    d. Vector display
    e. Dot trace
    f. Budget
    g. Score
6. What are you thinking when you start a new level?
7. How do you go about solving a level?
8. How would you improve your gameplay?
9. How does Surge move? Would she move this way in the real world?
10. What does this remind you of in the real world?
11. What are your own goals for the game?

12. Can you tell me about what you are learning?
13. Are you enjoying this game? Why or why not?

How do you feel right now?

How do you feel while playing?

Can you talk about your enjoyment or lack of enjoyment?

Version 002 – May 8th

1. Near the beginning of the session ask, “Can you tell me what this game is all about?”
    a. If they mention anything mechanistic ask them to explain what they mean by that (forces, motion, diagonal, sliding, etc.)
2. During any interesting levels, ask them, “Can you tell me what this level is all about?”
3. “Tell me a little bit about what you are learning?”
    a. Look for instances of “Force, vectors, arrow, etc” and probe on those things.
        i. What do you mean by Force? What’s a force?
4. “How has this level been going?”
    a. If students say “good” or “OK” ask them why.
        i. Have students compare levels: “How is this different than the last level?”
5. “What’s your strategy for solving the levels?”
    a. Probe deeper as students talk about anything mechanistic or give unclear details about their strategy.
6. Describe what these things do (interviewer points to each)
    a. Impulses in palette

b. Surge
c. Fuzzy
d. Vector display
e. Dot trace
f. Budget
g. Score

7. What would you do to become a better player of Surge?
    a. This question is looking for how students think about physics knowledge. Students will say things like "play more" "play harder" and then follow up with THIS:
        i. Why will [playing more] help you get better?
8. How does Surge move? Would she move this way in the real world?
9. What does this remind you of in the real world?
10. When you first start a new level, what do you do to solve it?
11. At the end of the session ask, "Can you tell me what this game is all about?"
12. Are you enjoying this game? Why or why not?

Version 003 – May 15th

Try your best to be systematic in every interview: similar approach, exact same ordering of questions (besides improvisational moments that arise due to interesting phenomena).

1. Approach Student and try to kneel at their level.

2. Ask them if you can watch them play for a bit. Observe them, but don't take out the camera just yet.
    a. Ask a few "test" questions about what they are doing to observe how comfortable and open they are. If they seem extremely timid or shy, move on to another student. Ideally, we would get data from anyone, but students that don't open up don't reveal their cognition at all to us.
3. Try to wait until they beat a level, or until they've interacted with the level for at least 3 minutes. If they are close to beating it, wait a bit longer to see if they do beat it.
4. Ask them if you can ask them some questions, and if they can please stop playing for a second. (This is necessary because otherwise they are trying to manage gameplay while talking and can't manage talking at the same time… cognitive overload.). Basically, they just need to know that it's OK to stop playing for a few minutes to chat. They can still show you things using the game, but we don't want them engaged in puzzle-solving.

ASK THE FOLLOWING:

5. TUESDAY 05-14-12:
    a. "So you've just started playing this game, can you tell me what it's all about?"
        i. If students mention force, motion, speed, etc. make them expand upon that
        ii. "What do you mean by speed? Can you explain that? Show me examples in the game?"
    b. "What is this level all about?"
    c. "Can you tell me how this level is similar to the level before it?"
    d. "Can you tell me how this level is similar to the level after it?"

e. “When you start a new level, what’s the first thing you think about accomplishing?”
f. “What would you do now to get better at this game?”
g. “Does this game relate to science? How so?”
h. “Pretend I’m a friend who has not learned about science, would this game be a good tool to teach me science? If yes, why? If no, why?”
    i. How would you teach a friend to play this game?
    ii. How would you teach a friend to solve this level?
i. “Does this game remind you of anything in the real world? And would Surge move this way if she were a ball in the real world and you pushed her?”

6. THURSDAY 05-17-12: Let’s not ask questions here
7. MONDAY 05-21-12:
    a. “So you’ve been playing this game now for several days, can you tell me what it’s all about?”
        i. If students mention force, motion, speed, etc make them expand upon that
        ii. “What do you mean by speed? Can you explain that? Show me examples in the game?”
    b. “What is this level all about?”
    c. “Can you tell me how this level is similar to the level before it?”
    d. “Can you tell me how this level is similar to the level after it?”
    e. “When you start a new level, what’s the first thing you think about accomplishing?”
    f. “What would you do now to get better at this game?”

g. “Does this game relate to science? How so?”

h. “Pretend I’m a friend who has not learned about science, would this game be a good tool to teach me science? If yes, why? If no, why?”

    i. How would you teach a friend to play this game?

    ii. How would you teach a friend to solve this level?

i. “Does this game remind you of anything in the real world? And would Surge move this way if she were a ball in the real world and you pushed her?”

8. TUESDAY 05-22-12: Final Quiz and Pizza and Surveys

I’m hoping that each camera can record 2 quality students per class… that’s 4 classes, which means case studies for 8 students incorporating both their pre and post intervention videos.

Part of me wants to stretch to 3 students per class… We can all talk about this…

REFERENCES

Aldrich, C. (2004). Simulations and the future of learning. San Francisco, CA: Wiley.

Csikszentmihalyi, M. (1975). Play and intrinsic rewards. Journal of Humanistic Psychology, 15(3), 41–63.

Deci, E. L. (1975). Intrinsic motivation. New York, NY: Plenum Press.

Elby, Andrew. (2011). Getting Started with Research on Epistemologies and Expectations. Department of Physics and Department of Curriculum and Instruction. University of Maryland, College Park.

Erlandson, D., Harris, E., Skipper, B., & Allen, S. (1993). *Doing Naturalistic Inquiry*. Newbury Park: Sage. (p. 161)

Ferster, C. B., & Skinner, B. F. (1957). Schedules of reinforcement. Psychological Reports, 3, 695.

Gee, J. P. (2003). What video games have to teach us about learning and literacy. New York, NY: Palgrave Macmillan.

Goffman, Erving. 1974. Frame analysis. New York: Harper & Row.

Lepper, M. R., & Greene, D. (1975). Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. Personality and Social Psychology, 31, 479–486.

Lepper, M., & Malone, T. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. Snow & M. Farr (Eds.), Aptitude, learning, and instruction (Vol. 3, pp. 255-286). Hillsdale, NJ: Erlbaum.

Loftus, G. R., & Loftus, E. F. (1983). Mind at play: The psychology of video games. New York, NY: Basic Books.

Shaffer, D. W., Squire, K. D., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. Phi Delta Kappan, 87(2), 104–111.

Tannen, D. (1993). What's in a frame? Surface evidence for underlying expectations. In D. Tannen (Ed.), Framing in Discourse (pp. 14-56). New York: Oxford University Press.