5-2014

E3: Emotions, Engagement, and Educational Digital Games

Ani Aghababyan
Utah State University

E^3: EMOTIONS, ENGAGEMENT, AND EDUCATIONAL DIGITAL GAMES

by

Ani Aghababyan

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Instructional Technology & Learning Sciences

Approved:

Taylor Martin, Ph.D Major Professor	Jürgen Symanzik, Ph.D Committee Member
Anne R. Diekema, Ph.D Committee Member	Adele Cutler, Ph.D Committee Member
Sarah Brasiel, Ph.D Committee Member	Ryan Baker, Ph.D Non-Voting Committee Member

Mark McLellan, Ph.D
Vice President for Research and
Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

ABSTRACT

E^3: Emotions, Engagement, and Educational Digital Games

by

Ani Aghababyan

Utah State University, 2014

Major Professor: Taylor Martin, PhD
Department: Instructional Technology and Learning Sciences

This study was conducted to investigate relationships between affect and engagement during student use of a digital educational game called Quantum Spectre. The study explored temporal interactions between several affective states and observed sequence of emotional states that preceded student academic disengagement. Participants included 50 Grade 5 students in Utah who played this educational game over the course of nine class sessions. The digital learning environment was designed around the physics concepts of refraction and reflection, which are Grade 6 concepts according to Utah Science Standards that can be introduced as early as the end half of Grade 5.

Previous research suggested an interesting relationship between frustration and confusion that requires more attention; the frequency of the occurrence of frustration and

confusion is influenced by the amount of external support provided. This study was designed to concentrate on significant patterns of frustration and confusion along with changes in student gameplay and engagement with the environment.

The results provide information on possible affect and behavior patterns that could be used in further research on affect and behavior detection in such open-ended digital game environments. Particularly, the findings show that students experience a considerable amount of confusion, frustration, and boredom, which hints at the possibility of a "vicious cycle" or persistence of negative affective states. Another finding highlights the need for remediation via embedded help, as the students referred to peer help often during their gameplay. However, possibly because of the low quality of the received help, students seemed to become frustrated or disengaged with the environment. Finally, findings suggest the importance of the decay rate of confusion; students' gameplay performance was associated with the length of time students remained confused or frustrated.

Overall, these findings show important transitional patterns that provide a better understanding of confusion to frustration and boredom transitions and contribute to the previously developed and hypothesized understanding of the interaction between affective states with negative valence.

(163 pages)

PUBLIC ABSTRACT

E^3: Emotions, Engagement, and Educational Digital Games

Ani Aghababyan

The use of educational digital games as a method of instruction for science, technology, engineering, and mathematics has increased in the past decade. While these games provide successfully implemented interactive and fun interfaces, they are not designed to respond or remedy students' negative affect towards the game dynamics or their educational content. Therefore, this exploratory study investigated the frequent patterns of student emotional and behavioral response to educational digital games.

To unveil the sequential occurrence of these affective states, students were assigned to play the game for nine class sessions. During these sessions, their affective and behavioral response was recorded to uncover possible underlying patterns of affect (particularly confusion, frustration, and boredom) and behavior (disengagement). In addition, these affect and behavior frequency pattern data were combined with students' gameplay data in order to identify patterns of emotions that led to a better performance in the game.

The results provide information on possible affect and behavior patterns that could be used in further research on affect and behavior detection in such open-ended digital game environments. Particularly, the findings show that students experience a considerable amount of confusion, frustration, and boredom. Another finding highlights the need for remediation via embedded help, as the students referred to peer help often during their gameplay. However, possibly because of the low quality of the received help,

students seemed to become frustrated or disengaged with the environment. Finally, the findings suggest the importance of the decay rate of confusion; students' gameplay performance was associated with the length of time students remained confused or frustrated. Overall, these findings show that there are interesting patterns related to students who experience relatively negative emotions during their gameplay.

ACKNOWLEDGMENTS

I would like to express my gratitude to the amazing committee who helped me shape this dissertation work: Dr. Taylor Martin, Dr. Ryan Baker, Dr. Sarah Brasiel, Dr. Jürgen Symanzik, Dr. Anne R. Diekema, and Dr. Adele Cutler. I greatly appreciate your research and field expertise, time, encouragement, support, and constant challenges. Your contributions have been invaluable for me as a developing researcher and for my research endeavors. In addition, I would like to thank Dr. Deborah Fields and Dr. Nicole Forsgren Velasquez for their inputs in the initial stages of this dissertation. I am grateful for your contribution and your time.

Special thank you to two individuals: my advisor, Dr. Taylor Martin, and our lab director, Dr. Sarah Brasiel. I am grateful to Dr. Martin for believing in me and selecting me to be her graduate student researcher, for challenging me and pushing me to work harder, for introducing me to a number of other researchers and teaching me how to be one, for introducing me to her family, and for being both my friend and my employer. I can never thank her enough for the trust she has in me. I would also like to thank Dr. Brasiel, who has been a true mentor in this whole process, has managed to comprehend at times incomprehensible sentences of mine, has been so understanding of my cultural background, and has shown interest in my development as a research scientist and a professional. I am very thankful for all her time and assistance that I have received ever since she joined the Active Learning Lab.

My acknowledgments go out to my lab mates, coworkers, and research acquaintances (e.g., Philip Janisiewis, XiaoTian Dai, Jason Maughn, to name a few), who

have influenced my dissertation work by constant questions, rigorous suggestions, feedback, and assistance. In addition, I would like to thank True Rubal from the IRB office for being so helpful and patient with me in regards to my dissertation IRB process. My warmest regards go to the elementary school and its teachers who welcomed my study and helped me conduct it and also to those amazing fifth graders who showed such interest in my work.

Of course, none of this would have had been possible without the constant encouragement of my parents and my dear friends who have studied with me, stayed up long nights to keep me company, encouraged me, and consoled me at times. My achievements are partially theirs.

Finally, I would like to dedicate this dissertation to my two grandfathers for being such amazing examples of dedication, hard work, and scholarité. They taught me how to strive for the best and work hard to achieve it, no matter the quantitiy and the size of obstacles on my path to success. Persistence is what I have learned from them and what has helped me stay focused on my education.

Ani Aghababyan

CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

Affect – This term refers to a student's emotional state or emotional experience within an educational environment. In this dissertation I use the word *affect* interchangeably with the following terms that appear in the literature: *affective state* (Pekrun, Goetz, Titz, & Perry, 2002; Rosenberg, 1998), *emotion*, *emotional state* (Baumeister & Bushman, 2007), and *cognitive-affective state*. In the affect literature, some theories differentiate between the words affect and emotion (e.g., Ekman, 2005). However, for the purposes of this dissertation, these two concepts are interchangeable.

Persistent frustration and *persistent confusion* – This is frustration and confusion that occurred two or more times in a consecutive order of discrete observations (e.g., observations over 15-second "clips" of student interaction with the game environment).

Simple frustration or *simple confusion* – This is frustration and confusion that occurred one time (i.e., a single occurrence) in a consecutive order of affect recordings.

Student engagement – This term (also *students' engagement* or *student's engagement*) refers to a student's engagement in an academic task or engagement with a learning environment.

Video game – These are electronic games that involve user interaction, interactive displays and visual feedback. In this dissertation I will be using video games as a reference to *digital game for learning, digital game-based environments,* or *educational digital game*. There is a clear distinction between commercial video games and educational video games. However, this dissertation concentrates only on the educational video games. Hence, unless otherwise specified, all the references to video games are synonymous to educational games for learning.

Descriptions of affect categories are taken from the Baker-Rodrigo Observation Method Protocol (BROMP) manual (Ocumpaugh, Baker, & Rodrigo, 2012) or the coding scheme developed by Baker, Corbett, Koedinger, & Wagner (2004):

Bored (B) – According to Baker et al. (2004), this affective category identifies with behaviors such as slouching; resting the chin on own palms; and statements such as "Can we do something else?" "This is boring!" or "This is not fun anymore."

Confused (CF) – This is a noticeable lack of understanding. According to the BROMP coding guide (Ocumpaugh et al., 2012), it may include student behavior such as scratching own head, repeatedly looking at the same interface elements, consulting with a peer student or seeking help from the teacher, peeking over to another student's screen to find solutions or see what the peer did, and statements like, "I'm confused!" or "Why didn't it work?"

Concentrating (C) – This engaged concentration manifested by visible "immersion, focus, and concentration on the system, with the appearance of positive engagement: leaning towards the computer; mouthing solutions; pointing to parts of screen" (Baker et al., 2004, p. 12).

Frustrated (F) – This occurs when the student faces an unresolvable "impasse" or the student has no action plan on how to overcome the barrier (N. Stein & Levine, 1991). According to Baker, D'Mello, Rodrigo, & Graesser (2010), this includes banging on the keyboard or throwing the mouse, pulling own hair, deep sighing, as well as statements such as, "What's going on?" "This is so frustrating!" or "There is no way to solve this!"

Delight (D) – Delight is a student's expression of pleasure with the results at the moment. According to the coding guide, it may include behavior such as clapping of hands and laughing with pleasure as well as statements such as, "Yes!" "I got it!" or "Yay, it worked!"

Surprise (S) – This is identified as "sudden jerking or gasping" and statements such as, "Huh?" or "Really!" (Baker et al., 2010).

Eureka (E) – Eureka is student's expression of sudden understanding of the task at hand. This is identified as student's "Ah ha" moment. This may include statements such as "Aaah" and "Ah ha".

? (= "other") – This code typically means a student could not be coded for affect because of physical absence or when the observer could not continue coding since that particular student noticed the field observation taking place.

Behavior categories are also coded according to the coding scheme presented in Baker et al. (2004) or the BROMP manual (Ocumpaugh et al., 2012):

On task (OT) – This refers to engagement in (working with) the assigned learning environment.

Other on-task conversation (OOC) – This typically refers to a student who is working on the task while casually chatting with a peer on topics related to the task (e.g., "What level are you on?" "See what I did there, I beat it") or to the teacher ("Have you played this game before?").

Off task (OfT) – This refers to a student who ceases to engage in the assigned task but rather engages in unrelated tasks. Karweit and Slavin (1982) referred to this as "disengaged behavior."

Receiving help (RH) – This refers to a behavior where the student is having an on-task conversation, receiving task-related help from another student who is working on the assigned task.

Giving help (GH) – This typically means a student is having an on-task conversation, providing task-related help to another student who is working on the assigned task.

? (= "other") – This code typically means a student could not be coded for behavior because of physical absence or the observer could not continue coding since that particular student noticed the field observation taking place.

Engaged and disengaged – Engagement is the level to which learners are involved in the academic task. Some authors have regarded engagement and disengagement as elusive constructs (Corno & Mandinach, 1993). In this dissertation, I have used the BROMP manual operationalization of the concepts *engaged* and *disengaged.* For the purposes of my study, engaged students are on task (i.e., visibly involved with this study's assignment, Quantum Spectre game). Disengaged students are off task (i.e., visibly distracted or disconnected from the study task, gameplay in Quantum Spectre).

Affect intensity – Also termed *emotional variability,* this is the intensity of a positive or negative emotion (Oosterwegel, Field, Hart, & Anderson, 2001). Larsen and Diener (1987) posited a consistent individual difference in the affect intensity manifested among different individuals.

CHAPTER 1

INTRODUCTION

Problem Statement

Digital educational games are popular for instruction and practice (Dempsey, Haynes, Lucassen, & Casey, 2002; Dempsey, Lucassen, Haynes, & Casey, 1997; Foreman, 2004; Mayo, 2009; O'Neil, Wainess, & Baker, 2005; Prensky, 2001; Rodrigo et al., 2008; Squire, 2007) in a variety of domains (e.g., science, technology, engineering, and mathematics). This instructional approach is mainly justified by the observation that these games can naturally motivate students to engage with the environment and learn (Barab et al., 2007; Entertainment Software Association, 2001; Gee, 2007a, 2007b; Kapp, 2012; Lenhart, Madden, Smith, & Macgill, 2007; Prensky, 2001). Digital games provide students with a safe space for failure that ideally helps instill confidence to persist (Juul, 2013), and they allow students to experience a variety of educational concepts via fun virtual environments.

From the perspective of educational institutions, digital games provide a unique advantage of simultaneous accessibility for thousands of children, along with an opportunity to customize learning pace and the ability to follow students' learning trajectories. Most importantly, compared to current formal learning approaches, digital game-based environments are posited to be good at keeping students motivated with the learning process (Prensky, 2005). As such, digital games are potentially powerful tools for learning (Federation of American Scientists, 2006) that also afford researchers the ability to investigate student–computer interactions down to clickstream granularity by

the analysis of log data. Nevertheless, many factors may influence student engagement with such digital learning environments. One of these factors is students' emotional experience; educational environments are social settings infused with emotional experiences.

Research has found engagement indispensable for the accomplishment of learning tasks regardless of the learning environment (Corno & Mandinach, 1983; Pintrich & Schrauben, 1992). A positive link has been observed between the level of students' engagement and their learning progress (Wigfield et al., 2008). Researchers have observed individual differences such as student self-efficacy, goal orientation, beliefs, attitude, and achievement goals as factors for the variance in the levels of engagement (Bandura, 1977; Zimmerman, 1995). There are different opinions of what motivates learners to stay engaged in digital learning environments (Dickey, 2005, 2006; Fisch, 2005; Waraich, 2004). Linnenbrink (2007) hypothesized emotions to be the key mediator between individual differences and levels of engagement.

According to Goetz, Pekrun, Hall, and Haaf (2006), academic emotions are directly linked to learning, classroom instruction, and achievement; these factors can either benefit or undermine students' engagement and learning. There is a complex interaction between affect and learning (Baker, D'Mello, Rodrigo, & Graesser, 2010); some affective states are positively correlated with learning outcomes, whereas others are associated with negative academic performance. Moreover, a student's affective state may be manifested differently depending on the learning environment, length and order of the affective states and other context-related factors. D'Mello (2013) found support for

this idea in his meta-analysis on affective states, where emotions were found to be "highly situation-dependent and contextually-coupled" (p. 30).

Due to this relationship between engagement and learning, lack of student engagement can be a threat to learning; disengaged students may not take full advantage of the learning opportunities offered by digital educational games. Given the possible negative impact of emotions on students' academic performance, it is important to understand students' affective responses to success and failure in learning environments and factors that may influence student emotions. Although there is plenty of research on different affective states, frustration and confusion are two affect states for which the research findings have been mixed with respect to their influence on learning (Baker et al., 2010; Craig, Graesser, Sullins, & Gholson, 2004; D'Mello, Picard, & Graesser, 2007; D'Mello, Taylor, & Graesser, 2007; Rodrigo et al., 2009). To further investigate frustration and confusion, in this study I used a combination of quantitative field observations and sensor-free, data-driven methods to infer important sequences of several observed affective states and disengagement. Understanding important sequences of affective states can help to determine points at which games might adapt to learners' affective state appropriately. The findings from this investigation may contribute to the improvement of digital educational game design.

Research Purpose and Questions

This dissertation investigated whether sequential patterns of student emotional states were exhibited in digital game environments that preceded boredom and off-task behavior. Discovery of such patterns might help to accurately (statistically) model the

order and temporal sequence of emotional states that lead to disengagement. In particular, I investigated the following questions:

1. What are some of the most frequent affect-behavior patterns?
2. What combinations of affective and behavioral states frequently precede boredom?
3. Are there sequential patterns in the occurrence of affective states, especially frustration and confusion, that are consistently associated with boredom?
4. What combinations of affective and behavioral states frequently precede off-task (disengagement) and receiving-help behavior?
5. Are there sequential patterns in the occurrence of affective states, especially frustration and confusion, that are consistently associated with off-task (disengagement) and receiving-help behavior?
6. Do students' affect sequences characterize their performance in the game?

Conjectures

In this work, I operationalized persistent frustration and persistent confusion as frustration and confusion that occurred two or more times in a consecutive order of discrete observations (e.g., observations over 15-second "clips" of student interaction with the game environment). Where "Ft" is an observation of frustration at time point "t," persistent frustration, for example, would occur in the following sequence <F1, F2, F3>. In addition, I operationalized simple frustration or simple confusion as frustration and confusion that occurred one time (i.e., a single occurrence) in a consecutive order of

affect recordings. For example, two instances of simple frustration occur in the sequence <X1, F2, X3, F4>, where "Xt" represents any affective state other than frustration.

Based on this operationalization, I speculated that frustration would occur following exhibition of persistent confusion: After two or more occurrences of confusion, student would experience frustration (e.g., <C1, C2, F3>). In addition, I also anticipated finding that persistent frustration would develop if, after the first exhibition of frustration, the student did not receive assistance. Related to the previous assumption, I expected that persistent frustration would transition into confusion, simple frustration would transform into concentration, and persistent frustration would transform into concentration after the students received help. Finally, I predicted that students would be bored only after experiencing persistent frustration, and off-task behavior would transition back to on-task behavior only after students receive external help.

Dissertation Outline

This dissertation is structured as follows. Chapter 2 covers a review of the literature on several areas of my research interests: theories of engagement, theories of affect, educational digital games, and finally, the current status of science education in the United States. Chapter 3 describes the research methodology, which includes study participants, materials, data collection instruments, and data sources. Chapter 4 describes the findings in regards to the sequential relationship between affective states and student engagement, and Chapter 5 discusses study contributions and implications of study findings as well as limitations and suggestions for future work.

CHAPTER 2

REVIEW OF THE LITERATURE

Whereas many factors impact student learning and engagement in digital learning environments, affective states are a prominent theoretical and empirical research topic. Findings have suggested significant associations between affect and various learning outcomes and student academic performance (Blair, 2002; Gumora & Arsenio, 2002; Raver, 2002; Stein & Kean, 2000). Many studies have investigated the association of emotions and student academic performance in intelligent tutoring systems (Pardos, Baker, San Pedro, Gowda, & Gowda, 2013; San Pedro, Baker, Bowers, & Heffernan, 2013), but similar investigations have yet to be carried out in digital game environments for learning.

Sources

This literature review identifies and synthesizes research in the fields of digital educational games, learning analytics, and educational data mining with particular attention to studies that identified ways to collect affect data while students are engaged in digital environments for learning. The research questions for this literature review are the following:

1. What are some of the prominent theories of engagement and emotion?
2. Is there a connection between emotions and students' learning, academic performance, and achievement?

3. What have prior studies found to be important to understand about digital learning environments?
4. Why is science an important area of education research?

To identify the approaches used in literature for measuring and inferring student affective states and to answer aforementioned review questions, I consulted multiple databases including EBSCOhost, ERIC, PsycINFO, and Web of Science. In addition, I consulted well-known researchers in the particular research fields to discover further relevant literature. For literature search I used different combinations of key terms and expressions, such as *affect, emotion, cognitive-affective state, motivation in education, engagement, field observation, affect detector, emotion detectors, video games, digital games, video games for learning*, and *educational games.* When a relevant article was identified, I also made sure to search through its reference list in order to widen my overview of literature. As a result I identified additional key terms such as *cognitive engagement, academic emotions, digital games for learning, game-based environments,* and *game-like environments*.

Based on my review of the literature, I next summarize the key findings that informed my study design. This literature review has five main parts: (a) theories of engagement, (b) theories of emotion, (c) frustration and confusion affective states, (d) education game environments, and (e) current conditions in science education.

Theories of Engagement

Some authors have considered student engagement a combination of behavior, affect, and cognition (Fredricks, Blumenfeld, & Paris, 2004; National Research Council

& Institute of Medicine, 2004; Newmann, Wehlage, & Lamborn, 1992). With respect to student cognition and engagement, negative emotions are posited to divert learners' cognitive resources to focusing on the object of the emotion rather than on the educational material (Blair, 2002).

Motivation and engagement are central to the understanding the influence of emotions on students' performance (Izard, Stark, Trentacosta, & Schultz, 2008). In particular, several researchers have indicated an association between motivation and student achievement (Ladd, Birch, & Buhs, 1999; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006) and the potential of motivation mediating the relationship that exists between emotion and achievement. Motivation theorists have long considered emotions as essential elements of their framework (Schunk, Pintrich, & Meece, 2008); however, only in the past decade have cognitive engagement and learning theories also acknowledged emotions as a central element (Goetz et al., 2006; Linnenbrink & Pintrich, 2002; Pekrun, Elliot, & Maier, 2006; Pekrun et al., 2002). Whereas motivation theories, such as intrinsic motivation theory, are important to explain students' reasons for engaging in online environments (Shroff & Vogel, 2009; Wighting, Liu, & Rovai, 2008), there are more explanatory factors than just intrinsic motivation. Therefore, this study focused on student affect as another possible determinant of engagement (i.e., the relationship between students' affect and their on-task engagement and disengagement).

Theories of Emotion

Baumeister and Bushman (2007) defined emotions as subjective states that are typically accompanied by a certain bodily reaction and represent an evaluative response

to stimuli. Izard et al. (2008) added to this definition the motivational component of emotions and their influence on human cognition. Other definitions of emotions by Damasio (2004) and Scherer (1984) described emotions as extending beyond affect while also including a motivational component. "Emotions are seen as multi-component, coordinated processes of psychological subsystems including affective, cognitive, motivational, expressive, and peripheral psychological processes" (Pekrun, 2006, p. 316).

In Pekrun's (2006) social cognitive control-value theory, student achievement emotions are tied to their cognitive appraisal (motivational beliefs) and academic success. These achievement emotions can impact student learning but also can be mediated by cognitive mechanisms such as student learning strategies or persistence (Pekrun et al., 2002). Based on these theories, it was important for my study to consider the possible connection between emotions and academic success. In fact, my primary research interest is in the relationship between academic success and potentially negative emotions, the latter of which has been negatively linked to motivational beliefs and academic success (Pekrun et al., 2002).

Because of its integrative and complementary overview of emotions and their categorization, Pekrun's framework on emotions (i.e., control-value theory) is most aligned with my primary research interest. This framework is built upon some of the assumptions of several models and theories of emotions, such as expectancy-value theories of emotion (Pekrun, 1984, 1988, 1992; Turner & Schallert, 2001), models of addressing an emotion's impact on learning and performance (Fredrickson, 2001; Pekrun, 1992; Pekrun et al., 2002; Zeidner, 1998, 2007), and theories of perceived control

(Patrick, Skinner, & Connell, 1993; Perry, 1991, 2003). Pekrun's (2006) framework includes two types of achievement emotions: activity emotions that are related to achievement activities and outcome emotions that are connected to the outcome of these activities. Some examples of achievement activity emotions are enjoyment, frustration, and boredom, whereas outcome emotions include sadness, shame, hope, and anticipatory joy.

Research by Baker and colleagues has offered additional insight into several affective states that are integrated in control-value theory (Baker et al., 2010; Lehman, D'Mello, & Graesser, 2012; Rodrigo & Baker, 2011; San Pedro, Baker, & Rodrigo, 2011). My research relates to their work on boredom, which suggested that boredom is worse than frustration because it is much harder to reengage students once they have become bored (Baker et al., 2010). In addition, boredom has been found to be a persistent affective state in many learning environments (Baker et al., 2010). However, since students do not generally get to a state of boredom immediately after being introduced to a learning environment, it is assumed that several other emotional states precede, and possibly contribute to, boredom. The ability to control for or resolve these specific emotional states within an educational environment could prevent boredom.

Frustration and Confusion Affective States

Baker et al. (2010) claimed that confusion and frustration may be unavoidable and relatively natural when students are faced with difficult learning material. According to the same study's findings (Baker et al., 2010), confusion is linked to learning gains, whereas frustration is associated with boredom (Perkins & Hill, 1985). Hence, in both

cases, it is crucial that the digital learning environment is able to handle students' confusion and frustration productively in order to strengthen deep learning and avoid boredom, which may result in poor learning and disengagement with the environment.

In Pekrun's (2006) framework related to the theory of emotions, frustration and boredom are activity emotions, but confusion is not represented in the framework. Literature has shown that confusion is not necessarily negative; it may impact learning negatively or positively (Rodrigo, Baker, & Nabos, 2010). In fact, resolvable confusion has been found to be quite enriching for the learning experience and may encourage deeper engagement (Kapur, 2008; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003). Unfortunately, some learners may give up their efforts when they experience confusion, since these learners tend to attribute their confusion to their lack of abilities (Dweck, 2002; Meyer & Turner, 2006). On the other hand, according to Schwartz and Martin (2004), Kapur (2008), and Kapur and Bielaczyc (2012), leaving students to experience failure may elicit productive failure, which has been shown to have a positive impact on learning (VanLehn, 1999; VanLehn et al., 2003). Moreover, inducing confusion has been shown to promote deeper exploration and learning (Lehman et al., 2012) and to be positively correlated with learning (Craig et al., 2004). Since confusion can be productive, it is not clear if and when interventions or remediation may be required to benefit learning. According to Mentis (2007), it is also not clear if and when frustration requires intervention, since it does not always require remediation. In regards to the length of an affective state, some researchers have brought up the idea of persistence of certain cognitive-affective states (D'Mello, Taylor, et al., 2007). D'Mello, Taylor, et al.

(2007) suggested the existence of a “vicious cycle” with regard to the persistence of some affective states over time; some affective states can be very persistent, creating a need for a response by the learning environment, especially when these emotional states have a negative valence.

With respect to frustration, the research findings have been inconsistent or have shown no effect on learning (Craig et al., 2004; Rodrigo et al., 2009). Depending on the length of the affective state or circumstances under which it occurs, affect can be manifested in many different ways (D’Mello & Graesser, 2011). Some studies have indicated the relevance of frustration to learning (Baker et al., 2010), but the characteristic of this relevance is not consistent (e.g., whether frustration is negatively or positively related to poor learning). Others have emphasized the possibility of frustration turning into boredom if persistent frustration is not resolved (D’Mello & Graesser, 2012). According to Gee (2007a), frustration can manifest in a pleasantly frustrating form within game environments, which implies that it may not always be a negative factor for learning outcomes.

Unfortunately, there is limited research on the sequential occurrence of confusion and frustration and the relevance of each to student engagement in digital educational game environments. Liu, Pataranutaporn, Ocumpaugh, and Baker (2013) suggested that the negative impact of these two affective states may be larger when they occur together. Understanding the relationship between frustration and confusion is important to consider when designing learning environments. Although both frustration and confusion are well researched in intelligent tutoring systems environments (Lehman et al., 2012; Lehman et

al., 2011), learning environments like digital games differ in that they do not have built-in support systems; disruption of engagement may be more destructive in digital educational game environments if frustration and confusion are not addressed. Given the dearth of treatment of these issues within digital education game environments, emotions, particularly academic affect (defined earlier), are central to my work because of their influence on student achievement and academic success.

Educational Game Environments

Video Games and Learning

To understand the domain of digital games for learning, it is important to begin with a review of research on video games in general and their impact on learning. Video games are one of the latest forms of multimedia software to penetrate the field of education as a potential teaching resource. In fact, a literature review on games and learning (Kirriemuir & McFarlane, 2004) highlighted that a growing body of research has identified games as the most frequently used interactive media among children (Beentjes, Koolstra, Marseille, & van der Voort, 2001; Feierabend & Klingler, 2001). According to market research, 91% of all children between the ages of 2 and 17 (approximately 64 million) play video games (NDP Group, 2011), which is a 9% increase over 2009. Most of these games are not simply tools for idle amusement but offer visual, audio, and kinesthetic experiences along with a storyline that keeps the users engaged while acquiring or improving a variety of types of skills, including visual, attention, problem solving, logical thinking, speed, accuracy, multitasking, and managing fear of failure (Higgins, 2000; Inkpen, Booth, Gribble, & Klawe, 1995; Whitebread, 1997). Some

researchers have found that playing successful games can promote students' systematic thinking skills (Squire, 2003) and collaborative problem solving (Steinkuehler & Chmiel, 2006).

Digital Games for Education

Digital games for education are a subset of video games that are either adapted commercial video games or specifically developed for educational purposes. Prensky (2001) suggested that effectiveness of educational games depends on the balance between a fun interface and educational value. A wide variety of digital educational game options is available for nearly any subject matter: mathematics (e.g., Klawe, 1999), geography (e.g., Virvou & Katsionis, 2006), history (e.g., Squire & Barab, 2004), engineering (e.g., Ebner & Holzinger, 2006), and science games (e.g., Magnussen, 2005), among others.

Engagement in Digital Games

Despite different opinions of what motivates learners to engage in video games in general, some authors have suggested that it is the narrative or the storyline that motivates users (Dickey, 2005, 2006; Fisch, 2005; Waraich, 2004). Others have claimed that it is video games' interactive experience that allures learners and intrinsically motivates them (Ebner & Holzinger, 2006; Hämäläinen, Manninen, Järvelä, & Häkkinen, 2006; Kambouri, Mellar, & Logan, 2006; Klawe, 1999; Squire & Barab, 2004). Yet other research studies have highlighted the importance of rewards systems and the act of playing itself (Amory, Naicker, Vincent, & Adams, 1999; Denis & Jouvelot, 2005; Jennings, 2001). These games, although fun, are developed to improve student

engagement and learning outcomes. Digital game developers have to determine the balance between designing a game with the right amount of play and learning activities.

There is empirical evidence that video games in general positively influence motivation and learning outcomes (Ebner & Holzinger, 2006; Mayo, 2009; Ricci, Salas, & Cannon-Bowers, 1996; Squire & Barab, 2004). However, many of these empirical studies are small scale (Becker, 2001), limiting the conclusions that can be drawn about the effectiveness of such environments (Dziabenko, Pivec, & Schinnerl, 2003). Hence, before researchers can accurately evaluate digital games for their effectiveness, it is important to understand mechanisms that are embedded in these digital environments. These can include the game's design, level of assistance, adaptive feedback, flexibility for customization, as well as its ability to convey intended concepts. Automated and personalized affective response systems should be incorporated into game design, as they may support important mechanisms for learning.

Emotions and Digital Games

McGonigal (2011) has suggested that some of the most intense emotional experiences are provoked during digital gameplay. If true, well-developed educational games with adaptive support systems should become one of the key ways to assist students to push past learning related obstacles and develop persistence, regardless of their negative emotional response to a potential failure within the learning environment. A learner's self-efficacy (i.e., beliefs about one's abilities and the effectiveness of one's effort) may influence that learner's affective response to the success and failure experienced within a learning environment, thus disengaging and discouraging the

learner (Dweck, 2002). Finding ways to change student beliefs about their abilities is important to increasing interest, engagement, and persistence, which have been shown to be related to positive long-term outcomes for students (Duckworth & Seligman, 2005). Digital games provide students with the opportunity to experience multiple failures but also with multiple opportunities to succeed. Because there is a balance between the opportunity to fail and the opportunity to succeed, learners may feel comfortable experiencing failure since they develop the confidence that they have everything necessary for their eventual success.

In summary, digital games are becoming prominent in the domain of education, and researchers should contribute to the improvement of these digital learning environments in order to capitalize on their advantages. While there is potential for such environments to improve learning outcomes for students in science, technology, engineering, and mathematics, this work focuses specifically on science education and better understanding patterns of student emotional response during gameplay in a science digital environment.

Science Education

Achievement data in science for elementary and secondary school students in the United States have displayed worrisome trends (Grigg, Lauko, & Brockway, 2006). According to tests scores from the 2006 Programme for International Student Assessment, 15-year-old students representing the United States scored below the average compared to other 30 industrialized nations (Organisation for Economic Co-operation and Development, 2007).

This lack of knowledge and low achievement are partially attributed to current approaches to science education that establish among students an absence of interest and motivation toward science education (National Research Council, 2005). Despite their initial (possibly innate) curiosity with which students come to school, current approaches to science education lack the ability to support and maintain that interest (National Research Council, 2011). Because of the intense concentration on school achievement tests, students lose their interest in science education, especially during their transition from elementary to middle school (Cavallo & Laubach, 2001; Cohen-Scali, 2003; Gibson & Chase, 2002; Ma & Wilkins, 2002). According to a national survey conducted among middle and high school students, only half of the respondents regarded science as important for their future academic success, and only about 20% voiced interest in a career in science (Project Tomorrow & PASCO Scientific, 2008).

Hence, there is a need to stimulate interest and increase engagement among students towards science. Computer games and similar digital learning environments provide an opportunity to meet this need. In particular, these environments allow learners to observe and manipulate natural phenomena via a virtual environment (e.g., mirror reflection, retraction, vaporization, force of gravity, etc.), which otherwise would be difficult or impossible to experience in a traditional classroom setting (Honey & Hilton, 2011). These environments have the potential to interest learners and maintain their motivation within the learning setting. In addition, from an instructional design perspective, these environments enable educators to adaptively tailor science instruction to the needs and performance of the student.

Summary

Since there is evidence of an important relationship between emotions and learning (Baker et al., 2010; Dragon et al., 2008; Lee, Rodrigo, Baker, Sugay, & Coronel, 2011; Liu et al., 2013; Sabourin, Rowe, Mott, & Lester, 2011), well-designed games that contain adaptive support systems may be one way to positively reinforce productive relationships between emotions and learning while inhibiting less productive relationships. If external help is always available, like in intelligent tutoring system environments, there will be the danger of overusing it (e.g., "gaming the system," as noted by Baker, 2011). In addition, "mystical" and "exploratory" aspects of games are presumably elements that make them engaging and are important to preserve. Therefore, additional research is needed to examine students' emotional response patterns in digital games in order to develop functional automated, adaptive, emotion-response systems in the context of digital game environments.

CHAPTER 3

METHODOLOGY

To investigate emotional states exhibited by students in an open-ended digital game environment (Quantum Spectre), I designed an exploratory study in which fifth-grade students interacted with a science digital learning environment during each of nine 45-minute class periods. I chose this game environment because of my experience with Quantum Spectre and other educational digital games and what I perceived to be an important content for fifth-grade students. In this section, I describe the research (a) participants, (b) materials, (c) data collection instruments and data sources, (d) procedures, and (e) data analysis methods used in the study.

Participants

Participants selected for this study included fifth-grade students from a public elementary school in a rural area in Utah. All fifth-grade students at this school were invited to participate in the study ($N = 53$). However, data were collected only for those students whose parents expressed written consent for their child's participation and students who participated in all nine sessions. Due to the school's class schedule, the study was conducted at two separate times to accommodate two classrooms. Classroom A participated in the morning, and Classroom B participated in the afternoon. The same study design and data collection procedures were used for both classrooms. The study

was conducted during students' keyboarding class periods, and students were not offered any compensation for their participation.

Two students were eliminated from the data analysis leaving the study with 51 participants' data; one student switched schools in the middle of the study, and the other student was absent for a majority of the study sessions due to family travel. In Table 1, I summarize participants' average age and gender information. Due to administrative restrictions, I was not able to acquire any other demographics data besides student gender and student age. Although the sample is small, research has suggested that the sample size is sufficient for quantitative field observations to observe student affect and behavior (Baker et al., 2010; D'Mello & Graesser, 2011).

Table 1

Study Participants by Classroom

Class statistic	Classroom A	Classroom B	Total
Total number of students	26	25	51
Total number of boys	14	15	29
Total number of girls	12	10	22
Average age	10	10	10

Materials: Quantum Spectre

The environment for this study is the Quantum Spectre (n.d.) digital game, which was designed and developed for educational purposes. The selection of this game was driven by several factors: It is single player, noncommercial, and was developed for

educational purposes to treat certain topics from science. It is a puzzle-style online game created by Boston's Educational Gaming Environments (EdGE) group at the Technical Education Research Center (TERC), a math and science research-focused organization (see Appendix A for game-level visualizations and descriptions). The game is designed for middle and high school students; however, it covers concepts introduced for the first time in fifth- and sixth-grade science curricula in the state of Utah, including reflection, angle of reflection, refraction, and optical spectrum. These topics are included in the required curricula for sixth-grade science education in Utah (Utah State Office of Education, 2002); however, they are practiced at the end of fifth grade at an introductory, conceptual level.

At each level, students have access to an inventory of resources for the level, such as flat and curved mirrors, lenses, filters, and beam-splitters, among others (see Figures 1–3) to guide laser beams into colored targets on the game board. In Figure 1 the students are given a flat mirror: in order to direct the red laser into the target, they have to position the mirror in a way that the reflection points at the target. The same is true for Figure 2 where students are given a curved mirror. There is a difference between the angles of reflection when using a flat vs. a curved mirror. Depending on the level, there will be one or more laser beams for the player to manipulate (see Appendix A for detailed explanations on game techniques). Each level requires the player to direct provided laser beams to targets while avoiding barriers. Players succeed at a level and move to the next one when they successfully reach all of a level's targets.

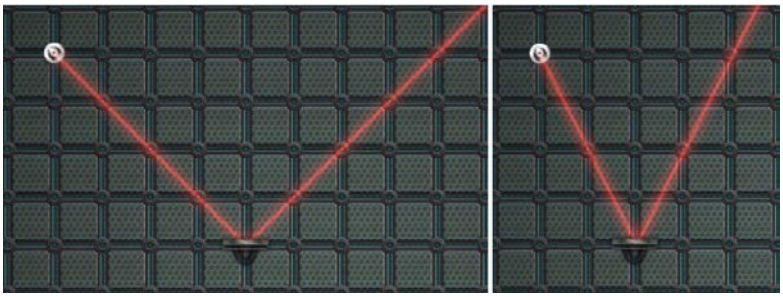


Figure 1. Flat mirrors in the Quantum Spectre game.

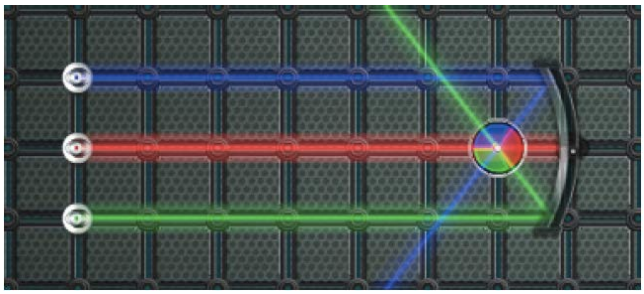


Figure 2. Curved mirrors in the Quantum Spectre game.

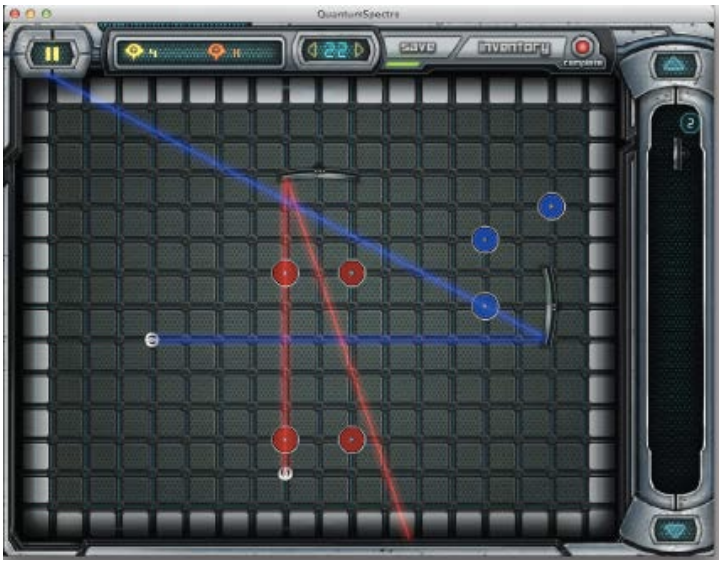


Figure 3. Lenses of two different colors in the Quantum Spectre game.

TERC has presented Quantum Spectre at the Game Arcade (Edwards, Bardar, Asbell-Clarke, & Larsen, 2013). However, no formal studies have been conducted until

recently, when EdGE at TERC launched its first formal study using Quantum Spectre. The researchers began formally collecting game data for the first time during the spring of 2014.

Data Collection Instruments and Data Sources

Baker-Rodrigo Observation Method Protocol (BROMP)

The BROMP is a method for conducting quantitative field observations developed by Ryan Baker and Mercedes Rodrigo (D'Mello, Picard, et al., 2007; Ocumpaugh et al., 2012). It is used to record observations of student behavior and affect in field settings. It provides synchronizable data that can be combined with technology-based learning-environment data (e.g., clickstream log data) to investigate student engagement and emotions along with their interaction with student performance within the game environment.

In this study, I followed the recommendations of the BROMP's developers; hence, I employed face-to-face observations in the same room as the students and at a presumably unobtrusive angle from the observed student in order to decrease the chance of interfering with a student's gameplay. To become a certified expert coder, my interrater reliability for using this method was conducted with an expert guide throughout BROMP coding training process. In comparison with the expert guide, my reliability as a coder was established with Cohen's (1960) kappa at values of $k = 0.78$ for cognitive affective states, which meets the accepted threshold of 0.75 for field coding (Bartel & Saavedra, 2001; Fleiss, 1981). Cohen's kappa is an agreement measure between two raters, an inter-rater reliability value used when observing categorical data.

Human Affect Recording Tool (HART)

Students' behavior and affect were observed according to the BROMP while data were recorded using the HART (see Figure 4). HART is a Google Android application (Baker et al., 2012) that implements the BROMP, described earlier.

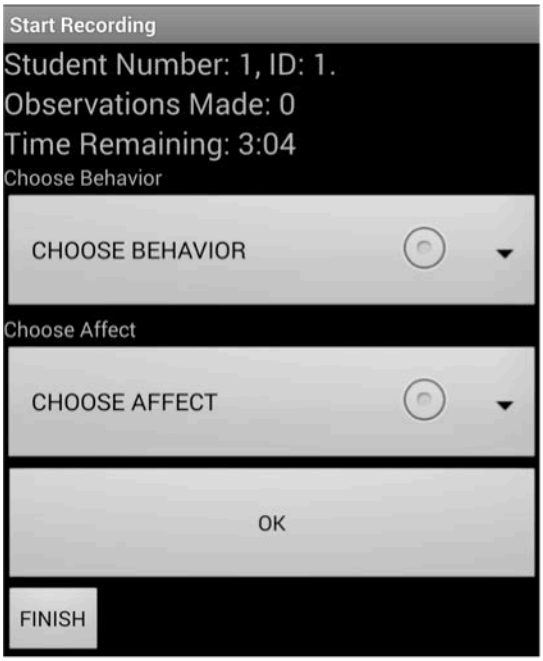


Figure 4. Human Affect Recording Tool (HART) application interface.

HART was used to record student behavior observations in the following categories: "on task," "other on conversation" (other on-task conversation), "off task," "receiving help," "giving help," and "?" (unknown). This coding scheme for behavior categories was developed by Baker at al. (2004). For affect, HART provides the following categories: "bored," "confused," "concentrating" (engaged concentration), "frustrated," "delight," "eureka," and "surprise." See Figures 5 and 6 for the application interfaces; see Tables 2 and 3 for complete descriptions of each of these categories

(Ocumpaugh et al., 2012). Graesser and colleagues developed this coding scheme for affect (D'Mello, Taylor et al., 2007).

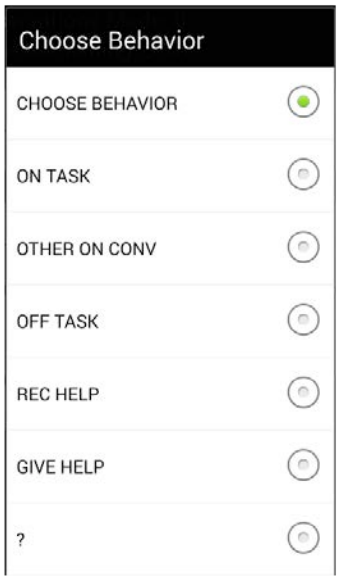


Figure 5. Human Affect Recording Tool (HART) interface for coding behavior.

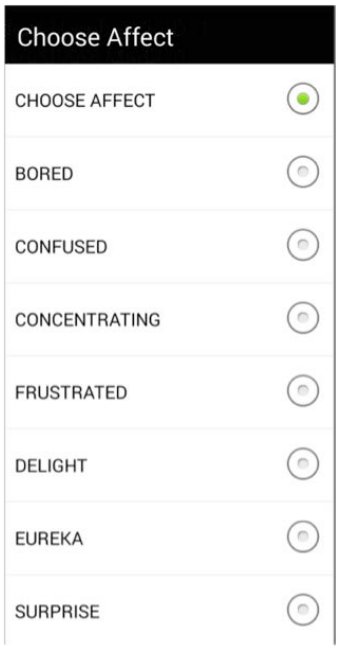


Figure 6. Human Affect Recording Tool (HART) interface for coding affect.

Table 2

Coding Scheme for Students' Behavior

Behavior category	Category description
On task (OT)	Student is working on the assigned task.
Other on conversation (OOC; otherwise operationalized as on-task conversation)	Typically refers to a student who is working on the task while casually chatting with a peer on topics related to the task (e.g., "What level are you on?") or to the teacher ("Have you played this game before?").
Off task (OfT)	Student is not engaged with the assigned task but rather is occupied with unrelated tasks.
Giving help (GH)	Student is providing task-related help to another student who is working on the assigned task.
Receiving help (RH)	Student is receiving task related help from another student.
? (Unknown)	Typically means a student who could not be coded because of physical absence or student noticed field observations taking place.

Note. Source: *Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual Version 1.0,* by J. Ocumpaugh, R. Baker, and M. Rodrigo, 2012, New York, NY: EdLab.

Table 3

Coding Scheme for Students' Affect

Affect category	Category description
Bored (B)	Identifies with behaviors such as slouching, resting the chin on own palms, statements such as "Can we do something else?" or "This is boring!"
Confused (CF)	Noticeable lack of understanding. May include student behavior such as scratching her own head, consulting with a peer student or seeking help from the teacher, peeking over to another student's screen to find solutions or see what the peer did, and statements like, "I'm confused!" or "Why didn't it work?"
Concentrating (C)	Manifested by visible "immersion, focus, and concentration on the system, with the appearance of positive engagement: leaning towards the computer; mouthing solutions; pointing to parts of screen" (Baker et al., 2004).
Frustrated (F)	Occurs when the student faces an unresolvable "impasse" or the student has no action plan on how to overcome the barrier. This includes banging on the keyboard or throwing the mouse, pulling own hair, deep sighing, as well as statements such as, "What's going on?" "This is so frustrating!" or "There is no way to solve it!"
Delight (D)	Expression of pleasure with the results at the moment: may include behavior such as clapping of hands and laughing with pleasure as well as statements such as, "Yes!" "I got it!" or "Yay, it worked!"
Surprise (S)	Identified as "sudden jerking or gasping" and statements such as "Huh?" or "Really!"
Eureka (E)	"Ah hah!" moments when students acquire new profound insights
? (Unknown)	Typically means a student who could not be coded because of physical absence or he noticed field observations taking place

Note. Source: *Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual Version 1.0,* by J. Ocumpaugh, R. Baker, and M. Rodrigo, 2012, New York, NY: EdLab.

Affect and Behavior Data

During affect data collection, I used all of the categories (see Tables 2 and 3); however, in order to answer the research questions, analyses mainly concentrated on patterns of frustration, confusion, boredom, off-task behavior, and giving- or receiving-help behavior. The affect recordings were saved with UNIX timestamps making it possible to synchronize with the gameplay data. In most analyses I analyzed *affect-behavior states*, in other words, combinations at one observation point of students' affect and behavior. For example, a student might be confused on task (CF-OT) early on in the observation session and later be bored off task (B-OfT).

Demographic Data

Students' gender and age information was obtained from the school's administration. Due to limited access, no other demographic information was made available. In addition to the information obtained from the school's administration, each student was given a BrainPlay (2013b) account to participate in this study. BrainPlay (2013a) is an online website that facilitates implementations of large-scale studies with science learning games and contains some student demographics, such as students' gender. BrainPlay's infrastructure facilitates continuous gameplay data collection from Quantum Spectre. Consent forms sent out to students' parents along with the Institutional Review Board protocol included information on such data collection.

Gameplay Data

The game selected for this research, Quantum Spectre, did not provide the user with any hints or instructions on how to play the game. In addition, it did not have any

support systems such as an interface for questions or recommendations for the next move. The learning environment captured gameplay data at a low level of granularity, including clickstream, sequences of actions, timestamps, movements on the board, among others (see Appendices B and C). The gameplay data contained necessary UNIX based time information to combine students' gameplay with their affect records.

The student study data were stored in the BrainPlay accounts of each student, which made data identifiable by the username (not student name) and synchronizable to the rest of the data collected for each student. Every time students played the game through their accounts, gameplay data were captured in BrainPlay, thus assuring the continuity of the collected data. Each student received an account with a predefined user ID in order to provide confidentiality. Along with a user ID for the game, I assigned students to the same computer for the entire period of the class (this was to maintain the same order for affect data collection via the BROMP tool). Each student's unique gameplay user ID was identical to his or her user ID in the affect data collection tool (see Affect and Behavior Data section).

The Quantum Spectre game progression prepared for this study had multiple levels of difficulty that increased as students progressed through the game. The raw log file contained information on beginning events for a level such as a "Level Start," "Rotate," "Move," "End Level," and so on. This included actors in fixed grid locations (i.e., placed on the board) as well as those in the inventory (i.e., in the left menu bar). A detailed list of log data, events, and features is provided in Appendices C, D, and E along with information on the data format.

Procedures

All the students played through the game progression individually. Although the students were allowed to ask questions and talk with each other during the entire time of their gameplay, they were not allowed to control another student's keyboard or mouse. This stipulation ensured that only one user did each move recorded in the gameplay log data. Other than this restriction, students were free to play their game progression as they pleased and ask questions or provide help whenever necessary.

Each session started with a 1-minute preparation period for everyone to be in their seats in front of their individual computers with the game interface on their screens. To make sure that the game was played uniquely in the classroom, the browser page was preloaded and students were not given access to the EdGE game interface page to play at home or after the classroom sessions. This ensured that no student had an advantage of familiarity with the game over other students. One teacher and one student teacher were available as a resource when students had questions.

I was the only observer recording observation data in a predefined order (see Table 4) using the BROMP. I observed students in sets of three for as many sets as the classroom period allowed. Observing a set of three students for one observation each is called a "round" (round is about the observations, while set is about the student group). I managed to record approximately four to five sets of three students during each 40- to 45-minute class period. I gave preference to a design of observing sets of three students over a previously employed design of observing each student in the classroom, in order to collect dense data points of affect and behavior data.

Table 4

Details of Coding Procedure

Time (minutes: seconds)	Student observation
Student Set 1	
0:00–0:15	Student 1
0:17–0:32	Student 2
0:34–0:49	Student 3
0:51–7:37	Repeat observations of Student Set 1 for nine rounds.
Student Set 2	
8:00–8:15	Student 4
8:17–8:32	Student 5
8:34–8:49	Student 6
8:51–15:37 approx.	Repeat observations of Student Set 2 for nine rounds.
Student Set 3	
16:00–16:15	Student 7
16:17–16:32	Student 8
16:34–16:49	Student 9
16:51–23:37 approx.	Repeat observations of Student Set 3 for nine rounds.
Student Set 4	
24:00-25:15	Student 10
25:17-25:32	Student 11
25:34-25:49	Student 12
25:51-31:37 approx.	Repeat observations of Student Set 4 for nine rounds.
Student Set 5	
32:00-43:15	Student 13
34:17-34:32	Student 14
34:34-34:49	Student 15
34:51-39-37 approx.	Repeat observations of Student Set 5 for nine rounds.

I stood near the first set of three students and observed the affect and behavior of the first student in that set for 15 seconds and then recorded this student's predominant affect and behavior states. According to the BROMP, there are occasions when the observer notices several emotional states within the same 15-second segment (Ocumpaugh et al., 2012). The protocol suggests using one's best judgment for identifying the emotional state that is dominant for that segment or, in case of difficulty, using the "?" category in order to avoid misidentification. This recording process took approximately 2 seconds. Then I moved on to the second student of the same set and recorded his or her predominant affect and behavior after observing for 15 seconds. Finally, I observed the third student in the set and recorded this student's predominant affect and behavior in the same manner. This process of observing three students in a set is called a "round." A round took about 49–51 seconds to complete. I observed the same set of three students for a total of nine rounds before moving to the next set of three students. I observed this first set of students for approximately 8 minutes in order to complete nine rounds of observations before moving on to the next set.

Sometimes the class period ended before I had a chance to complete all nine rounds for the set. However, I made sure to end each day's observation on the last student in the set. Because of a 24-hour gap between the last observation of the day and the first observation of the following session, I did not restart from the last set I observed the previous session. This means that some of my sets did not get a full nine rounds of observations. I had to account for this in my analyses.

On the following data collection day, I started with a new set and continued the same data collection procedure. In order to make sure that all the participants received equal chance of exposure to the data collection, my data collection design varied the order of sets that I was to observe each day. However, this variation was not at random every time: the relevant literature did not specify need for such random selection for every new set). For example, if I started from the front of the room on the first day, the next day I might start from the back of the classroom. Observation order and timing chart were defined a priori (see Appendix E).

It is important to highlight that I created an organized selection of student sets instead of randomizing it at the start of every session. I believe, this approach made the field observations less obtrusive and visible to students.

Data Analysis Methods

I conducted several different analyses for the purposes of this dissertation. For these analyses, the collected data were first screened for invalid or missing data. Since there were no missing data due to using "?" (Unknown) coding, I proceeded with the next step. Next, I analyzed the affect and behavior frequency data descriptively by classroom for normality, independence, and variance. Given that my data were not distributed normally (see Appendix G-J), I sought out a nonparametric alternative for analysis of variance. For this purposes I used the Kruskal-Wallis test. The Kruskal-Wallis test is a nonparametric alternative to analysis of variance. Therefore, instead of running a one-way analysis of variance (O'Brien, 1979), I selected the Kruskal-Wallis nonparametric test (Hollander & Wolfe, 1973) to examine the null hypothesis that both classrooms were

similar, or alternative hypotheses (e.g., they had different concentrations or that there was more concentration in Classroom 1 or Classroom 2). This analysis was conducted for exploratory reason: there was no a priori hypothesis whether the classrooms will have the same or variable distribution of affect or behavior. Therefore, I did not pursue one tailed t-test. The general purpose of the variance test was to identify whether there is any variance between the classroom affect and behavior distributions.

I used the R programming language to count the frequencies of each affective and behavioral state for each student throughout the entire 9 sessions. Seven of the eight affect states and five of the six behavior states were used in this analysis. The Kruskal-Wallis test reports a chi square (χ^2). I ran the Kruskal-Wallis test to ensure that the samples were similar before combining the data. The test showed that none of the variables was statistically significant, thus supporting the null hypothesis and suggesting that the samples were similar and I could combine them for my further analysis. However, given the "fishing for significance" approach of this analysis (Haines, 1981), where 12 significance tests were conducted, the Bonferroni (1936) correction method was suggested to correct for multiple comparisons. After conducting this test, none of the 12 variables showed statistical significance. Hence, there was no significant variance in the distribution of affect and behavior between the two classrooms. While Bonferroni correction is considered to be a more conservative test, none of the variables were significantly different; hence, the use of Bonferroni correction test was appropriate.

For further analyses, I coded the data depending on the need for student-based or sequence-based analyses. For all analyses, I used the data from all 51 students whose

parents consented to this study. Because there was a minimum of 24 hours gap between each new set of observations, I conducted some analyses on the level of student per day, which resulted in 272 sequences (9days * approximately 10sets of students per day * 3 students in each set). Other analyses I conducted at the level of student, which resulted in 51 sequences (one sequence per student). Table 5 shows three hypothetical examples of affect-behavior sequences.

Table 5

Examples of Student Affect-Behavior Sequences

Observation	Sequence 1	Sequence 2	Sequence 3
Observation 1	CF-OT	CF-OfT	C-RH
Observation 2	CF-OT	CF-OT	CF-OT
Observation 3	F-OT	F-OT	CF-OT
Observation 4	D-RH	B-OfT	D-RH
Observation 5	E-RH	E-RH	E-OT
Observation 6	C-OT	C-OT	C-OT
Observation 7	D-OT	D-OT	C-GH
Observation 8	Unk-A-Unk-B	C-OT	C-OT
Observation 9	Unk-A-Unk_B	C-OT	-

Note. B = bored; CF = confused; C = concentrating; F = frustrated; D = delight; E = eureka; OT = on task; OfT = off task; GH = giving help; RH = receiving help; Unk-A = unknown affect; Unk-B = unknown behavior.

To understand the basic characteristics of my data, I used traditional statistical methods. To describe the answers to my research questions, I used frequent sequence pattern mining, Markov models, and sequential pattern clustering with optimal matching. Below I describe each of these methods in detail.

Educational Data Mining

The uniqueness of scale of educational game datasets renders many traditional statistical methods inapplicable for analyses such as the ones I needed to conduct for understanding the underlying patterns of affect and behavior (Azarnoush, Bekki, Runger, Bernstein, & Atkinson, 2013). Researchers considering educational game datasets have used sequence analyses in order to gain a more granular perspective of the data and existing patterns therein (Pahl & Donnellan, 2003; Sanjeev & Zytkow, 1995; Shen, Yang, & Han, 2003; F.-H. Wang, 2002; W. Wang, Weng, Su, & Tseng, 2004; Zaïane & Luo, 2001; Zaïane, Xin, & Han, 1998). Sequence pattern analyses are concerned with the underlying patterns and orders of events in the dataset (Agrawal & Srikant, 1995; Zhou, Xu, Nesbit, & Winne, 2010). Once student data are converted into a simple ordered list of items (see Appendix F), there are numerous ways to investigate this sequential data.

Frequently Observed Sequence Patterns

The main goal of this exploratory study was to contribute to the field's knowledge of underlying sequential patterns of student affective states such as frustration and confusion and their interaction with student engagement. Hence, for this study it was important to discover whether the data contained patterns of affect and behavior that were temporally ordered, particularly those that contained confusion, frustration, or boredom. To discover patterns, I used a frequent sequencing algorithm called the Sequential Pattern Discovery Using Equivalence (SPADE) classes (Zaki, 2001) to delve deeper into underlying patterns observed after my initial graphical analysis of student affect and behavior sequences.

Frequent sequencing is explained as an association discovery over a temporal database (Agrawal, Mannila, Srikant, Toivonen, & Verkamo, 1996; Savasere, Omiecinski, & Navathe, 1995). This analysis allows the user to discover inter-event patterns (i.e., sequences) within many different input-sequences. For this analysis I used the arulesSequences package (Buchta, Hahsler, & Diaz, 2014) in the R programming environment to implement the SPADE algorithm to identify frequent sequential patterns. SPADE mines associations via temporal joins (database joins) and lattice algorithm (Agrawal et al., 1996; Savasere et al., 1995; Zaki, 2001). Research has suggested that the SPADE algorithm is more efficient in terms of execution time and reduces the number of database scans needed (Zaki, 2001) compared to previous frequent sequencing algorithms such as AprioriAll (Agrawal & Srikant, 1995) and GSP (Srikant & Agrawal, 1996).

For the implementation of this algorithm, two parameters need to be considered: support and maximum gap. Support is a value between 0 and 1 that indicates the frequency of the state. The support value is chosen a priori based on the research questions. As suggested in the literature, the support value depends on the research question, since if the support value threshold is set too high, it will inhibit finding the rules involving rare states in the data (Kumar, Srinivas, & Rao, 2012). Research on students' affective states in digital learning environments (D'Mello, 2013) has suggested that concentration is significantly more frequent than confusion and boredom, which are in turn significantly more frequent than frustration. Hence, I expected to observe significantly more concentration than any of these three affective states of interest. Therefore, for this analysis, I set support at 0 to unveil all the sequences, particularly the

sequences that contain frustration, a relatively infrequent affective state. As described previously, the maximum gap is the maximum order difference between any two elements. Hence, I set the maximum gap to equate to 1, which would make sure that the algorithm did not skip point B while going from point A to point C but rather would make sure to keep the sequential order and consider all the consecutive states without skipping.

To prepare the dataset, I took all the sequences for all students for all the study sessions and created a basket format: transaction data is put into temporal/sequential format with order information for each of the sates. This created more than 2,000 lines of single state affect-behavior data organized in temporal order for each student and each session. Then I input this data into SPADE algorithm, assigned maximum gap and support information discussed earlier and allowed the algorithm to find subsequences from initial sequences (without altering the temporal information) and provide frequency information for each of these new subsequences. This means that the algorithm searches the data for subsequence patterns and identifies how many of the initial sequences contain these new subsequences. This results in more than 5,000 subsequences of different length (maximum of 9 observations).

Once all the new sequences were identified, to answer my research questions for this analysis, I selected all the sequences that contained at least one instance of confusion, frustration, or boredom (e.g., CF-OT, CF-OfT, F-OT, F-OfT, B-OfT, etc.). This was done via a programming script that automatically selected all those sequences that contained at least one state with above mentioned affect categories. After creating this new sequence

data, I conducted inter-sequence distance analysis and hierarchical clustering described below.

Inter-sequence Distance Analysis

Inter-sequence distance analysis (Sabherwal & Robey, 1993) with optimal matching (Bailey, 1994; Tryon, 1939) and hierarchical clustering techniques were conducted with several different datasets in this dissertation.

Inter-sequence distance analysis can be conducted using different algorithms: Hamming distance (Hamming, 1950) or optimal matching (using Needleman-Wunsch; Abbott & Tsay, 2000) algorithms. Given the optimal matching algorithm's adaptiveness to matching similarities in different parts of a sequence, I have employed an inter-sequence distance method using the Needleman-Wunsch algorithm for optimal matching (Needleman & Wunsch, 1970). This algorithm creates a distance matrix consisting of the cost (value) of transforming and assimilating one sequence to another. This transformation process can be done using insertion, deletion or substitution approaches (Sabherwal & Robey, 1993). The algorithm automatically calculates and selects the approach that is the most economical for the transformation process (i.e., transformation approach that has the lowest cost).

For this analysis, the TraMineR (Gabadinho, Ritschard, Müller, & Studer, 2011; TraMineR, 2014) sequence analysis R (R Development Core Team, 2013) package was used to generate, describe and visualize sequences of student affect and behavior categories. In order to understand how TraMineR computes sequence dissimilarity, some of the technical terms require further interpretation: *sequence distances*, *optimal*

matching, and *transformation cost*. Sequence distances are computed in a pairwise manner. Each sequence is matched with all the other sequences through using one of three actions: insertion, deletion, or substitution of each and every dissimilar state between the pair of sequences. This transformation process generates the dissimilarity matrix that contains numeric values for transformation cost that is computed based on how many actions will be taken in order to assimilate pairs of sequences. The dissimilarity can be computed using several different metrics such as optimal matching, distance based on the longest common prefix or on the longest common subsequence, and hamming distance. For my purposes optimal matching was the best solution (explained earlier); it is also a method used in social sciences for time-ordered sequence data (Abbott & Tsay, 2000; Wu, 2000). The optimal matching function can use either a constant or a transition rate for transformation costs; however, the literature does not provide a distinctive guideline as to what is the appropriate cost (Wu, 2000). Therefore, for this analysis, I arbitrarily assigned each of the transformation approaches an equal cost (see Chapter 5 Limitations section for more details on transformation cost).

To further explain the process, below I present an example of such process. In Table 6 I display three sequences with different length: sequence 1 has nine observations with last two observations being Unknown. Sequence 2 has nine observations and sequence 3 has only eight observations.

Table 6

Example of Sequence for Later Transformation

Observation	Sequence 1	Sequence 2	Sequence 3
Observation 1	CF-OT	CF-OT	C-OfT
Observation 2	CF-OT	CF-OT	CF-OT
Observation 3	F-OT	F-OT	CF-OT
Observation 4	D-RH	D-RH	D-RH
Observation 5	E-RH	E-RH	E-OT
Observation 6	C-OT	C-OT	C-OT
Observation 7	D-OT	D-OT	C-GH
Observation 8	Unk-A-Unk-B	C-OT	C-OT
Observation 9	Unk-A-Unk_B	C-OT	-

Note. B = bored; CF = confused; C = concentrating; F = frustrated; D = delight; E = eureka; OT = on task; OfT = off task; GH = giving help; RH = receiving help; Unk-A = unknown affect; Unk-B = unknown behavior.

Sequence 1 and Sequence 2 are almost identical except for observations 8 and 9. To be considered different, either the affect or the behavior portion of the observation should be different (e.g., CF-OT vs. CF-OfT) or both should be different (e.g., D-RH vs. B-OfT). Once the differences between the sequences are identified, the algorithm has several ways to transform sequence 1 into sequence 2, sequence 1 into sequence 3 and sequence 2 to sequence 3. One of the options is through substitution, substituting the eighth and the ninth observations for sequence 1 with C-OT to conform to sequence 2. Assuming that each substitution cost is 1, this approach results in a transformation cost of 2.

In case of assimilating sequence 1 and sequence 3, the algorithm could use substitution to alter observations 1, 3, 5, and 7 and use deletion for observation 8 and 9. Another option is to substitute the eighth observation for sequence 3 with an Unk-A-Unk-B and insert the ninth observation as Unk-A-Unk-B. The final decision is based on the cost of this transformation and whether we would like to shorten our sequences.

After calculating the cost for each sequence compared to every other sequence, a pairwise distance matrix is generated. Table 7 displays this example transformation matrix. In Table 7, I have presented a matrix composed of three sequences from Table 6 that show the similarity cost between each pair (e.g., the similarity between sequences 1 and 2 costs only 2, and the similarity between sequences 1 and 3 costs 6, etc.).

Table 7

Transformation Matrix of Example Sequence (See Table 6)

	Sequence 1	Sequence 2	Sequence 3
Sequence 1	0	2	6
Sequence 2	2	0	5
Sequence 3	6	5	0

After calculating the dissimilarity matrix, I employed a hierarchical clustering approach for categorizing my sequences into meaningful groups. The clustering analysis was conducted on the numerical data from the transformation process in order to identify groupings of sequences.

Clustering

Cluster analysis is a common method for classifying data into categories (Lorr, 1983). It is based on the Levenshtein distance approach (Levenshtein, 1966). Given the nature of my data, I selected hierarchical clustering method over other clustering methods (Rencher, 2002; Romesburg, 1984). This clustering method is guided by a bottom-up approach and arranges solutions in a hierarchical structure: It starts out with small clusters that contain individual states and progressively expands into larger clusters with items that are closely related. Most importantly, this method does not require an a priori decision on the number of clusters, since the number of clusters was not hypothesized in this exploratory study.

In this analysis, I used sequence clustering (Sabherwal & Robey, 1993) using optimal matching (Bailey, 1994; Tryon, 1939) and hierarchical clustering techniques (Lorr, 1983) with several different data files (e.g., all the subsequences generated as a result of the frequent sequencing analysis; all the 272 unique sequences; and finally, all the student cluster sequences). For this analysis, I used TraMiner sequence analysis R package. This process is the continuation of the inter-sequence distance analysis, where the dissimilarity matrix numbers are clustered in order to group sequences.

Since I conducted my cluster analysis on a distance matrix, hierarchical clustering methods was rendered most appropriate for my analysis (Ulrich & McKelvey, 1990). Based on TraMineR documentation, I used Ward's clustering method approach suggested by the authors (Ward, 1963). This method uses sum of squares for distance measure. While generally in hierarchical clustering, distance of data points identifies the

new groupings, Ward's method relies on the sum of squares. Sum of squares starts out at 0 with every data point being in its own cluster. According to previous studies, Ward's method is considered to have higher overall performance compared to other hierarchical methods (Blashfield 1976; Desmarais & Lemieux, 2013; Hands & Everitt, 1987; Kuiper & Fisher, 1975).

Although hierarchical clustering avoids the issue of predetermining the optimal number of clusters a priori, it requires the researcher to determine this number postanalysis based on a dendrogram or on agglomeration coefficients. For hierarchical clustering, the literature (Everitt, 1980; Kiran, Serra, & Cousty, 2012) suggested to stop clustering further down either when the distance between the objects within each cluster becomes too small as to suggest a large similarity between objects, or when there is a large difference between the numbers of instances in the clusters. Therefore, for each of the cluster analysis in this dissertation, I have selected a range of three to eight clusters before identifying the best cluster based on the cluster memberships, dendrogram and the agglomeration coefficient.

While each of the previous analyses investigated the affect-behavior sequences in isolation of student information, my study was investigating affect patterns of a specific sample, fifth-grade students at a local elementary school. Therefore, another clustering analysis was conducted on student level: after clustering all the sequences, this information was combined with individual student IDs by creating a sequence of cluster memberships for each student. An example sequence could be {1,3,5,2}. This sequence notation can be interpreted as follows: A student has affect sequences within identified

Cluster 1 (the confusion-frustration cluster), then Cluster 3 (the persistent cluster), Cluster 5 (the requiring-help cluster), and finally Cluster 2 (the concentration transition cluster). Some students were observed more than others; therefore, they had more cluster memberships than others. After compiling these data, the same sequence cluster analysis was conducted to cluster students based on their sequence information.

Markov Models

I conducted additional analyses of sequences using Markov sequence model. This model assumes that the probability of each item within a sequence depends on the item chosen for the previous position. For the purposes of this dissertation the analysis did not go into selecting either the Markov chain or the hidden Markov model for states that are either fully or partially observable, since prediction was not a goal of this study. Instead, the analysis used the Markov property to analyze the sequence states. This study used a transition matrix built upon the idea of the Markov property (Markov, 1954) to identify interesting association patterns. The notion of a Markov property for sequences assumes that the next observation in the sequence depends on the current state (or the current states depended on the previous state) and not on the entire previous sequence. This allowed me to find associations between combinations of affect and behavior, which is the first step for building a model of possible order of affective states based on observed patterns. The basis of the Markov model is the Markov process (Dynkin, as cited in "Markov Process," 2014). In the theory of conditional probability, Markov process is a stochastic or random process that satisfies a specific property (Gihman & Skorohod, 1975). The Markov property is otherwise called "memoryless" since only the present

(current) state matters for the prediction of the future state; the past is independent from the future ("Markov Process," 2014). According to the assumptions of the Markov process, a guess about a future state (e.g., frustration) does not depend on knowledge of all prior states but depends only on the current state (e.g., confusion).

To prepare the data for the Markov process, I generated a Markov stochastic matrix (also called a probability or transition matrix) to find the probabilities of the next affective and behavioral states based on the previous state (e.g., the probability of student disengagement based on previous affective and behavioral states). A Markov matrix requires a selection of parameters or states to calculate the transition probability. In this study analysis states could take on a discrete number of values (8*6 = 48); however, there were only 28 unique combinations of affect and behavior states that were actually observed in the data (see Figure 7).

Summary

This dissertation uses educational data mining techniques and traditional statistical methods to examine student engagement levels and their emotional response during gameplay in a digital environment for science learning. In this study I investigated significant patterns in the interaction of frustration and confusion and the relationship between occurrences of these emotions and student engagement. In the next chapter, I provide the findings from my analyses methods.

CHAPTER 4

RESULTS

In this chapter I provide the results organized by research question. In this study, an alpha level of .05 was used to test statistical significance.

Classroom Variability

The null hypothesis is that both classrooms would have a similar distribution of the frequencies of affect and behavior data (no significant difference between the distributions of types of affect and behavior between the two samples). To determine the appropriate statistical test for this hypothesis, I began by testing for normality of the distribution of all affect and behavior states in both classrooms. Histograms of affect and behavior data indicated that there was a mixed distribution of the data: normally distributed, skewed, and bimodal (see Appendices G, H, I, and J). In addition, Table 8 summarizes all the affect and behavior categories with their frequencies (I did not include Unknown states).

Table 8

Affect and Behavior Frequencies

Affect or behavior	Classroom A	Classroom B	Total
Affect			
Bored	31	71	102
Confused	203	235	438
Concentrating	806	784	1,590
Frustrated	100	109	209
Delight	58	38	96
Surprise	0	2	2
Eureka	14	9	23
Unknown affect	10	10	20
Behavior			
On task	906	887	1,793
Other on conversation	61	76	137
Off task	47	80	127
Giving help	73	69	142
Receiving help	126	136	262
Unknown behavior	9	10	19

Note. The ordering of affect and behavior categories in this Table is based on the ordering of affect and behavior categories in the Tables 2 and 3.

Analysis 1: Most Frequent Affect-Behavior Patterns (Research Question 1)

The first analysis was to answer Research Question 1: What are some of the most frequent affect-behavior patterns? I hypothesized that there would be an interaction between certain affective states and student behavior (e.g., confusion, frustration, boredom, off-task behavior, etc.). To test this hypothesis, I used graphical tools such as the R package TraMineR (Gabadinho, Ritschard, Müller, & Studer, 2011) to visualize

student affect and behavior sequences and to provide information on frequencies of each state (both affect and behavior indices).

Further, to conduct exploratory analysis, I created visualizations of all of the sequences, especially frequently occurring sequences in the data. For this purpose, I used the graphical tools from the R package TraMineR (2014) package. In Figure 7 I present the label for all the visualizations in this chapter. Due to the extensive number of variables, I presented this legend separately from the actual graphs (this legend is the basis for all of the visualizations in this dissertation, unless otherwise specified). I used sequential plots to graph (see Figure 8) all the 272 sequences.

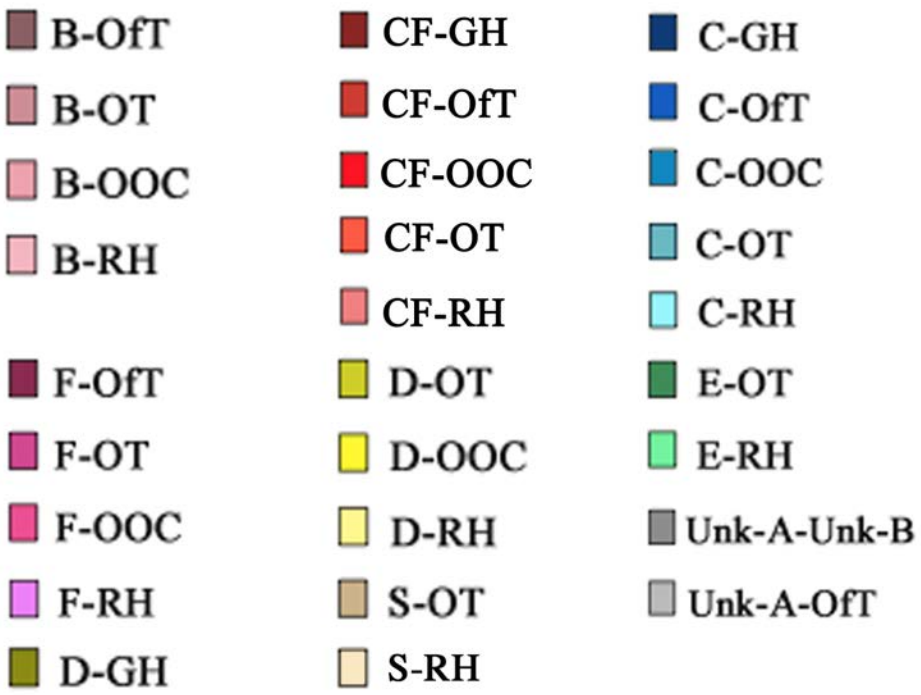


Figure 7. Categorized color identifiers for all the 28 unique affect-behavior states. B = bored; CF = confused; C = concentrating; F = frustrated; D = delight; S = surprise; E = eureka; Unk-A = unknown affect; OT = on task; OOC = on-task conversation; OfT = off task; GH = giving help; RH = receiving help; Unk-B = unknown behavior. These states are organized into groups based on affect and behavior order identified in Tables 2 and 3. Within groups they are organized by the color scheme.

In Figure 8 I show all 272 sequences ordered by the most frequent affect-behavior states in the beginning (left side) portion of all observations. This graph displays C-OT to again be the most frequent state in the first several observations. The second most frequent affect-behavior state is CF-OT in the first few observations. Next show up F-OT and C-RH states, which appear with roughly equal frequency. This graph shows quite diverse sequences in terms of affect-behavior states.

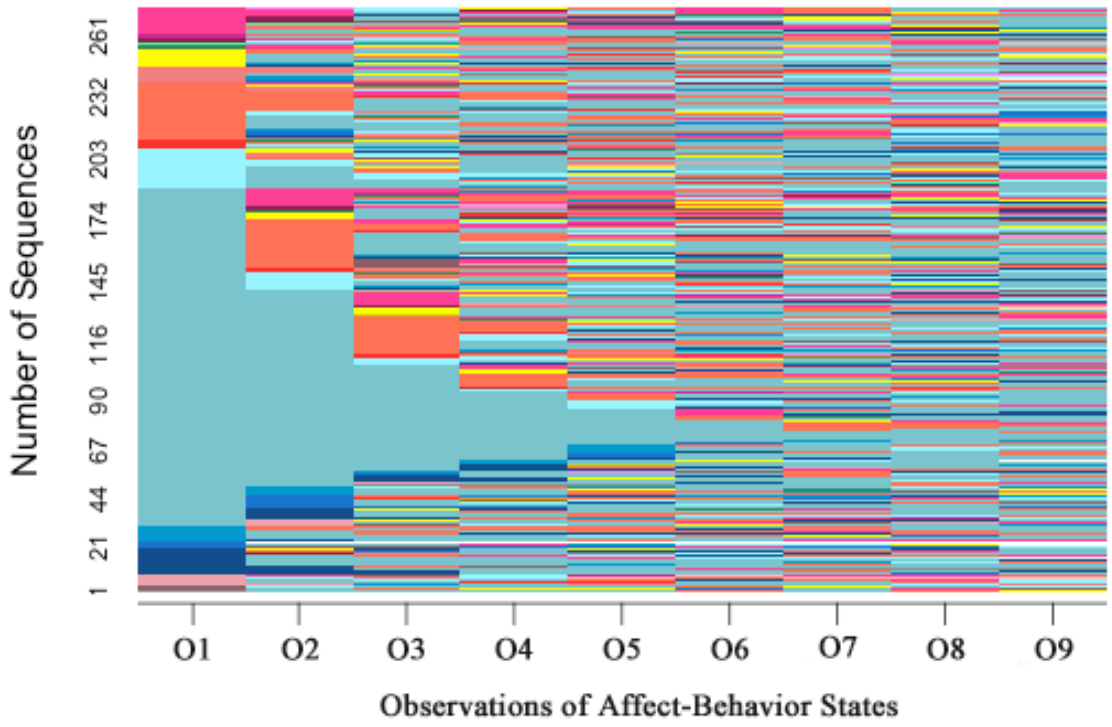


Figure 8. Visualization of all 272 affect-behavior sequences (sorted from the start). See Figure 7 for color schemes.

Finally, I used the R package TraMineR to plot several summary sequence graphs that shed light on some of the sequence characteristics present in the data. For example, Figure 9 displays the percent of states at each observation. Hence, in Observation 1 it is

visible that C-OT was the most frequently occurring state. The second most frequently occurring state in Observation 1 was CF-OT. While this type of visualizations may easily mislead the reader, I used this visualization to simply combine the line states per each observation segment.

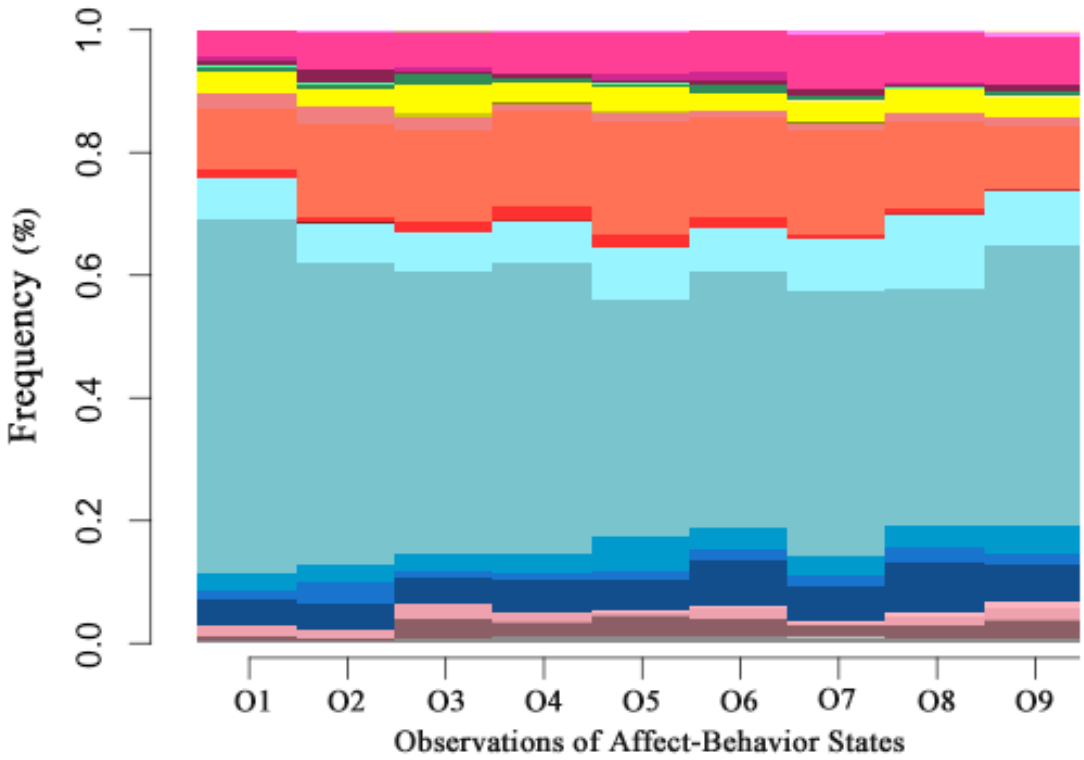


Figure 9. Sequence state frequencies. See Figure 7 for color schemes.

In Figure 10, I display the entropy index (based on 272 sequences). This index indicates the level of diversity in the sequence—how many different types of categories there are in the dataset for each observation. The entropy index remains relatively stable throughout the entire sequence. Maximum value for entropy indicates the abundance of diversity in the data. Minimum value for entropy index indicates the lack of diversity. For

example, in this study the entropy index of all sequences was of medium level, which indicates that although not abundant, the diversity in the data is relatively level, which in this context means that students experienced and exhibited a variety of affect and behavior types. The TraMineR package uses a line graph instead of a dot plot since it considers the sequence to be temporal, even though the time variable here means order and not necessarily time elapsed.

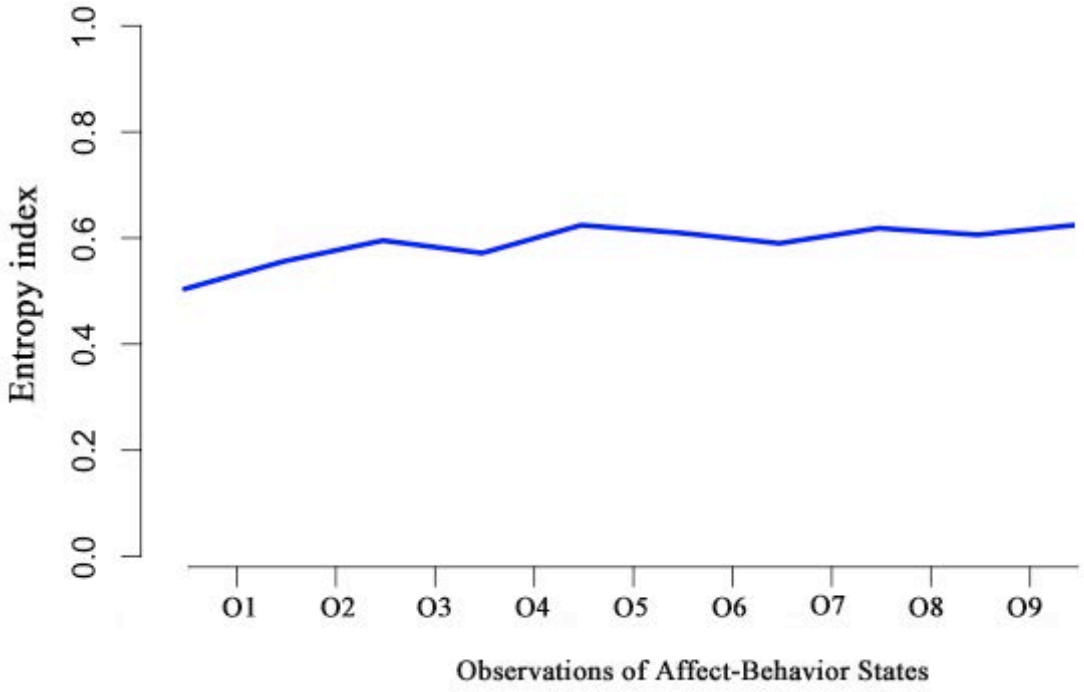


Figure 10. Entropy index of state distributions.

This first set of analysis was to investigate and identify the most frequent individual affect-behavior states present in the data. I found that C-OT was the predominant individual affect-behavior state followed with some instances of C-GH and C-RH. This is an interesting finding that shows that the students were mostly

concentrating on task but also had the need to seek help. In addition, this analysis also identified quite a few B-OT and B-RH states. This was indicative of the fact that students still got bored despite the fun interface and showed need for help. It is possible that there is a need for a more rigorous support system.

Analysis 2: Student Affect-Behavior and Boredom

Research Questions 2 through 5

2. What affect-behavior sequences precede boredom?
3. Are there sequential patterns in the occurrence of frustration and confusion that are consistently associated with boredom?
4. What combinations of affective and behavioral states frequently precede off-task (disengagement) and receiving-help behavior?
5. Are there sequential patterns in the occurrence of affective states, especially frustration and confusion that are consistently associated with off-task (disengagement) and receiving-help behavior?

Cluster Findings

In this analysis, I used the clustering method described in Chapter 3 on all the frequent subsequences that contain CF, F, or B states in order to group these subsequences based on similarity of their content.

Figure 11 presents the tree-like structure of the data (dendrogram) when the hierarchical clustering method is applied. The y-axis shows the height at which clusters split into smaller groupings. It is important to emphasize that this clustering is done on

the transformation matrix numbers that identify the similarity/difference between the sequences.

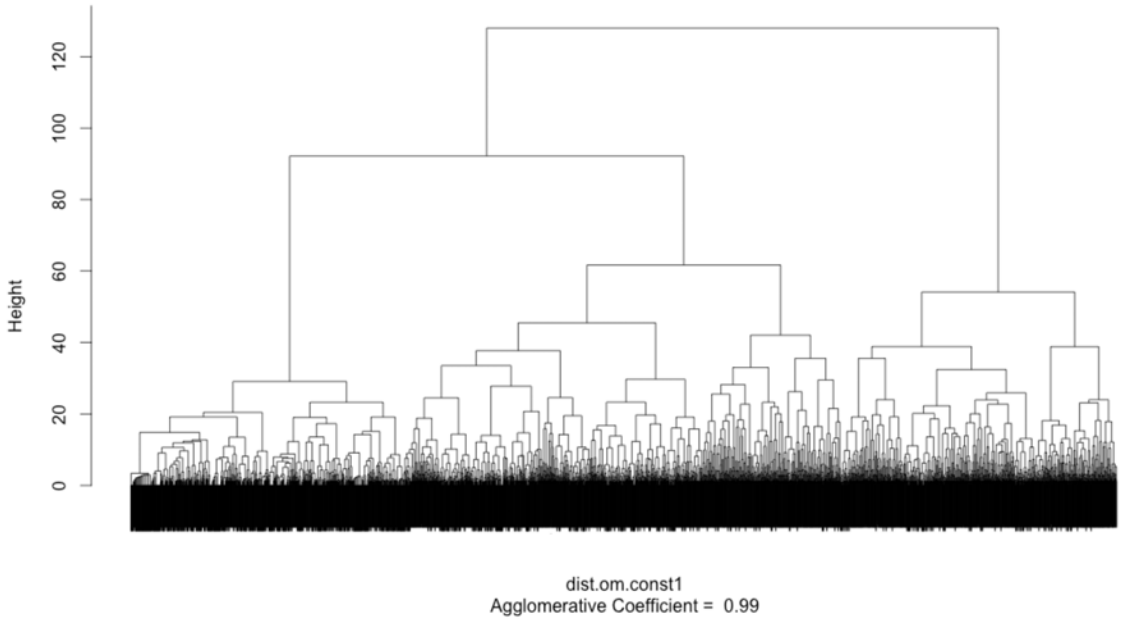


Figure 11. Cluster dendrogram for frustration, confusion, and boredom sequences.

Based on to the cluster membership numbers for solutions with three to seven clusters along with the dendrogram, I chose to examine the interpretability of a four-cluster solution (Everitt, 1980; Kiran et al., 2012; see more in Inter-sequence distance method in Chapter 3). Figure 12 is the graphical representation of the four-cluster solution. After investigating the cluster contents, I named each cluster. I found the following clusters: short sequences, constant transition, persistently confused and frustrated, and consistently concentrating.

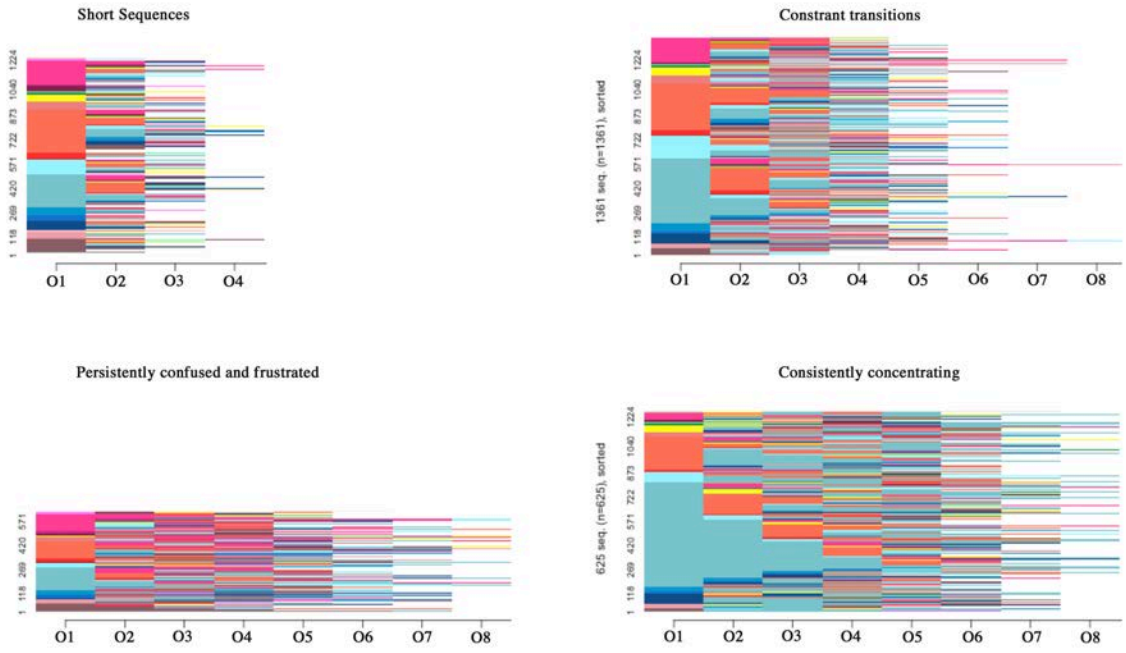


Figure 12. Four-cluster solution for sequences that contain frustration, confusion, or boredom. See Figure 7 for color schemes.

The short sequences cluster. This cluster seems to be grouping based on sequence length. Whereas this could be considered a limitation in any other clustering case, in this context it was quite useful for my research purposes: Short sequences inhibited me from investigating the states before and after the occurrences of frustration, confusion, and boredom. In terms of content, most of these sequences started out with confusion, frustration, or concentration. However, rarely did they show a transition back to concentration. Overall, this cluster showed transitions predominantly into or from confusion to frustration and boredom.

The constant transition cluster. In this cluster, like in the next cluster, constant transitions happen between C-OT and CF-OT states for the sequence observations: A third of the observations started with C-OT but mostly lasted for only one observation

point. In the next state within the sequence, only a few sequences still displayed C-OT; others transitioned between CF-OT, C-OT, F-OT, B-OT, and C-RH. Hence, these sequences alternated equally between C-OT and CF-OT with occasional B-OT and F-OT patterns within the sequences. Therefore, there is no consistent pattern.

The overview of the cluster also showed that most of the time boredom was followed either by another boredom state or frustration but rarely followed by concentration. While not a conclusive result, this may suggest that boredom can be persistent, that is, repeat more than once and hence be distractive to learning. However, the most alarming characteristic of this cluster may be that the B-OfT state seemed to persist once it showed up in the sequence. Hence, this may suggest that once the student is bored and also off task, he or she may run the risk of remaining in that state for several states, which by itself is a loss of learning opportunities.

The persistently confused and frustrated cluster. This group presents a group of sequences that mostly contained persistent affective states such as confusion and frustration that repeated more than one observation period within the sequence. While in the observation 1, a third of the sequences started out with C-OT, they quickly moved into confusion and frustration and remained in these states for the remaining portion of the sequences. The other two thirds of the sequences mainly started out with either CF-OT or F-OT and either repeated those same states in the next observation or alternated between CF-OT and F-OT. These cluster sequences may suggest that although according to the literature confusion may not necessarily be a negative affective state, in this cluster it still showed persistent occurrence and might require attention. In addition, in some of

the cases in this cluster, confusion seemed to be following C-OT states or other confusion states, whereas frustration seemed to be succeeding C-RH. This may be an interesting observation to research in the future studies.

Another interesting characteristic of this cluster is the relatively large number of F-OT states within sequences that succeeded C-OT and sometimes C-RH or F-RH states. However, F-OfT was mostly the result of F-OT state. What were unexpected were the observed transitions from B-RH to F-OT. This suggests that while bored, if students receive help, they may go back into on-task behavior even if still frustrated (this might be a sign of persistence or effort).

The consistently concentrating cluster. This group of sequences primarily started with C-OT (about two thirds of observation 1 sequences were C-OT). However, the interesting characteristic of these sequences is that they reveal a lot of simple confusion within the sequences: Confusion states lasted only one observation point and transformed back to C-OT the next observation period. This is an interesting interplay, because it may show that CF-OT does not necessarily require a remedy to transform back to C-OT.

For this section I was looking at the overall picture of the patterns in the data on an exploratory level. It is critical to highlight that while cluster descriptions offered above presented certain observations of affect and behavior interplay, this analysis was not quantified further due to the fact that these are sub sequences (portions) of the initial 272 sequences that were automatically split into new sequences due to possible repetitive characteristic and do not necessary have high frequency in the actual sequences.

Analysis 3: Student Affect-Behavior and Engagement

Research Questions 2 through 5

This section of analyses covers some of the same research questions answered in the Analysis 2. However, here I look at some of the results in a more granular view.

Markov Model

The previous sequential pattern analyses looked into sequences composed of more than two affect-behavior states. However, to extend this exploratory study, I decided to extend the analysis into looking into sequences of two affect-behavior observations. Therefore, I conducted additional analyses of sequences using the Markov sequence model, which is a relatively accurate representation of the evolution of sequences (see Data Analysis Methods section for more detail).

Findings

To create the transition matrix, I generated a transition table with information that contains all the observed combinations of affect-behavior$_1$ and affect-behavior$_2$ (where 1 is the current state of the affect-behavior pair and 2 is the future state). This transition table contains the frequencies of all affect-behavior combinations (e.g., CF-OT: 209 occurrences), the number of times an affect-behavior state was observed as the initial or current state (e.g., 328 times CF-OT was the initial state in the combination of current and future states), and also the ratio of current state to future state (see Appendix K for a portion of the transition table as an example).

Overall, there were above 2,176 total transitions between affect-behavior states (from affect-behavior$_1$ to affect-behavior$_2$). However, only 210 of them were unique combinations. For example, there were 550 counts of C-OT → C-OT and only 33 counts of C-OT → D-OT.

The transition table was used to create the graph in Figure 13. In Figure 13, all the sequence transitions that had counts above 10 (2.2% of total 2,207 sequences) are displayed. The x-axis represents the frequency of sequence transitions, and the y-axis represents the sequence transition description. In addition, the bars were ordered in descending order of transition frequency and colored according to the initial state (e.g., B-OfT). Each unique instance of affect-behavior present in this subset of the data was assigned a color. Each affect-behavior was grouped within a bigger category based on the affect variable. For example, all the affect-behavior combinations of C-OfT, C-OT, C-OOC, C-GH, and C-RC were grouped in the same color palette, shades of green.

The purpose of this analysis was to see patterns in the transition data. There was no a priori established expectation as to how many of each of the affect and behavior categories will show up in the data. Rather, I tried to visualize affect and behavior pair transitions. Figure 13 shows that C-OT to C-OT was the predominant sequence transition that appeared in the data. The next most frequent transition was from C-OT to CF-OT. A few bars below is the transition from C-OT to F-OT. Figure 13 highlights that the transition from C-OT to F-OT was the sixth most frequent transition (see Figure 13, second blue bar from the top), despite the relative infrequency of the frustration affective state.

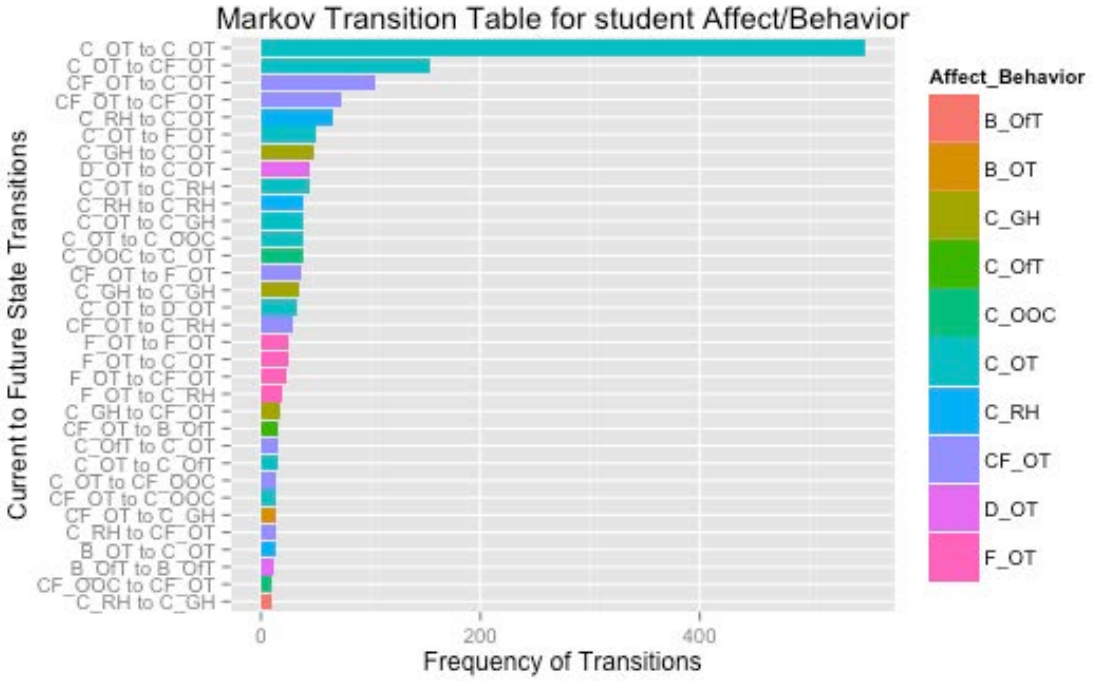


Figure 13. Transitions between affect-behavior states. This graph uses an alternative color schema specifically generated for this visualization. The colors are assigned by the starting state. B = bored; CF = confused; C = concentrating; F = frustrated; OT = on task; OOC = on-task conversation; OfT = off task; GH = giving help; RH = receiving help.

After generating the transition table, the calculation of the transition matrix depended on the number of unique affect-behavior states; investigation of this study's data showed that there were 28 distinct affect-behavior combinations (e.g., C-OT; CF-OT etc.). Based on this information, a transition matrix was generated with 28 variables. In Table 9 I provide the resulting Markov transition matrix. The rows in this transition matrix represent the current affect-behavior states in the sequence (precedent state), and the columns in the matrix represent the potential future state in the sequence. For example, the first row in Table 9 is interpreted as follows: 55.5% of the time the C-OT state was followed by another C-OT state, whereas about 0.3% of the time C-OT was

followed by the CF-OfT state, and 15.7% of the time it was followed by CF-OT state. It is important to highlight that in this matrix the sum of each row values equates to either 100% or 0% (it equates to 0% only if there are no records of that particular affect-behavior combination).

Table 9

Markov Transition Matrix: Percentage of Transitions

	C-OT	CF-OfT	CF-OT	CF-OOC	F-RH	C-GH	C-OOC	C-RH	C-OfT	D-OT	E-OT	F-OT	S-OT	B-OT	B-OfT	CF-RH	F-OOC	F-OfT	B-OOC	D-GH	?-?	E-RH	CF-GH	B-RH	D-OOC	?-OfT	D-RH	S-RH
C-OT	55.7	0.3	15.7	1.3	0.2	4.0	3.9	4.5	1.5	3.4	0.9	5.3	0.1	0.9	0.9	0.8	–	0.3	0.1	–	0.3	–	–	–	0.2	–	–	–
CF-OfT	–	–	50.0	–	–	–	–	–	–	50.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
CF-OT	32	–	22.3	2.1	–	4.0	4.3	8.5	2.1	2.7	1.2	11.1	–	0.9	5.1	1.5	–	0.6	–	–	0.3	0.3	–	0.6	–	–	–	0.3
CF-OOC	28.1	–	31.3	3.1	3.1	–	3.1	18.8	–	–	3.1	3.1	–	–	–	6.3	–	–	–	–	–	–	–	–	–	–	–	–
F-RH	16.7	–	–	–	16.7	16.7	–	16.7	–	–	–		–	–	16.7	–	–	16.7	–	–	–	–	–	–	–	–	–	–
C-GH	40.3	–	14.3	0.8	–	29.4	4.2	–	0.8	4.2	–	1.7	–	–	0.8	0.8	–	–	–	–	–	–	0.8	–	0.8	0.8	–	–
C-OOC	51.4	–	13.5	–	–	6.8	6.8	8.1	2.7	2.7	–	4.1	–	1.4	1.4	–	1.4	–	–	–	–	–	–	–	–	–	–	–
C-RH	38.9	–	7.8	1.2	–	5.4	2.4	23.4	1.2	4.8	0.6	4.7	–	1.2	0.6	5.3	0.6	1.2	–	–	–	–	–	0.6	–	–	–	–
C-OfT	41.0	–	7.7	2.6	–	–	5.1	7.7	15.4	–	–	10.3	–	2.6	7.7	–	–	2.6	–	–	–	–	–	–	–	–	–	–
D-OT	57.9	–	6.6	–	–	14.5	2.6	2.6	2.6	6.6	1.3	2.6	–	–	–	–	–		–	–	1.3	–	–	–	–	–	1.3	1.3
E-OT	41.2	–	5.9	–	–	5.9	–	5.9	–	5.9	–	26.3	–	–	–	–	–		–	–	–	5.9	–	–	–	–	–	–
F-OT	16.9	–	15.5	–	2.0	2.7	2.7	13.5	2.7	4.7	0.7	17.3	–	2.7	2.7	2.0	4.0	5.3	1.3	0.7	0.7	0.7	–	0.7	–	–	–	–
S-OT	–	–	–	–	–	–	–	–	–	100.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B-OT	41.9	–	6.5	–	–	–	3.1	6.5	–	3.1	–	9.4	–	15.6	12.5	–	–	–	–	–	–	–	–	–	–	–	–	–
B-OfT	10.9	–	10.9	2.2	2.2	6.5	–	17.4	2.2	2.1	–	8.5	–	8.5	21.3	4.3	–	–	–	–	–	–	–	–	–	–	2.2	2.2
CF-RH	21.1	–	10.5	–	–	5.3	5.3	21.1	2.6	5.3	–	5.3	–	2.6	2.6	15.8	–	–	–	–	–	–	–	2.6	–	–	–	–
F-OOC	20.0	–	10.0	10.0	–	–	–	20.0	–	–	–	30.0	–	–	–	–	–	–	–	–	–	–	–	10.0	–	–	–	–
F-OfT	25.0	–	6.3	–	–	–	6.3	6.3	6.3	5.9	–	23.5	–	–	5.9	–	–	5.9	–	–	6.3	–	–	–	–	–	–	–
B-OOC	–	–	–	–	–	–	–	–	–	–	–	–	–	–	50.0	–	–	–	–	–	50.0	–	–	–	–	–	–	–
D-GH	–	–	50.0	–	–	–	–	50.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
?-?	23.5	–	–	5.9	–	–	–	11.8	–	–	–	–	–	–	–	–	–	–	–	–	58.8	–	–	–	–	–	–	–
E-RH	75.0	–	25.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
CF-GH	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	100.0	–	–	–	–	–	–	–	–	–	–	–	–
B-RH	–	–	20.0	–	–	–	–	–	–	–	–	40.0	–	–	–	–	–	–	–	–	–	–	–	40.0	–	–	–	–
D-OOC	100.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
?-OfT	100.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
D-RH	–	–	100	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S-RH	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Note. Rows represent the current affect-behavior states, and columns represent the potential next future state in the sequence. A dash represents 0%. B = bored; C = concentrating; CF = confused; D = delight; E = eureka; F = frustrated; S = surprise; GH = giving help; OfT = off task; OOC = on-task conversation; OT = on task; RH = receiving help; ? = unknown.

Below are some of the findings organized according to the affect-behavior states: it is important to highlight that some of these cases may contain small percentages or

small frequencies (given the small study sample), however in the affect research these occurrences represent important value since a consistent affect-behavior transition model has yet to be developed. The findings are organized according to the initial observation: the observations are grouped together (e.g., all the findings for frustration are grouped together):

1. F- RH (16.67%) → F-RH. Continuing the theme in the previous rule, 16.7% of the time students stayed in F-RH. This may suggest that previously provided help was not sufficient. Given the absence of an embedded support system, this may be a problem if it continues (it may lead to further discouraging situations).
2. F-OfT (6.25%) → C-OfT. An interesting pattern showed up in 6.25% of the sequences, when students went from F-OfT back to C-OfT. They not only did not return to their learning environment but became further engaged in an off-task behavior by concentrating on it. This may suggest that after being frustrated, students get disengaged and try to find other things to engage in besides the learning task. This was quite a big portion of sequences and hence should be investigated further in future work.
3. F-OfT (23.5%) → F-OT. A fourth of the time F-OfT behavior was followed by F-OT. While very interesting, it is hard to explain this transition without further investigation of the exact gameplay instance. This may suggest that students were being persistent on their own but also may suggest that students were going from off-task behavior back to on-task for other, unknown reasons.

4. F-OT (17.3%) → F-OT. About 17% of the time, data showed persistent frustration transitions: Students transitioned from F-OT to yet another instance of F-OT.
5. F-RH (16.67%) → F-OfT. Approximately 17% of the time, students transitioned from F-RH to F-OfT. This suggests that 17% of the time after receiving help, students got disengaged. It might have been the result of poor help or might suggest that the students were not paying attention to the provided help because they were extremely frustrated. In either case, students' frustration seems to require remediation.
6. F-OT (5.3%) → F-OfT. For about 5% of the time, students transitioned from F-OT to F-OfT, which may suggest that students were not receiving help to resolve their frustration and were transitioning to disengagement from the learning environment.
7. F-OfT (5.9%) → F-OfT. About 6% of the time, students exhibited persistent F-OfT behavior. The students were frustrated enough to become disengaged and remain disengaged. These are the situations that require immediate remediation.
8. F-RH (16.67%) → B-OfT. Approximately 17% of the time, students went from frustrated and receiving help to bored and off task. Here it seems that despite received support, students still transitioned to boredom and immediately disengaged with the learning task. Seventeen percent is a relatively high proportion, which should require further investigation as to why the provided support is not sufficient to keep the student engaged with the environment.

9. CF- OT (11.1%) → F-OT. About eleven percent of the time students transitioned from CF-OT to F-OT. This may suggest that students were confused with the impasse and were getting stuck, thus transitioning into frustration but still putting enough effort into the learning task as to not get disengaged. It is a question of how long they will remain on task; hence, it is important to consider providing them assistance at certain points.
10. C- OfT (15.38%) → C-OfT. It seemed that 15.38% of the time disengaged but concentrated students remained disengaged and concentrated on off-task behavior. This requires more in-game investigation to see why this disengagement is occurring. However, this investigation is not possible in the boundaries of this analysis.
11. C-RH (4.7%) → F-OT. Although this transition took place only about 5% of the time, it is important to note that concentration while receiving help may transition to F-OT. This may be the result of poor help that the student received or a result of other factors. However, it seems that the quality of received help is crucial here to make sure students do not transition from concentration to frustration immediately.
12. B-RH (40%) → F-OT. Forty percent of the time, the bored and receiving-help state was followed by a frustrated and on-task state. This is very important since it may suggest that 40% of the time when students are bored but also receiving help, the help brings them back to on-task behavior while still frustrated. Two important take away messages from this finding are that receiving help may help

students go back to engagement, but the quality of received help may still keep them frustrated instead of concentrated. So it is important to look into implementing quality and accurate support systems that provide students with thorough assistance in times of impasse.B-OT (16.13%) → B-OT. About sixteen percent of the time students remained in a bored, on-task state. This suggests that ether their efforts were not enough to solve the problem and go back to concentration or that their beliefs in self-efficacy made them believe that they could not pass, so they still remained on task but were bored with their learning environment for at least two consecutive instances. This may impede learning. Another consideration was the ease of the game levels, however, given that the game was designed for middle school students and the study was conducted with elementary school students, this assumption is not very plausible.

13. B-OfT (8.5%) → B-OT. Unlike the previous case, in 8.5% of the time students went from B-OfT to B-OT. This may suggest that for some reason, while still bored, students were putting effort into giving another try to the learning environment and transitioning into being on task.
14. B-OOC (50%) → B-OfT. For 50% of the time, students went from B-OOC to B-OfT. This may suggest that students while initially still conversing with peers about the learning environment got completely disengaged with the environment and remained bored.
15. B-OfT (21.3%) → B-OfT. For about 21% of the time, students transitioned from B-OfT to another instance of B-OfT. Similarly to the previous result, this may

suggest that students had no desire to reengage with the environment and were overall bored and not even engaged or concentrating on any other relevant task.

16. B-OT (12.5%) → B-OfT. For about 12.5% of the time, students transitioned from B-OT to B-OfT, which is the dangerous situation discussed in affect literature, where the student's opportunity for learning is reduced.

Analysis 4: Student Affect and Behavior Characteristic

Research Question 6

This fourth analysis served to answer Research Question 6: Do students' affect sequences characterize their performance in the game? Cluster analysis at sequence and student levels was involved.

Clustering: Sequence Level

Cluster analysis was conducted on sequence-level data using the same method described earlier in the Analysis 2 section using all 272 unique sequences. Subsequently, in the cluster analysis on student level, these sequence-level clusters were linked back to the student gameplay in order to identify whether these clusters could help identify high-performing student and whether there was any relationship between students' affective performance and their gameplay performance. For instance, if students show predominant tendencies of confusion throughout the study, are they students who fell behind or those who reached the final game level?

Findings

Cluster analysis of all the sequences revealed how these sequences were grouped. Figure 14 presents the dendrogram of the data of the hierarchical clustering method applied to these data (i.e., 272 sequences). The y-axis shows the height at which clusters split into smaller groupings. In hierarchical clustering, height indicates the distance between the objects in different clusters. Depending on the selected clustering method, this parameter will be calculated differently (e.g., ward's, between linkage, within linkage, etc.).

Using the information about cluster membership counts and cluster distances (as discussed earlier in Chapter 3, methods section; Everitt, 1980; Kiran, Serra, & Cousty, 2012), a five-cluster solution was selected as performing the best in terms of identifying most meaningful cluster groupings (3 to 8 cluster solutions were performed before selecting 5 cluster solution).

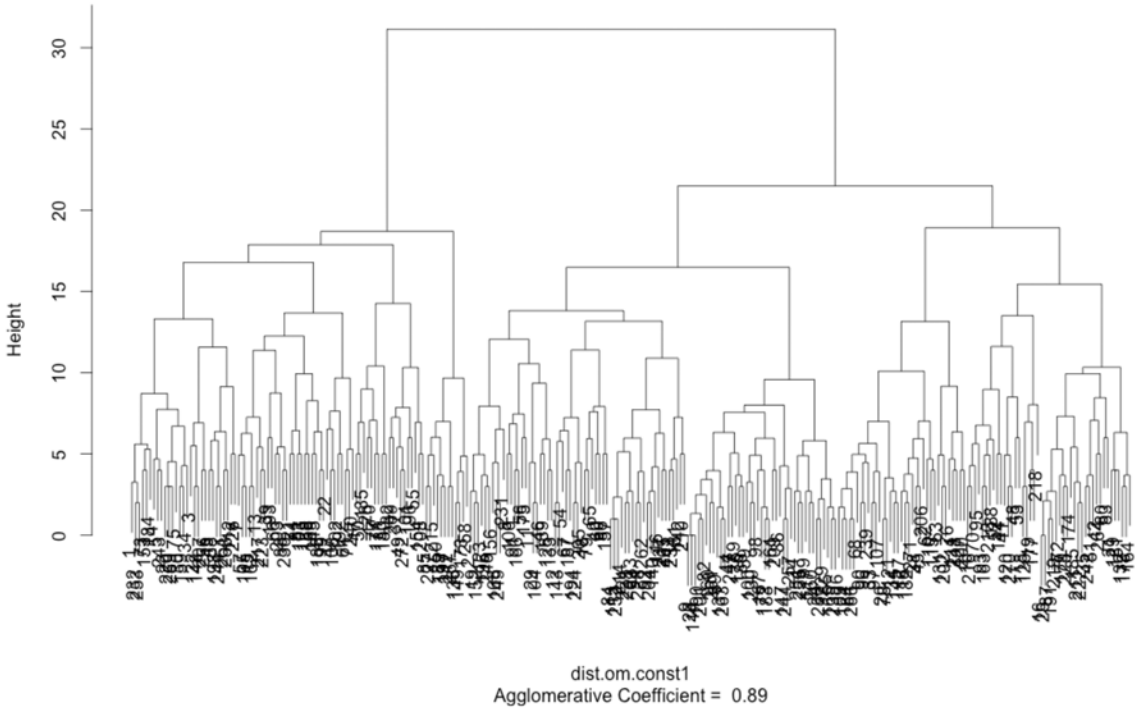


Figure 14. Dendrogram of sequence clusters.

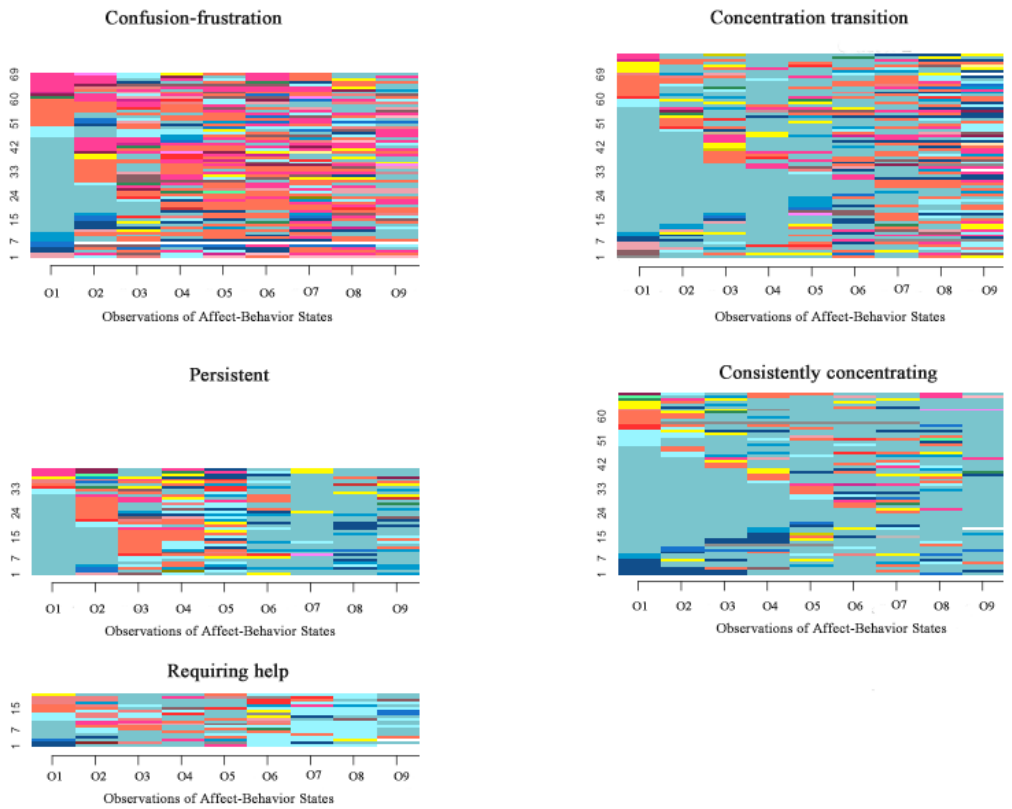


Figure 15. Five-cluster solution. These clusters have maximum length of nine observations. Some of the sequences may have white states, which means that those rare sequences were shorter than nine observations.

After reviewing all five clusters, I named each cluster based on their visual affect-behavior content. I found the following clusters: confusion-frustration cluster, concentration transition cluster, persistent cluster, consistently concentration cluster, and requiring help cluster. Below I descried each of the clusters in detail.

Cluster 1: The confusion-frustration cluster. This cluster starts off with a large chunk of sequences displaying concentration and on-task behavior. Beyond Observation 2 up until the last observation of the row, the predominant affective state was confusion with some sparks of frustration and boredom.

Cluster 2: The concentration transition cluster. This cluster starts off with much more concentration than Cluster 1. However, halfway through the observations, it transitions into mixed confusion, frustration, and some sparks of delight.

Cluster 3: The persistent cluster. This cluster starts out predominantly with concentration but immediately transitions into a pattern of persistent confusion. However, this persistent confusion does spread beyond Observation 4, at which point students show delight and return to concentrating and sometimes giving help.

Cluster 4: The consistently concentrating cluster. This is very similar to the concentration transition cluster, except that it predominantly shows concentration throughout the entire observation sequence. It starts out with mostly concentration on task and ends with the same pattern of concentration on task.

Cluster 5: The requiring-help cluster. This cluster shows a predominant presence of concentration and requiring-help (C-RH) states. In addition, this is by far the smallest cluster in terms of cluster membership.

Clustering: Student Level

The cluster analysis of all the student sequences of cluster memberships revealed how students were grouped based on their affective and behavioral observations. Figure 16 presents the dendrogram of the hierarchical clustering method applied to these data (selection was done as discussed earlier in Chapter 3, methods section; Everitt, 1980; Kiran et al., 2012). The y-axis shows the height at which clusters split into smaller groupings.

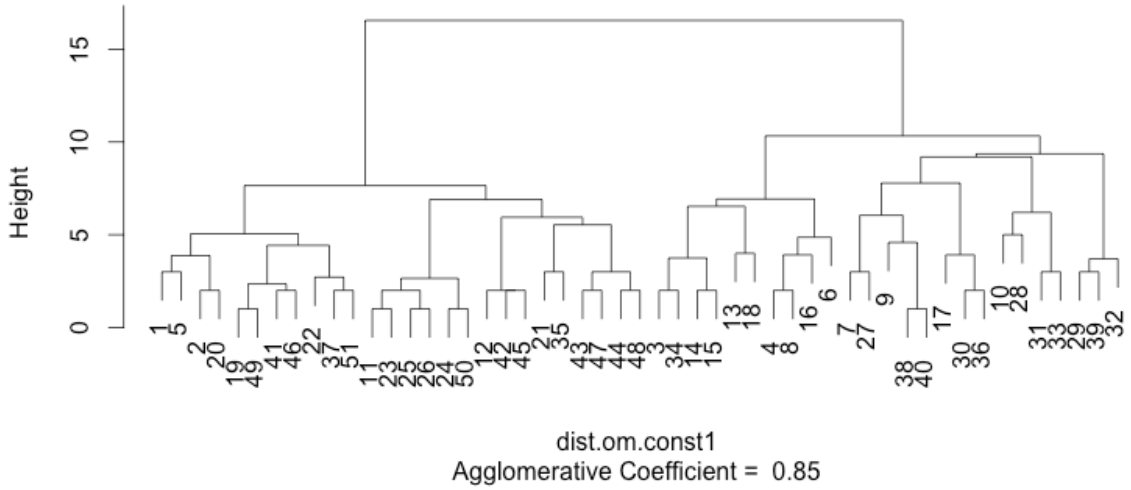


Figure 16. Dendrogram of student cluster membership sequences.

Using the information about cluster membership counts and cluster distances (see Chapter 3, methods section; Everitt, 1980; Kiran et al., 2012), a three-cluster solution was selected as performing the best in terms of identifying most meaningful student groupings. See Figure 17.

Cluster 1: Struggling. In the struggling cluster, students seem to show equal number of the requiring-help cluster and the confusion-frustration cluster and a bit higher presence of concentration transition cluster.

Cluster 2: Concentrated students. In the concentrated students cluster, students mostly relate to the consistently concentrating cluster.

Cluster 3: Persistent students. In the persistent students cluster, students show membership in the confusion-frustration cluster and concentration transition cluster. This suggests that these students showed a relatively higher level of confusion, especially towards the end time points; however, they started off mostly with concentration and on-task behavior.

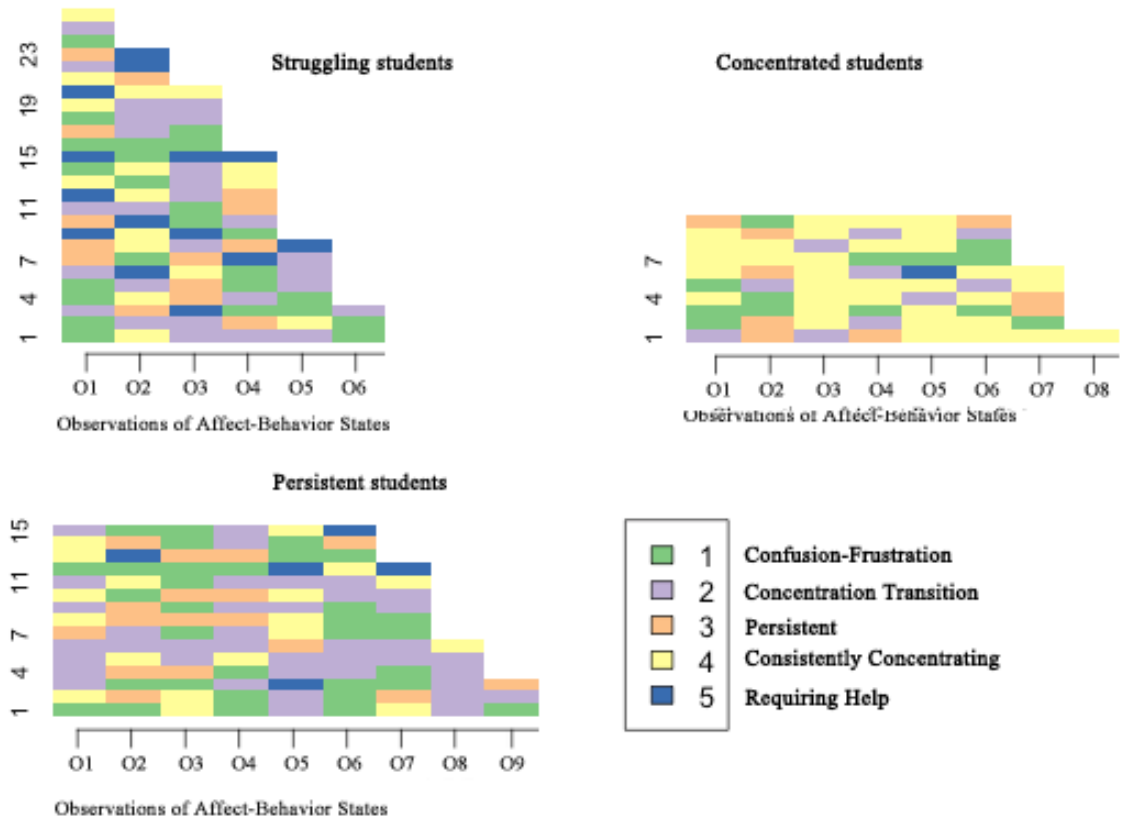


Figure 17. Three-cluster solution.

Relating Cluster Membership to Game Performance

For this study there were six worlds, and each world had 30 levels. One might think that students in the concentrated group would show the most advanced results in the game relative to their peers (i.e., have reached the final level of the game). Another conjecture could be that students in the struggling cluster would be further behind in the game levels than their peers (i.e., between the first and second worlds of the game), whereas the persistent students might start up slowly and reach only the third world of the game.

To investigate these conjectures, an R function was used to connect student cluster memberships to their gameplay data in order to investigate their performance in the game. In this case, gameplay data consisted of two variables: number of total game worlds completed and the number of the final level completed (e.g., World 6, Level 8). Table 10 shows the results of these analyses.

Findings

According to the results, although there is a difference in membership numbers between clusters, the cluster with the highest number of students (the struggling cluster, *n* = 26) completed the lowest number of average levels completed. In fact, the cluster with the lowest number of students (the concentrated students cluster, *n* = 10) completed the highest number of average levels completed. The standard deviation numbers indicate that in each of the clusters, the data points are spread out from the mean.

Table 10

Student Gameplay Analysis

Cluster membership	Student count	Sum of game levels completed	Average of game levels completed	Standard deviation
Struggling students (cluster 1)	26	2,801	107.7	31.5
Concentrated students (cluster 2)	10	1,469	146.9	33.2
Persistent students (cluster 3)	15	1,922	128.0	20.5

To visually depict this image, I used the R package rggobi graphical package that is based on GGobi (Swayne, Cook, & Buja, 1998; Swayne, Temple Lang, Buja & Cook, 2003;), a software package for the visualization and exploration of high-dimensional data. Figure 18 shows that whereas the struggling students cluster had a relatively higher number of total levels played, the concentrated students cluster had the highest student representation in the World 6 (final game world for this study) and no student representation in the Worlds 2 or 3. Tying this information back to the cluster characteristics suggests that the concentrated students cluster showing an abundance of concentration was in fact the leading group in terms of students who reached the final World 6 of the game. In addition, the struggling cluster, which was characterized by students who were mainly in the confusion-frustration and receiving-help domains, showed to have the most representation in World 2 and World 3, which means that the students who fell behind were mostly represented in this cluster. Finally, the persistent group was predominantly in the 5th World, not quite reaching the final World but not lingering in the initial Worlds. While this analysis allows me to tie student gameplay data to their affect data, these findings do not claim that students who have completed less levels are certainly in struggling cluster. The findings show that students, who have completed less game levels, also happen to appear in the struggling cluster.

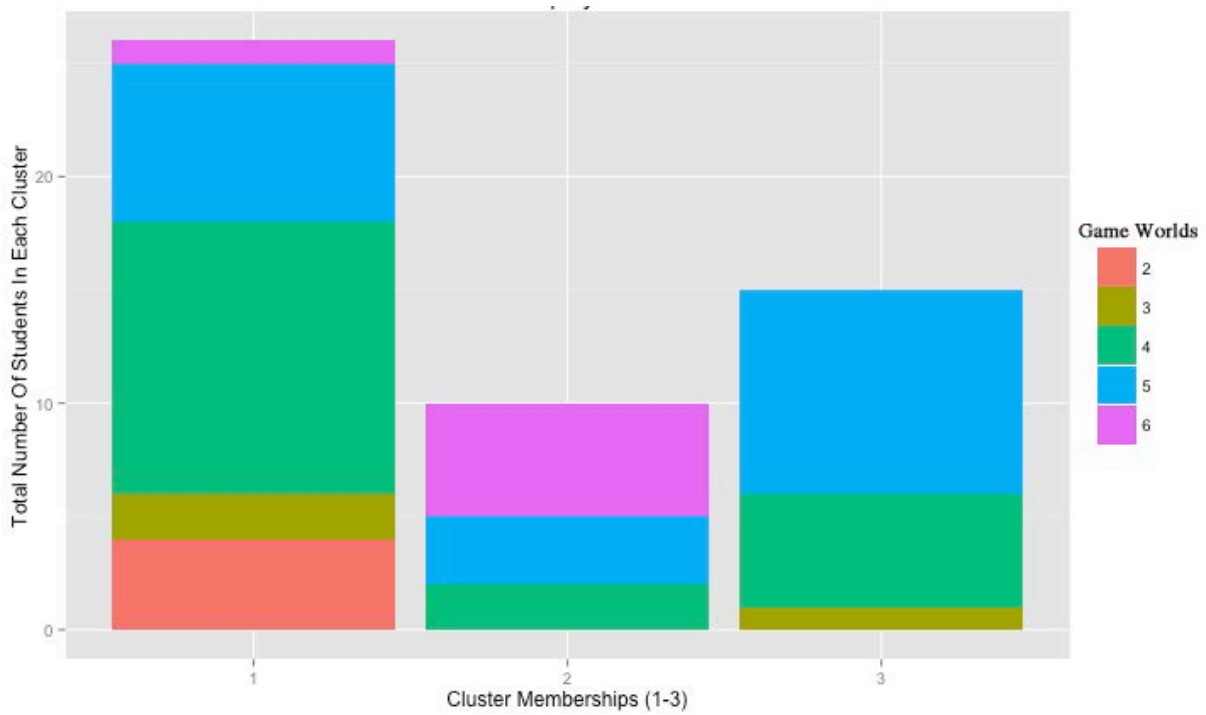


Figure 18. Student gameplay performance.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

Summary

In summary, in this dissertation, I have presented sequential analyses of student affect data to shed light on potentially existent patterns of emotions during students' gameplay. As identified earlier, this study used an exploratory research method that advocates for a bottom-up approach of theory generation. As described in my study design, I started the study with some conjectures and expectations about the data, and then I conducted observations, investigated underlying patterns, and tied the results back to conjectures and research questions in order to summarize my findings.

As Izard (2002) stated, emotions "contain useful information that can guide cognition and action" (p. 815). Hence, this study provided valuable insight into understanding some of the underlying patterns of emotion that form the information that guides students' academic behavior.

Revisiting the Research Questions

Research Question 1

The results of the first analysis displayed some of the most frequent sequences and overall picture of the data. However, these findings suggest that the most frequent 10 sequences comprise only about 4.8% of the data. Hence, this is suggestive of a lack of consistency in the frequent sequences. My findings of low-frequency nature of some of

the data points (e.g., frustration, confusion, boredom, etc.) is consistent with what has been found in prior research (D'Mello, 2013). The infrequency of some of these states may be the result of a small sample size. However, this study was designed to be an exploratory study that aimed to discover whether such data may have consistent patterns that would allow for targeting a general temporal model of affective states exhibited in digital games.

Research Questions 2–5

This analysis unveiled several interesting characteristics of the sequence data. In regards to boredom, most of the time boredom seemed to occur after a student was in a F-RH state. This may be an alarming finding given that the student was receiving help before transitioning to boredom. This may be indicative of poor help quality. For some of the clusters, once boredom showed up, it became a persistent state (once it occurred, it repeated for several consecutive observations). In addition, findings highlight that B-RH was usually a persistent state. This may suggest that either the student was trying to invest some effort into succeeding in the game but was not receiving quality support to cease being bored, or the student was hearing a peer talk but was not invested in the game enough to care about understanding or using the provided help. This may have important implications in the design of educational digital games and also for the development of embedded support systems that can both provide help and keep the students' attention on the help content.

Another finding suggests that about 40% of the time students in the B-RH state transitioned to F-OT. This might be the result of a student's effort to persist. However,

another 17.3% of the time, a student in the F-OT state would remain in that state for at least another observation point. One factor for this pattern might be that the provided help (B-RH) was not long enough for the students to actually grasp the solution and subsequently transition from F-OT to C-OT on their own. On the other hand, it also may suggest that the support got the student back in the environment, because the student felt the pressure to act upon the provided guidance, but the guidance was not sufficient to motivate the student to reengage in the environment (e.g., was not personalized, was not clear, etc.). In addition, the lack of quality help may be related to student's self-efficacy and may make them doubt their abilities. It is also important to mention that the 40% may not represent too many cases, but in the process of developing a affect-behavior model for affect research, this occurrence informs the research about such possibility.

In regards to frustration, Figure 13 from Analysis 3 showed an interesting pattern: not only were there quite a few transitions to F-OT from C, but also the transition from concentration to frustration took place without any intermediate affective state such as confusion, which is what some of the affect literature has proposed. Finally, the other important transition was the transition from CF-OT to B-OfT. This finding refutes my initial hypothesis that the transition to boredom would take place with the intermediate state of frustration.

Finally, Analysis 2 showed that some of my initial hypotheses did not hold true, whereas others found some evidence in the data.

1. I hypothesized that persistent frustration will develop into confusion. Unfortunately, most cases of persistent frustration would be the end of the

sequence, which would inhibit me from observing what happens next. However, a few cases that had further states showed equal numbers of transitions into concentration as into confusion. Hence, the answer to this hypothesis remains inconclusive.

2. I also hypothesized that simple frustration will transform into concentration, and persistent frustration will transform into confusion after the students receive help. Again, there were not enough data to support the conjecture of the transition of persistent frustration into confusion. However, there were similar amounts of simple frustration cases that transitioned to concentration, confusion, and even to boredom. Overall, there were not enough data to investigate this conjecture in more detail.
3. Finally, I hypothesized that boredom will occur primarily after persistent frustration. This conjecture did not receive data-based verification. In fact, most of the time the boredom state occurred after the student was in an F-RH state, frustrated and receiving help. This may be the result of poor help; hence, it is important that future studies to look into the quality of support provided (peer vs. professional support). For some of the clusters, once boredom showed up, it became a persistent state (once it occurred, it repeated for several consecutive observations). For some others, that was not the case. In addition, B-RH is usually a persistent state: Even though the students were receiving help while bored, they remained in that bored and receiving-help state for several observations. This may suggest that the student is trying to

invest some effort into succeeding in the game but is not receiving sufficient support to cease being bored.

Research Question 6

The analysis for RQ6, while not conclusive, suggests an interesting grouping of sequences and students' affective characteristics. According to the results, this analysis suggests that the students who showed persistent concentration throughout their affect sequences were also the students who reach the highest level of the game environment and did not get lost in the initial levels. On the other hand, the students who showed persistent transitions between confusion and concentration had a better game performance than the students who were struggling (going back and forth between confusion, frustration, and receiving-help states). Hence, although in one case (i.e., persistent students) confusion seemed to be a positive factor, in other cases a combination of confusion and frustration over time resulted in poor performance in the game. This may be suggestive of unhelpful assistance or being in a state of confusion or frustration for too long or of too high intensity.

Overall, previous studies suggested that confusion transitions into frustration (Kort, Reilly, & Picard, 2001), whereas there is no jump from confusion to boredom, and frustration leads to boredom (Perkins & Hill, 1985). This study's findings highlight the double-sided nature of the relationship between these affective states: in this data set there was evidence that at times confusion transition to frustration, but also frustration transitions to confusion, confusion transitions to boredom, boredom leads to frustration, and concentration skips confusion when transitioning into frustration. Hence, this

suggests that to build a model of affective states in an educational digital game environment, there is a need for more extensive study with far denser observational data points.

Limitations

Generalizability Limitations

While the generalizability of this study's findings to a much wider population than one Utah county's fifth-grade students may be limited due to the specificity of the study sample, it is a consideration that could be addressed in future research. In addition, another limitation related to the sample size is the density of the data. This study was designed to accommodate dense data, but it will be useful to consider using more than one observer and thus decrease the seconds gap between observations.

Data Limitations

It is a concern that some of the students were scheduled to leave the class earlier due to other school related responsibilities (e.g., lunch helpers). Hence, some of the students had fewer observations than others. In addition, these students may have been chosen as lunch helpers due to their better academic performance, hence, a randomization of student sets could have addressed this potential bias. This is something to consider for future studies.

Limitations of Sequence Clustering Analysis

One of the limitations of this package and this analysis approach is that TraMineR uses 1 as the default value for substitution, insertion, and deletion cost (Gabadinho et al.,

2011). I was not able to find a theoretical basis for such a cost value; therefore, I suggest maintaining that number for all substitutions, insertions, and deletions. It is important to highlight that insertion and deletion should equate to 1 (or any other arbitrary number), whereas substitution cost should be the addition of both insertion and deletion (e.g., if insertion/deletion cost = 1, then substitution cost = insertion cost + deletion cost = 2).

One of the biggest flaws of the optimal matching technique is the selection of appropriate transformation cost (Wu, 2000). In my analysis I set all three methods of insertion, deletion, and substitution to equate to 1; hence, they all had equal weight in defining my data dissimilarity. However, in the future work it would be beneficial to reconstruct the matrix into considering an alternative weight calculation: If a state in the first sequence is partially identical to the state in the second sequence (e.g., C-OT vs. C-OfT), then this transformation would have a cost of 1. However, if a state in the first sequence is completely dissimilar to the state in the second sequence (e.g., C-OT vs. CF-OfT), then this transformation would have a cost of 2. This is a suggestion for consideration in future work.

Limitations of Hierarchical Clustering Analysis

In each case of the hierarchical clustering analysis, due to sequence length discrepancy (even though small), the clustering may take place based on length differences instead of content differences. Based on the results, it seems that was the case only in Analysis 2. However, in that analysis, I considered this limitation advantageous, because it allowed me to separate the short sequences from the long sequences, which I could use for my pattern identification.

In addition, in this dissertation work I have selected to incorporate a clustering approach suggested by TraMineR package. However, upon inspecting some of my results (see Figure 15), I have noticed that there are some sequences that are very similar. Nonetheless they appear in different clusters. For example, there are 3 sequences with complete concentration states for all 9 observations that were places in Consistently concentrating cluster. At the same time, another sequence with 8 observations of concentration and one observation of confusion affect-behavior instance was places in the Concentration transition cluster. This unexpected results can be attributed to the Ward's method used in my clustering algorithm. Upon the review of TraMineR package documentation, I have realized that this finding is obviously related to the method used in this algorithm since based on their example sequence analysis (Figure 9.2 pg 102 TraMineR, 2014), I found similar unexpected clustering results. Therefore, while I have chosen Ward's clustering method for my hierarchical clustering approach, I believe it will be worthwhile to investigate other methods in future studies. It will be useful to compare the results of all possible methods for such analysis and contrast the interpretability of each of the findings.

Contributions

These findings have several potential implications: theoretical, instructional, and game development implications. From theoretical perspective, this work contributes to our knowledge of student emotions and engagement in game-based environments. Particularly, this work demonstrates that although there may be no strongly defined affect sequence patterns in this small dataset, there are important transitional patterns that provide a better understanding of transitions from confusion to frustration and boredom and their accompanying behavior. This knowledge on possibly introducing vicious cycles of negative affect will drive our instructional support strategies. This study also has instructional implications suggesting teachers identify the best time for instructional interventions. In addition, this study provided an understanding of the classroom dynamics during gameplay. Finally, this study offers design implications for educational game development and adaptive systems designs. Based on the results, I would like to emphasize two of the main educational implications of this study:

1. The study findings show that students do indeed experience a considerable amount of frustration, confusion, and boredom even within interactive game interfaces. Some of these states are simple, whereas others occur more than once (i.e., persistent). Hence, this study helps to understand the circumstances (in terms of affect and behavior) that bring students out of these affective states or do not instigate negative valence.
2. Students seem to be reaching out for help a considerable amount of time. However, their received help is not always helping them return to the

concentration state. Therefore, this study emphasizes the importance of investigating the quality of help and evaluating peer-help effectiveness compared to professional assistance from within the game. After all, optimal learning environments that are able to respond to students' emotional response to the digital learning environment will promote students engagement (Shernoff, 2013).

Future Work

This dissertation raises some questions and suggestions for future work. First, it will be beneficial to replicate this study with a much larger dataset and much more dense data collection design (reducing the potential 51-second gap between each observation). This will ensure that sequences are uninterrupted and the transitions are completely reliable. In addition, considering classroom dynamics, in the future it would be beneficial to randomize the observation order before each session. However, for this particular study a decision was made to shuffle observations between the student sets in an organized manner (selecting students from front, middle, and end of the seating chart) in order to be less obtrusive. In regards to specific analyses, in the case of inter-sequence distance analysis, it would be beneficial to alter the cost of transformation based on theoretical understanding of affect and their distance from each other: Is frustration on task closer to confusion on task or to delight on task? In addition, it will be worthwhile to consider whether there should be a cost adjustment difference between combinations with entirely identical states and combinations with partially identical states. For example, is the transformation of concentration on task to concentration off task similar to the

transformation of concentration on task to confusion receiving help, or should the cost of the first be 1 while the cost of the latter is 2?

Finally, while I did not intend to investigate gender differences within the scope of this study, overall observations highlighted some apparent differences between both genders in terms of gameplay strategies. A future study could delve deeper into student gender differences in regards to their efforts and persistence.

The results of this study showed that there is a demand for help during student gameplay. Moreover, some of the results suggested that it may be important to evaluate the quality of this help as it may influence a student's further disposition toward the learning environment. Hence, it is assumed that further research in the use of help systems is warranted. In fact, future work can compare the use of embedded support systems to a peer support system. It would be beneficial to evaluate the difference between the efficiency and success rate of these two support systems.

REFERENCES

Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology review and prospect. *Sociological Methods & Research, 29*(1), 3-33.

Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., & Verkamo, A. I. (1996). Fast discovery of association rules. In U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), *Advances in knowledge discovery and data mining* (pp. 307-328). Menlo Park, CA: AAAI Press.

Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the 11th International Conference on Data Engineering* (pp. 3-14).

Amory, A., Naicker, K., Vincent, J., & Adams, C. (1999). The use of computer games as an educational tool: Identification of appropriate game types and game elements. *British Journal of Educational Technology, 30*, 311-321.

Azarnoush, B., Bekki, J. M., Runger, G. C., Bernstein, B. L., & Atkinson, R. K. (2013). Toward a framework for learner segmentation. *Journal of Educational Data Mining, 5*(2), 102-126.

Bailey, K. (1994). *Typologies and taxonomies*. Thousand Oaks, CA: Sage.

Baker, R. S. J. d. (2011). Gaming the system: A retrospective look. *Philippine Computing Journal, 6*(2), 9-13.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the Cognitive Tutor classroom: When students "game the system." In *Proceedings of ACM CHI 2004: Computer–Human Interaction* (pp. 383-390).

Baker, R. S. J. d., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68,* 223-241.

Baker, R. S. J. d., Gowda, S. M., Wixon, M., Kalka, J., Wagner, A. Z., Salvi, A., . . . Rossi, L. (2012). Towards sensor-free affect detection in Cognitive Tutor Algebra. In *Proceedings of the 5th International Conference on Educational Data Mining* (pp. 126-133). Chania, Greece.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Barab, S. A., Zuiker, S., Warren, S., Hickey, D. Ingram-Goble, A., Kwon, E.-J., . . . Gerring, S. C. (2007). Situationally embodied curriculum: Relating formalisms and contexts. *Science Education, 91*, 750-782.

Bartel, C. A., & Saavedra, R. (2001). The collective construction of work group moods. *Administrative Science Quarterly, 45,* 197-231

Baumeister, R. F., & Bushman, B. J. (2007). Angry emotions and aggressive behaviors. In G. Steffgen & M. Gollwitzer (Eds.), *Emotions and aggressive behavior* (pp. 61-75). Ashland, OH: Hogrefe & Huber.

Becker, K. (2001). Teaching with games—The Minesweeper and Asteroids experience. *Journal of Computing in Small Colleges, 17*(2), 22-32.

Beentjes, J. W. J., Koolstra, C. M., Marseille, N., & van der Voort, T. H. A. (2001). Children's use of different media: For how long and why? In S. Livingstone & M. Bovill (Eds.), *Children and their changing media environment: A European comparative study* (pp. 85-111). Mahwah, NJ: Erlbaum.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57,* 111-127.

Blashfield, R. K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. *The Psychological Bulletin, 83,* 377–388.

Bonferroni, C. E. (1936). *Teoria statistica delle classi e calcolo delle probabilita.* Florence, Italy: Libreria Internazionale Seeber.

BrainPlay. (2013a). *About BrainPlay*. Retrieved from http://www.BrainPlay.com/notices/about

BrainPlay. (2013b). *Welcome to BrainPlay!* Retrieved from http://www. BrainPlay.com/user/signup

Buchta, C., Hahsler, M., & Diaz, D. (2014). *arulesSequences: Mining frequent sequences. R package version 0.2-6.* Retrieved from http://CRAN.R-project.org/package=arulesSequences

Cavallo, A. M., & Laubach, T. A. (2001). Students' science perceptions and enrollment decisions in differing learning cycle classrooms. *Journal of Research in Science Teaching, 38*, 1029-1062.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Cohen-Scali, V. (2003). The influence of family, social, and work socialization on the construction of the professional identity of young adults. *Journal of Career Development, 29*, 237-249.

Corno, L., & Mandinach, E. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist, 18,* 88-108.

Craig, S., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media*, *29,* 241-250.

Damasio, A. R. (2004). Emotions and feelings: A neurobiological perspective. In A. S. R. Manstead, N. Frijda, & A. Fischer (Eds.), *Feelings and emotions* (pp. 49-57). Cambridge, England: Cambridge University Press.

Dempsey, J. V., Haynes, L. L., Lucassen, B. A., & Casey, M. S. (2002). Forty simple computer games and what they could mean to educators. *Simulation & Gaming: An Interdisciplinary Journal, 33,* 157-168.

Dempsey, J. V., Lucassen, B. A., Haynes, L. L, & Casey, M. S. (1997). *An exploratory study of forty computer games* (COE Technical Report No. 97-2). Mobile, AL: University of South Alabama.

Denis, G., & Jouvelot, P. (2005). Motivation-driven educational game design: Applying best practices to music education. In *ACE '05: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology* (pp. 462-465). New York, NY: ACM Press.

Desmarais, M.C. & Lemieux, F. (2013). *Clustering and Visualizing Study State Sequences*. In 5th Conference on Educational Data Mining, EDM2013, Memphis, TN

Dickey, M. D. (2005). Three-dimensional virtual worlds and distance learning: Two case studies of Active Worlds as a medium for distance education. *British Journal of Educational Technology, 36*, 439-451.

Dickey, M. D. (2006). "Ninja Looting" for instructional design: The design challenges of creating a game-based learning environment. In *Proceedings: ACM SIGGRAPH 2006 educators program* (Article No. 17). doi:10.1145/1179295.1179313

D'Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology, 105,* 1082-1099.

D'Mello, S. K., & Graesser, A. C. (2011). The half-life of cognitive-affective states during complex learning. *Cognition and Emotion, 25,* 1299-1308.

D'Mello, S. K., & Graesser, A. C. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22,* 145-157.

D'Mello, S., Picard, R., & Graesser, A. (2007). Towards an affect-sensitive autotutor. *IEEE Intelligent Systems, 22*(4), 53-61.

D'Mello, S. K., Taylor, R., & Graesser, A.C. (2007). Monitoring affective trajectories during complex learning. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 203-208). Austin, TX: Cognitive Science Society.

Dragon, T., Arroyo, I., Woolf, B. P., Burleson, W., El Kaliouby, R., & Eydgahi, H. (2008). Viewing student affect and learning through classroom observation and physical sensors. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems: 9th International Conference, ITS 2008, proceedings* (pp. 29-39). Berlin, Germany: Springer-Verlag.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*, 939-944.

Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield & J. Eccles (Eds.), *The development of achievement motivation* (pp. 57-88). San Diego, CA: Academic Press.

Dziabenko, O., Pivec, M., & Schinnerl, I. (2003). *Game scenario: UniGame: Social skills and knowledge training.* Retrieved from http://www.unigame.net/html/case _studies/Game_scenario.pdf

Ebner, M., & Holzinger, A. (2006). Successful implementation of user-centered game based learning in higher education: An example from civil engineering. *Computers & Education, 49*, 873-890.

Edwards, T., Bardar, E., Asbell-Clarke, J., & Larsen, J. L. (2013, June). *Quantum Spectre*. Paper presented at the Games+Learning+Society 9.0 (GLS 9.0) conference, Madison, WI.

Ekman, P. (2005). Basic emotions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 45-60). Chichester, England: John Wiley & Sons.

Entertainment Software Association. (2001). *State of the industry 2000–2001.* Retrieved from http://www.theesa.com

Everitt, B. (1980). *Cluster analysis* (2nd ed.). London, England: Heineman Educational Books.

Federation of American Scientists. (2006). *Report: Summit on educational games: Harnessing the power of video games for learning.* Washington, DC: Available from http://www.fas.org/gamesummit/

Feierabend, S., & Klingler, W. (2001). Children and media 2000. PC/Internet gain importance. *Media Perspektiven, 7*, 345-357.

Fisch, S. M. (2005). Making educational computer games "educational." In *Proceedings of the 2005 Conference on Interaction Design and Children* (pp. 56-61). doi:10 .1145/1109540.1109548

Fleiss, J. L. (1981) *Statistical methods for rates and proportions* (2nd ed.). New York, NY: Wiley.

Foreman, J. (2004, September/October). Game-based learning: How to delight and instruct in the 21st century. *EDUCAUSE Review.* Retrieved from http://www .educause.edu/ir/library/pdf/ERM0454.pdf

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56,* 218-226.

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*. *40*(4), 1-37.

Gee, J. P. (2007a). *Good video games + good learning: Collected essays on video games, learning, and literacy*. Bern, Switzerland: Peter Lang.

Gee, J. P. (2007b). *What video games have to teach us about learning and literacy.* New York, NY: Palgrave Macmillan.

Gibson, H., & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education, 86*, 693-705.

Gihman, I. I., & Skorohod, A. V. (1975). *The theory of stochastic processes.* Berlin, Germany: Springer.

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*, 289-308.

Grigg, W. S., Lauko, M. A., & Brockway, D. M. (2006). *The nation's report card: Science 2005.* Washington, DC: National Center for Education Statistics.

Gumora, G., & Arsenio, W. F. (2002). Emotionality, emotion regulation, and school performance in middle school children. *Journal of School Psychology, 40,* 395-413.

Hämäläinen, R., Manninen, T., Järvelä, S., & Häkkinen, P. (2006). Learning to collaborate: Designing collaboration in a 3D game environment. *The Internet and Higher Education, 9*(1), 47-61.

Haines, S. J. (1981). Six statistical suggestions for surgeons. *Neurosurgery, 9*(4), 414-418.

Hamming, R. W. (1950), Error detecting and error correcting codes. *Bell System Technical Journal 29*(2), 147–160.

Hands, S., & Everitt, B. (1987). A Monte Carlo study of the recovery of cluster structure in binary data by hierarchical clustering techniques. *Multivariate Behavioral Research, 22,* 235–243.

Higgins, S. (2000). The logical Zoombinis. *Teaching Thinking, 1*(1), 15-18.

Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical methods*. New York, NY: Wiley.

Honey, M. A., & Hilton, M. (Eds.). (2011). *Learning science through computer games and simulations.* Washington, DC: National Academies Press.

Inkpen, K. M., Booth, K. S., Gribble, S. D., & Klawe, M. M. (1995). Give and take: Children collaborating on one computer. In J. M. Bowers & S. D. Benford (Eds.), *CHI 95: Human Factors in Computing Systems, Denver, CO, ACM conference companion* (pp. 258-259).

Izard, C. (2002). Translating emotion theory and research into preventive interventions. *Psychological Bulletin, 128,* 796-824.

Izard, C., Stark, K., Trentacosta, C., & Schultz, D. (2008). Beyond emotion regulation: Emotion utilization and adaptive functioning. *Child Development Perspectives, 2*, 156-163.

Jennings, M. (2001). Best practices in corporate training and the role of aesthetics: Interviews with eight experts. In *Proceedings of the 2001 ACM SIGCPR Conference on Computer Personnel Research* (pp. 215-219). doi:10.1145/371209 .371239

Juul, J. (2013). *The art of failure: An essay on the pain of playing video games*. Cambridge, MA: The MIT Press.

Kambouri, M., Mellar, H., & Logan, K. (2006). Adult learners and ICT: An intervention study in the UK. In W. Nejdl & K. Tochtermann (Eds.), *EC-TEL, 2006, LNCS 4227. Book series: Lecture notes in computer science* (pp. 213-226). Berlin, Germany: Springer.

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education.* San Francisco, CA: Pfeiffer.

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*, 379-424.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences, 21*(1), 45-83.

Karweit, N., & Slavin, R. E. (1982). Time-on-task: Issues of timing, sampling, and definition. *Journal of Experimental Psychology, 74*, 844-851.

Kiran, B. R., Serra, J., & Cousty, J. (2012). Climbing: A unified approach for global constraints on hierarchical segmentation. In *Computer Vision–ECCV 2012, workshops and demonstrations* (pp. 324-334). Berlin, Germany: Springer.

Kirriemuir, J., & McFarlane, A. (2004). *Literature review in games and learning*. Retrieved from the Futurelab website: http://telearn.archives-ouvertes.fr/docs/00/ 19/04/53/PDF/kirriemuir-j-2004-r8.pdf

Klawe, M. (1999). Computer games, education and interfaces: The E-GEMS project. In *Proceedings of the Graphics Interface Conference* (pp. 36-39). Available from http://pdf.aminer.org/000/235/915/computer_games_education_and_interfaces_th e_e_gems_project.pdf

Kort, B., Reilly, R., & Picard, R. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy—Building a learning companion. In *Proceedings of the IEEE International Conference on Advanced Learning Technology: Issues, achievements and challenges* (pp. 43-48). Available from http://affect.media.mit.edu/projectpages/lc/icalt.pdf

Kuiper, F. K., & Fisher, L. (1975). A Monte Carlo comparison of six clustering procedures. *Biometrics, 31*, 777–783.

Kumar, K. M., Srinivas, P. V. S., & Rao, C. R. (2012). Sequential pattern mining with multiple minimum supports by MS-SPADE. *International Journal of Computer Sciences, 9*(5), 61-73.

Ladd, G. W., Birch, S. H., & Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child Development, 70*, 1373-1400.

Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality, 21*(1), 1-39.

Lee, D. M., Rodrigo, M. M., Baker, R. S. J. d., Sugay, J., & Coronel, A. (2011). Exploring the relationship between novice programmer confusion and achievement. In *Proceedings of the 4th Bi-Annual International Conference on Affective Computing and Intelligent Interaction.* Available from http://www .columbia.edu/~rsb2162/acii_lee_et_alv5.pdf

Lehman, B., D'Mello, S. K., & Graesser, A. C. (2012). Interventions to regulate confusion during learning. In S. Cerri, W. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems* (pp. 576-578). Berlin, Germany: Springer-Verlag.

Lehman, B., D'Mello, S. K., Strain, A. C., Gross, M., Dobbins, A., Wallace, P., . . . Graesser, A. C. (2011). Inducing and tracking confusion with contradictions during critical thinking and scientific reasoning. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial Intelligence in Education: 15th international conference* (pp. 171-178). Berlin, Germany: Springer.

Lenhart, A., Madden, M., Smith, A., & Macgill, A. R. (2007). *Teens and social media.* Retrieved from the Pew Research Center website: http://www.pewinternet.org/ 2007/12/19/teens-and-social-media/

Levenshtein, V. I. (February 1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady 10*(8), 707–710.

Linnenbrink, E. A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 107-124). San Diego, CA: Elsevier Academic Press.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational Psychologist, 37*, 69-78.

Liu, Z., Pataranutaporn, V., Ocumpaugh, J., & Baker, R. S. J. d. (2013). Sequences of frustration and confusion, and learning. In *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 114-120). Available from http://www.educationaldatamining.org/EDM2013/papers/rn_paper_18.pdf

Lorr, M. (1983). *Cluster analysis for social scientists: Techniques for analyzing and simplifying complex blocks of data.* San Francisco, CA: Jossey-Bass.

Ma, X., & Wilkins, J. L. (2002). The development of science achievement in middle and high schools: Individual differences and school effects. *Evaluation Review, 26*, 395-417.

Magnussen, B. F. (2005). The eddy dissipation concept: A bridge between science and technology. In *Proceedings of the ECCOMAS thematic conference on computational combustion* (pp. 21-24). Available from http://folk.ntnu.no/ivarse/edc/BFM_ECOMAS2005_Lisboa.pdf

Markov, A. A. (1954). *Theory of algorithms* (J. J. Schorr-Kon, Trans.). Moscow, Russia: Imprint Moscow, Academy of Sciences of the USSR.

Markov process. (2014). In *Encyclopædia Britannica*. Retrieved from http://www .britannica.com/EBchecked/topic/365797/Markov-process

Mayo, M. J. (2009). Video games: A route to large-scale STEM education? *Science, 323*, 79-82.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. New York, NY: Penguin Press.

Mentis, H. M. (2007). Memory of frustrating experiences. In D. Nahl & D. Bilal (Eds.), *Information and emotion* (pp. 197-210). Medford, NJ: Information Today

Meyer, D., & Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review, 18,* 377-390. doi:10.1007/ s10648-006-9032-1

National Research Council. (2005). *America's lab report: Investigations in high school science.* Washington, DC: National Academies Press.

National Research Council. (2011). *Learning science through computer games and simulations.* Washington, DC: National Academies Press.

National Research Council and Institute of Medicine. (2004). *Engaging schools: Fostering high school students' motivation to learn.* Washington, DC: National Academies Press.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*, 443-453.

Newmann, F. W., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. W. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11-39). New York, NY: Teachers College Press.

NPD Group. (2011). *The video game industry is adding 2–17-year-old gamers at a rate higher than that age group's population growth.* Retrieved from https://www.npd.com/wps/portal/npd/us/news/pressreleases/pr_111011

O'Brien, R. G. (1979). A general ANOVA method for robust tests of additive models for variances. *Journal of the American Statistical Association, 74,* 877-880.

Ocumpaugh, J., Baker, R. S. J. d., & Rodrigo, M. M. T. (2012). *Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. training manual version 1.0.* New York, NY: EdLab.

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal, 16,* 455-474.

Oosterwegel, A., Field, N., Hart, D., & Anderson, K. (2001). The relation of self-esteem variability to emotion variability, mood, personality traits, and depressive tendencies. *Journal of Personality, 69*, 689-708.

Organisation for Economic Co-operation and Development. (2007). *Programme for International Student Assessment (PISA) 2006: Science competencies for tomorrow's world* (vol. I). Paris, France: Author. Retrieved from http://www.oecd.org/document/2/0,3343,en_32252351_32236191_39718850_1_1_1_1,00.html

Pahl, C., & Donnellan, C., 2003. Data mining technology for the evaluation of web-based teaching and learning systems. In M. Driscoll & T. Reeves (Eds.), *Proceedings of the World Congress on E-Learning in Corporate, Government, Healthcare, and Higher Education, Montreal, Canada* (pp. 1-7). Chesapeake, VA: Association for the Advancement of Computing in Education.

Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., & Gowda, S. M. (2013). Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 117-124). New York, NY: ACM Press.

Patrick, B. C., Skinner, E. A., & Connell, J. P. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain. *Journal of Personality and Social Psychology, 65,* 781-791.

Pekrun, R. (1984). An expectancy-value model of anxiety. In H. M. van der Ploeg, R. Schwarzer, & C. D. Spielberger (Eds.), *Advances in test anxiety research* (Vol. 3, pp. 53-72). Lisse, The Netherlands: Swets & Zeitlinger.

Pekrun, R. (1988). *Emotion, motivation, und persönlichkeit* [Emotion, motivation and personality]. Munich, Germany: Psychologie Verlags Union.

Pekrun, R. (1992). The expectancy-value theory of anxiety: Overview and implications. In D. G. Forgays, T. Sosnowski, & K. Wrzesniewski (Eds.), *Anxiety: Recent developments in self-appraisal, psychophysiological and health research* (pp. 23-41). Washington, DC: Hemisphere.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315-341.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*, 583-597.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist, 37,* 91-106.

Perkins, R. E., & Hill, A. B. (1985). Cognitive and affective aspects of boredom. *British Journal of Psychology, 76*, 221-234.

Perry, R. P. (1991). Perceived control in college students: Implications for instruction in higher education. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 7, pp. 1-56). New York, NY: Agathon.

Perry, R. P. (2003). Perceived (academic) control and causal thinking in achievement settings. *Canadian Psychologist, 44,* 312-331.

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. Schunk & J. Meece (Eds.), *Student perceptions in the classroom: Causes and consequences* (pp. 149-183). Hillsdale, NJ: Erlbaum.

Prensky, M. (2001). *Digital game-based learning.* New York, NY: McGraw-Hill.

Prensky, M. (2005). Computer games and learning: Digital game-based learning. *Handbook of Computer Game Studies, 18,* 97-122.

Project Tomorrow & PASCO Scientific. (2008). *Inspiring the next generation of innovators: Students, parents, and teachers speak up about science education.*

Irvine, CA: Project Tomorrow. Retrieved from http://www.tomorrow.org/SpeakUp/pdfs/Inspiring_the_next_generation_of_innovators.pdf

Quantum Spectre. (n.d.). Home page. Retrieved from http://edgeatterc.com/edge/games/quantum-spectre/

R Development Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report, Society for Research in Child Development, 16*, 3-18.

Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed.). New York, NY: Wiley.

Ricci, K. E., Salas, E., & Cannon-Bowers, J. A. (1996). Do computer-based games facilitate knowledge acquisition and retention? *Military Psychology, 8*, 295-307.

Rodrigo, M. M. T., & Baker, R. S. J. d. (2011). Comparing learners' affect while using an intelligent tutor and an educational game. *Research and Practice in Technology Enhanced Learning, 6*(1), 43-66.

Rodrigo, M. M. T., Baker, R. S. J. d., D'Mello, S., Gonzalez, M. C. T., Lagud, M. C., Lim, S. A., . . . Viehland, N. J. (2008). Comparing learners' affect while using an intelligent tutoring system and a simulation problem solving game. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems: 9th International Conference, ITS 2008, proceedings* (pp. 40-49). Berlin, Germany: Springer.

Rodrigo, M. M. T., Baker, R. S. J. d., Jadud, M. C., Amarra, A. C. M., Dy, T., Espejo-Lahoz, M. B. V., . . . Tabanao, E. S. (2009). Affective and behavioral predictors of novice programmer achievement. In *Proceedings of the 14th ACM-SIGCSE Annual Conference on Innovation and Technology in Computer Science Education* (pp. 156-160). Available from http://www.jadud.com/dl/pdf/2009-iticse-rodrigo.pdf

Rodrigo, M. M. T., Baker, R. S. J. d., & Nabos, J. Q. (2010). *The relationships between sequences of affective states and learner achievements.* In *Proceedings of the 18th International Conference on Computers in Education*. Available from http://www.columbia.edu/~rsb2162/ICCE2010-RBN .pdf

Romesburg, H. C. (1984). *Cluster analysis for researchers.* Belmont, CA: Lifetime Learning.

Rosenberg, E. (1998). Levels of analysis and the organization of affect. *Review of General Psychology, 2*, 247-270. doi:10.1037//1089-2680.2.3.247

Sabherwal, R., & Robey, D. (1993). An Empirical Taxonomy of Implementation Processes Based on Sequences of Events in Information Systems Development. *Organization Science, 4*(4), 548-576.

Sabourin, J., Rowe, J., Mott, B., & Lester, J. (2011). When off task in on-task: The affective role of off-task behavior in narrative-centered learning environments. In *Proceedings of the 15th International Conference on Artificial Intelligence in Education* (pp. 534-536). Springer Berlin Heidelberg.

San Pedro, M. O., Baker, R. S. J. d., Bowers, A. J., & Heffernan, N. T. (2013). Predicting college enrollment from student interaction with an intelligent tutoring system in middle school. In *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 177-184). Available from http://educationaldatamining.org/EDM2013/proceedings/paper_16.pdf

San Pedro, M. O., Baker, R., & Rodrigo, M. M. (2011). Detecting carelessness through contextual estimation of slip probabilities among students using an intelligent tutor for mathematics. In *Proceedings of the 15th International Conference on Artificial Intelligence in Education* (pp. 304-311). Available from http://www.columbia.edu/~rsb2162/SPBR2011-v2.pdf

Sanjeev, A. P., & Zytkow, J. M. (1995). Discovering enrollment knowledge in university databases. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining, Montreal, Canada* (pp. 246-251). Montreal, Quebec, Canada.

Savasere, A., Omiecinski, E., & Navathe, S. (1995). An efficient algorithm for mining association rules in large databases. In *Proceedings of the 21st International Conference on Very Large Databases: VLDB'95* (pp. 432-444). Zurich, Switzerland.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293-317). Hillsdale, NJ: Erlbaum.

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). *Motivation in education: Theory, research, and applications* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129-184.

Shen, R. M., Yang, F., & Han, P. (2003). A dynamic self-organizing e-learner communities with improved multi-agent matchmaking algorithm. *Lecture Notes in Computer Science, 2903,* 590-600.

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement.* New York, NY: Springer.

Shroff, R. H., & Vogel, D. R. (2009). Assessing the factors deemed to support individual student intrinsic motivation in technology supported online and face-to-face discussions. *Journal of Information Technology Education, 8*, 59-85.

Srikant, R., & Agrawal, R. (1996). Mining sequential patterns: Generalizations and performance improvements. *In Proceedings of the 5th International Conference on Extending Database Technology* (pp. 3-17). London, England: Springer-Verlag.

Squire, K. D. (2003). *Gameplay in context: Learning through participation in communities of Civilization III players* (Unpublished doctoral dissertation). Indiana University, Bloomington.

Squire, K. D. (2007). Open-ended video games: A model for developing learning for the interactive age. In K. Salen (Ed.), *The ecology of games: Connecting youth, games, and learning* (pp. 167-198). Cambridge, MA: The MIT Press.

Squire, K. D., & Barab, S. A. (2004). *Replaying history: Learning world history through playing Civilization III*. Bloomington, IN: Indiana University.

Stein, M. B., & Kean, Y. M. (2000). Disability and quality of life in social phobia: Epidemiologic findings. *American Journal of Psychiatry, 157*, 1606-1613.

Stein, N. L., & Levine, L. J. (1991). Making sense out of emotion. In W. Kessen, A. Ortony, & F. Kraik (Eds.), *Memories, thoughts, and emotions: Essays in honor of George Mandler* (pp. 295-322). Hillsdale, NJ: Erlbaum.

Steinkuehler, C., & Chmiel, M. (2006). Fostering scientific habits of mind in the context of online play. In S. A. Barab, K. E. Hay, N. B. Songer, & D. T. Hickey (Eds.), *Proceedings of the International Conference of the Learning Sciences* (pp. 723-729). Mahwah, NJ: Erlbaum.

Swayne, D. F., Cook, D., & Buja, A. (1998). XGobi: Interactive dynamic data visualization in the X Window System. *Journal of Computational and Graphical Statistics, 7(1),* 113–130.

Swayne, D. F., Lang, D. T., Buja, A., & Cook, D. (2003). GGobi: Evolving from XGobi into an extensible framework for interactive data visualization. *Journal of Computational Statistics and Data Analysis, 43*(4), 423-444.

TraMinerR. (2014). *TraMinerR: Sequence analysis in R.* Retrieved from http://mephisto .unige.ch/traminer

Tryon, R. C. (1939). *Cluster analysis: Correlation profile and orthometric (factor) analysis for the isolation of unities in mind and personality*. Ann Arbor, MI: University of Michigan.

Turner, J. E., & Schallert, D. L. (2001). Expectancy-value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology, 93,* 320-329.

Ulrich, D., & McKelvey, B. (1990). General Organizational Classification: An Empirical Test Using the United States and Japanese Electronic Industry. *Organization Science, 1*(1), 99-118.

Utah State Office of Education. (2002). *Core Standards for Science.* Retrieved from http://schools.utah.gov/CURR/science/Core-Standards/UT-3-6-Science-Core -Standards.aspx

VanLehn, K. (1999). Rule learning events in the acquisition of a complex skill: An evaluation of cascade. *The Journal of the Learning Sciences, 8*(1), 71-125.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction, 21*, 209-249.

Virvou, M., & Katsionis, G. (2006). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education, 50,* 154-178.

Wang, F.-H. (2002). On using data-mining technology for browsing log file analysis in asynchronous learning environment. In P. Barker & S. Rebelsky (Eds.), *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 2005-2006). Chesapeake, VA: Association for the Advancement of Computing in Education.

Wang, W., Weng, J., Su, J., & Tseng, S. (2004, October). *Learning portfolio analysis and mining in SCORM compliant environment.* Paper presented at the 34th ASEE/IEEE Frontiers in Education Conference, Savannah, GA.

Waraich, A. (2004, June). *Using narrative as a motivating device to teach binary arithmetic and logic gates.* Paper presented at the 9th annual SIGCSE Conference on Innovation and Technology in Computer Science Education, Leeds, England.

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236–244.

Whitebread, D. (1997). Developing children's problem-solving: The educational uses of adventure games. In A. McFarlane (Ed.), *Information technology and authentic learning* (pp. 13-37). London, England: Routledge.

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R., & Davis-Kean, P. (2006). Development of achievement motivation. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology*: *Vol. 3. Social, emotional, personality development* (6th ed., pp. 933-1002). New York, NY: Wiley.

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., Mcrae, A., & Barbosa, P. (2008). Role of reading engagement in mediating the effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools, 45,* 432-445.

Wighting, M. J., Liu, J., & Rovai, A. P. (2008). Distinguishing sense of community and motivation characteristics between online and traditional college students. *Quarterly Review of Distance Education, 9*, 285-295.

Wu, L. L. (2000). Some comments on "Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect." *Sociological Methods & Research, 29,* 41-64.

Zaïane, O. R., & Luo, J. (2001). Towards evaluating learners' behaviour in a web-based distance learning environment. In *Proceedings of the IEEE International Conference on Advanced Learning Technologies* (pp. 357-360).

Zaïane, O., Xin, M., & Han, J. (1998). Discovering web access patterns and trends by applying OLAP and data mining technology on web logs. In *ADL '98: Proceedings of the Advances in Digital Libraries Conference* (pp. 19-29). Washington, DC: IEEE Computer Society.

Zaki, M. J. (2001). SPADE: An efficient algorithm for mining frequent sequences. *Machine Learning, 42*(1-2), 31-60.

Zeidner, M. (1998). *Test anxiety: The state of the art.* New York, NY: Plenum.

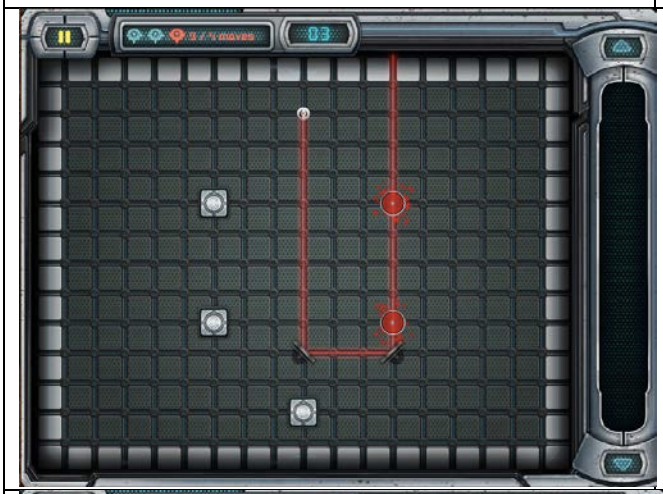
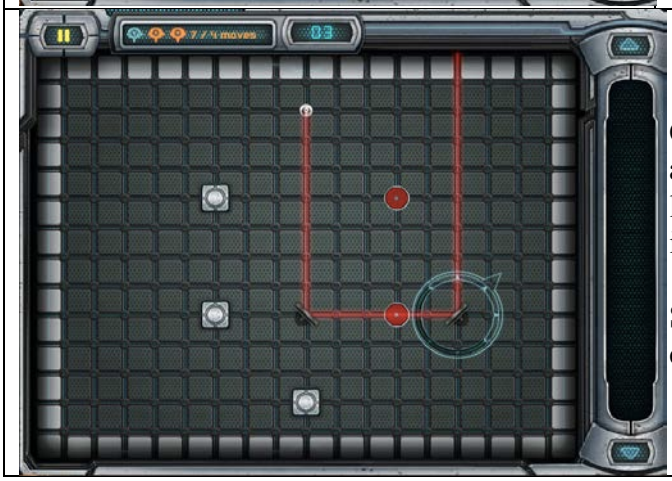
Zeidner, M. (2007). Test anxiety in educational contexts: Concept, findings, and future directions. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 165-184). San Diego, CA: Elsevier Academic Press.

Zhou, M., Xu, Y., Nesbit, J. C., & Winne, P. H. (2010). Sequential pattern analysis of learning logs: Methodology and applications. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. C. de Baker (Eds.) *Handbook of educational data mining* (pp. 107-121). Boca Raton, FL: Taylor & Francis.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). Cambridge, England: Cambridge University Press.

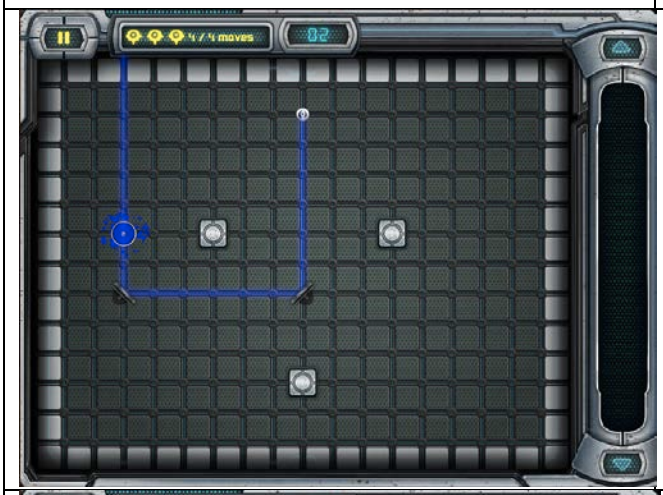
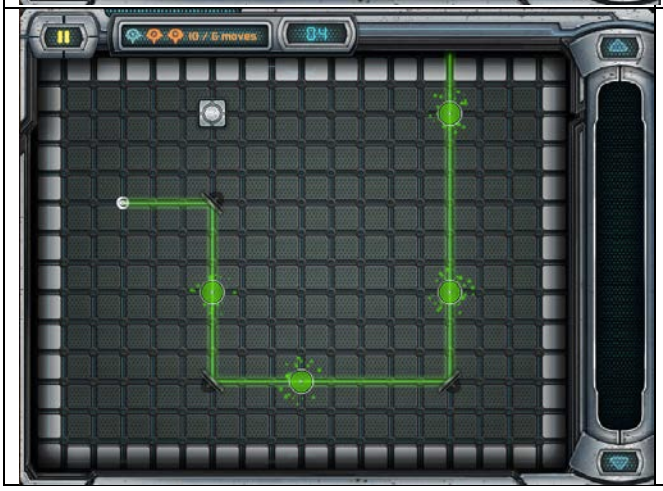
APPENDICES

Appendix A. Game-Level Visualizations

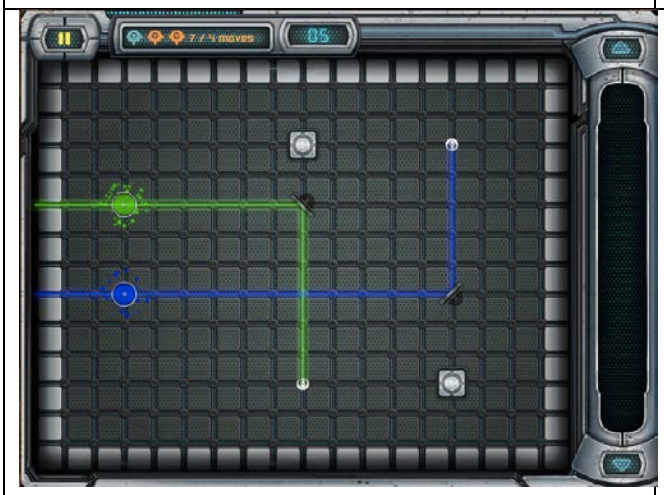
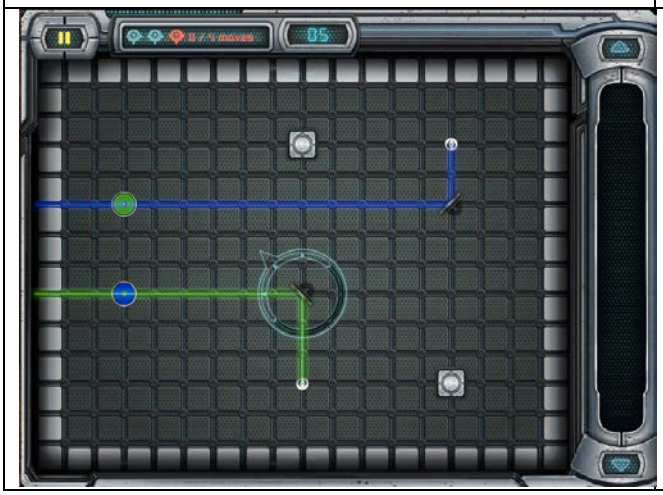
Level visualization	Description
	Quantum Spectre level with one laser and one target of the same color as the laser. Student leads the laser into both targets. This leads to success.[1] Student is given two mirrors to conquer this level.
	Quantum Spectre level with one laser and one target of the same color as the laser. Student sets up the mirror pieces incorrectly. This leads to failure.[2] Student is given two mirrors to conquer this level.

[1] In this game, when the student correctly reaches all the targets, the targets start sparkling and the level changes by taking the student to the next level.

[2] In this game, wrong actions such as the ones that lead to failure, do not stop or restart the game. Instead, the student is given time to find the right approach to overcome the level.

Level visualization	Description
	Quantum Spectre level with one laser and one target of the same color as the laser. Student leads the laser around the barrier into the target, thus, succeeding the level. Student is given two mirrors to conquer this level.
	Quantum Spectre level with one laser and one target of the same color as the laser. Student leads the laser into a barrier, which blocks the laser from reaching the target. This leads to failure. Student is given two mirrors to conquer this level.
	Quantum Spectre level with one laser and four targets of the same color as the laser. Student is directing the laser through all the targets. This leads to success.[3] Student is given three mirrors to conquer this level.

[3] When the student correctly reaches all the targets, the targets start sparkling and the level changes by taking the student to the next level.

Level visualization	Description
	Quantum Spectre level with two lasers and two targets (one for each color of laser). Student is correctly directing the lasers into appropriate targets (color coding). This leads to success. Student is given two mirrors to conquer this level.
	Quantum Spectre level with two lasers and two targets (one for each color of laser. Student is directing the lasers into the inappropriate targets (color coding). This leads to failure. Student is given two mirrors to conquer this level.

Appendix B. Features of the Game Data: Single Line per Event

Feature	Specific features
	Within each game-level round
F.1: Numbers of each type of object placed in the game grid so far	F1.1: Number of lasers of each color type
	F1.2: Number of single color targets of each type
	F1.3: Number of dual color targets of each type
	F1.4: Number of tri-color targets
	F1.5: Number of filters of each type
	F1.6: Number of barriers
	F1.7: Number of flat mirrors, single sided
	F1.8: Number of flat mirrors, double sided
	F1.23: Number of convex lenses—double focal points, focal length 2
	F1.24: Number of convex lenses—double focal points, focal length 3
F.2: Total number of moves so far	F2.1: As percentage of optimal number of moves
	F2.2: Number of location changes of single-sided flat mirrors after initial placement
	F2.3: Number of rotations of single-sided flat mirrors after initial placement
	F2.4: Number of location changes of double-sided flat mirrors after initial placement
	F2.5: Number of rotations of double-sided flat mirrors after initial placement
	F2.6: Number of location changes of convex mirrors after initial placement
	F2.7: Number of rotations of convex flat mirrors after initial placement
	F2.8: Number of location changes of concave mirrors after initial placement
	F2.9: Number of rotations of concave flat mirrors after initial placement
	F2.10: Number of location changes of double-sided convex/concave mirrors after initial placement
	F2.11: Number of rotations of double-sided

Feature	Specific features
	convex/concave flat mirrors after initial placement
	F2.12: Percentage of locations of mirrors in optimal placements & orientations
	F2.13: Number of location changes of beam splitter after initial placement
	F2.14: Number of rotations of beam splitter after initial placement
	F2.15: Number of location changes of concave lenses after initial placement
	F2.16: Number of rotations of concave lenses after initial placement
	F2.17: Number of location changes of convex lenses after initial placement
	F2.18: Number of rotations of convex lenses after initial placement
	F2.19: Percentage of locations of lenses in optimal placements & orientations
	F2.20: Total number of location changes of mirrors after initial placement, divided by num objects
	F2.21: Total number of rotations of mirrors after initial placement, divided by num objects
	F2.22: F2.20 divided by F2.21 (0 when div/0 would occur)
	F2.23: Percentage of total number of actions that are placements
	F2.24: Percentage of total number of actions that are location changes
	F2.25: Percentage of total number of actions that are rotations
	F2.26: Max number of moves for any specific object
	F2.27: Max number of rotations for any specific object
	F2.28: Specific object being moved or rotated (FlatMirror1, FlatMirror2, ConvexMirror2, etc.)

Feature	Specific features
F3: Game outcome (Solution Found within optimal number moves, Solution Found using > number optimal moves, Restart, Quit)	F3.1: Successful solution? Yes/No
	F3.2: Successful solution within optimal number of moves? Yes/No
	F3.3: Number of moves beyond optimal
	F3.4: Rating: 1, 2, or 3 “stars”/spectre
	F3.5: Percentage of maximum rating possible (Rating/[3*number levels played])
	F3.6: F3.1.3 divided by optimal number of moves
F4: Time elapsed since the start of the round	F4.1: Time since last click
	F4.2: Time since last game object removed from inventory
	F4.3: Time since last location change
	F4.4: Time since last rotation
	F4.5: Time to first move (moving something from inventory to game grid)
	F4.6: Time since last move (moving/rotating something in game grid or moving something to/from inventory to game grid)
	F4.7: Average time taken for location changes so far
	F4.8: Average time taken for rotations so far
Within game level (across all rounds played so far)	
F5: Number of times played this level so far	F5.1: Number of restarts so far
	F5.2: number of rounds completed to 1-star status
	F5.3: number of rounds completed to 2-star status
	F5.4: number of rounds completed to 3-star status: perfect solution
	F5.5: number of rounds played without completing
F6: Number of restarts before first successful solution	
F7: Number of quits so far	
F8: Number of quits before first successful solution	
F9: Number of successful solutions so far	

Feature	Specific features
F10: Total time spent playing this level so far	F10.1: Average duration of rounds so far
F11: Average number moves per round so far	F11.1: As percentage of optimal number moves
	F11.2: Average number location changes of mirrors per round so far
	F11.3: Average number rotations of mirrors per round so far
	F11.4: Average number location changes of lenses per round so far
	F11.5: Average number rotations of mirrors per round so far
	Overall game
F12: Highest game level reached	F12.1: Number of game levels played to 1-star status (as highest level)
	F12.2: Number of game levels played to 2-star status (as highest level)
	F12.3: Number of game levels played to 3-star status (as highest level)
F13: Highest game level attempted	
F14: Number of rounds played (regardless of game level)	
F15: Total number restarts	
F16: Total number quits	
F17: Total number levels successfully completed on first attempt	
F18: Total number sessions	
F19: Consistent with optimal locations & orientations of each object in the game space?	
F20: Total time played	

Appendix C. The Output Format of the Game Logfile

Column name	Column description
ID	Unique ID for each log row in the database.
Tag	Optional string to allow for easier filtering of events. For example, if you're capturing data at a school called Flapjack and you have 2 different setups of the game, you could put "Flapjack1" and "Flapjack2" into here and then specify this in the DBCT website when creating your filters. This would need to be set up prior to the event in a server-side config file. You can also set the tag for your session by supplying it on the URL in the same way as you can specify your PlayerID.
PlayerID	Use the same scheme as Impulse for now; at some point we would use a BrainPlay login reference of sorts.
SessionID	Unique 64-bit number representing the session.
SessionLogID	Unique sequential number starting at 0 for each session. This determines the true sequence of events, as timestamps are not reliable for this.
Timestamp	UNIX-style, seconds since 1970 + milli/micro second fraction.
Level Time	Starts at 0 on the Level Start Event and counts upwards only while the game is actually being played. Measured in seconds.
Tutorial Mode	Integer, 0 = off, 1 = on
Event Type	N/A
Data 1–6	Six generic integer columns, usage is specific to Event Type

Note. Assembled from the information provided by EdGE at TERC in regards to their Quantum Spectre game.

Appendix D. Game Log Event Types

Event type	Description
Level Start	Data1 = Total Score/Rating at level start Data2 = Good Moves Data3 = Optimal Moves
Level End	Data1 = Reason (0 = Success, 1 = Restart, 2 = Quit) Data2 = Rating (0 = No rating,1 = 1 Spectre icon, 2 = 2 Spectre icon, 3 = 3 Spectre icon) Data3 = Total number of moves taken in level
Actor Status	Data1 = ActorID of the actor whose status is being specified Data2 = Grid X (null if inventory) Data3 = Grid Y (null if inventory) Data4 = Rotation (null if inventory). Zero degrees is down/south and it rotates clockwise (0 = down/south, 90 = left/west, 180 = up/north and 270 = right/east). Data5 = Count. Only applies to inventory items, used for setting up levels.
Move	Data1 = ActorID of the actor being moved Data2 = Previous Grid X (null if inventory) Data3 = Previous Grid Y (null if inventory) Data4 = Grid X (null if inventory) Data5 = Grid Y (null if inventory) Data6 = Rotation (null if inventory)
Rotate	Data1 = ActorID of the actor being rotated Data2 = Grid X (null if inventory) Data3 = Grid Y (null if inventory) Data4 = Initial rotation Data5 = Final rotation
Tutorial	This needs to be sent when a tutorial popup is shown and again when it is hidden. Data1 = Visible (0 = No, 1 = Yes)

Note. Assembled from the information provided by EdGE at TERC in regards to their Quantum Spectre game.

Appendix E. HART Tool Data by Session and by Classroom

Session	Classroom 1	Classroom 2
Session 1	First Set: 1, 2, 3 Second Set: 4, 5, 6 Third Set: 7, 8, 9 Fourth Set: 10, 11, 12	First Set: 1, 2, 3 (27, 28, 29) Second Set: 4, 5, 6 (30, 31, 32) Third Set: 7, 8, 9 (33, 34, 35) Fourth Set: 10, 11, 12 (36, 37, 38) Fifth Set: 13, 14, 15 (39, 40, 41)
Session 2	First Set: 2, 3, 4 Second Set: 5, 6, 7 Third Set: 8, 9, 10 Fourth Set: 13, 14, 15 Fifth Set: 16, 17, 18 Sixth Set: 19, 20, 21 Seventh: 22, 23, 24	First Set: 19, 20 (45, 46) Second Set: 22, 23, 24 (48, 49, 50) Third Set: 1, 2, 3 (27, 28, 29) Fourth Set: 4, 5, 6 (30, 31, 32)
Session 3	First Set: 1, 3, 4 Second Set: 5, 6, 7 Third Set: 8, 10, 13 Fourth Set: 14, 15, 16 Fifth Set: 17, 18, 21	First Set: 1, 2, 3 (27, 28, 29) Second Set: 5, 6, 7 (31, 32, 33) Third Set: 8, 9, 10 (34, 35, 36) Fourth Set: 11, 12, 13 (37, 38, 39) Fifth Set: 14, 16, 17 (40, 42, 45)
Session 4	First Set: 1, 2, 3 Second Set: 4, 5, 6 Third Set: 7, 8, 9 Fourth Set: 10, 13, 14 Fifth Set: 15, 16, 17 Sixth: 18, 20, 21	First Set: 1, 2, 3 (27, 28, 29) Second Set: 4, 5, 6 (30, 31, 32) Third Set: 7, 8, 9 (33, 34, 35) Fourth Set: 10, 12, 13 (36, 38, 39) Fifth Set: 14, 17, 26 (40, 43, 52) Sixth Set: 16, 18, 19 (42, 44, 45)
Session 5	First Set: 26 Second Set: 22, 24, 25 Third Set: 18, 19, 20 Fourth Set: 15, 16, 17 Fifth Set: 10, 13, 14	First Set: 23, 24, 25 (49, 50, 51) Second Set: 20, 21, 26 (46, 48, 52) Third Set: 17, 18, 19 (43, 44, 45) Fourth Set: 13, 14, 16 (39, 40, 42) Fifth Set: 8, 10, 12 (34, 36, 38) Sixth Set: 6, 7, 9 (32, 33, 35) Seventh Set: 1, 2 (27, 28)

Session 6	First Set: 1, 2, 3	First Set: 1, 2, 3 (27, 28, 29)
	Second Set: 8, 9, 10	Second Set: 4, 5, 6 (30, 31, 32)
	Third Set: 13, 14, 15	Third Set: 7, 8, 9 (33, 34, 36)
	Fourth Set: 16, 17, 18	Fourth Set: 11, 12, 13 (37, 38, 39)
	Fifth Set: 19, 20, 21	Fifth Set: 14, 16, 26 (40, 42, 52)
Session 7	First Set: 20, 21, 22	First Set: 11, 24, 25 (37, 50, 51)
	Second Set: 16, 17, 18	Second Set: 22, 23 (48, 49)
	Third Set: 13, 14, 15	Third Set: 18, 19, 20 (44, 45, 46)
	Fourth Set: 9, 10, 12	Fourth Set: 15, 16, 17 (41, 42, 43)
		Fifth Set: 12, 13, 14 (38, 39, 40)
Session 8	First Set: 2, 3, 4	First Set: 1, 2, 3 (27, 28, 29)
	Second Set: 5, 6, 7	Second Set: 4, 5, 6 (30, 31, 32)
	Third Set: 8, 9, 10	Third Set: 7, 8, 9 (33, 34, 35)
	Fourth Set: 12, 13, 14	Fourth Set: 10, 12, 13 (36, 38, 39)
		Fifth Set: 14, 15, 16 (40, 41, 42)
Session 9	First Set: 1, 2, 3	First Set: 1, 2, 3 (27, 28, 29)
	Second Set: 4, 5, 6	Second Set: 4, 6, 7 (30, 32, 33)
	Third Set: 7, 8, 9	Third Set: 10, 13, 14 (36, 39, 40)
	Fourth Set: 10, 12, 13	Fourth Set: 15, 18, 19 (41, 44, 45)
	Fifth Set: 16, 17, 18	Fifth Set: 20, 22 (46, 48)

Note. This table reflects the raw data outputted from HART application. For Classroom 1 both HART application IDs and student user IDs were identical. For Classroom 2, the HART application ID restarted from 1, while the student IDs continued from 27 (as a continuation from Classroom 1 IDs). Therefore, in this Table I provide both the HART application IDs and the student user IDs if the latter is different from the former (see Classroom 2). In addition, in Sessions 2, 5, 7, and 9, Classroom 2 user 21(47) was deleted due to not consenting after the data collection.

Appendix F. Examples of a Gameplay Data Export

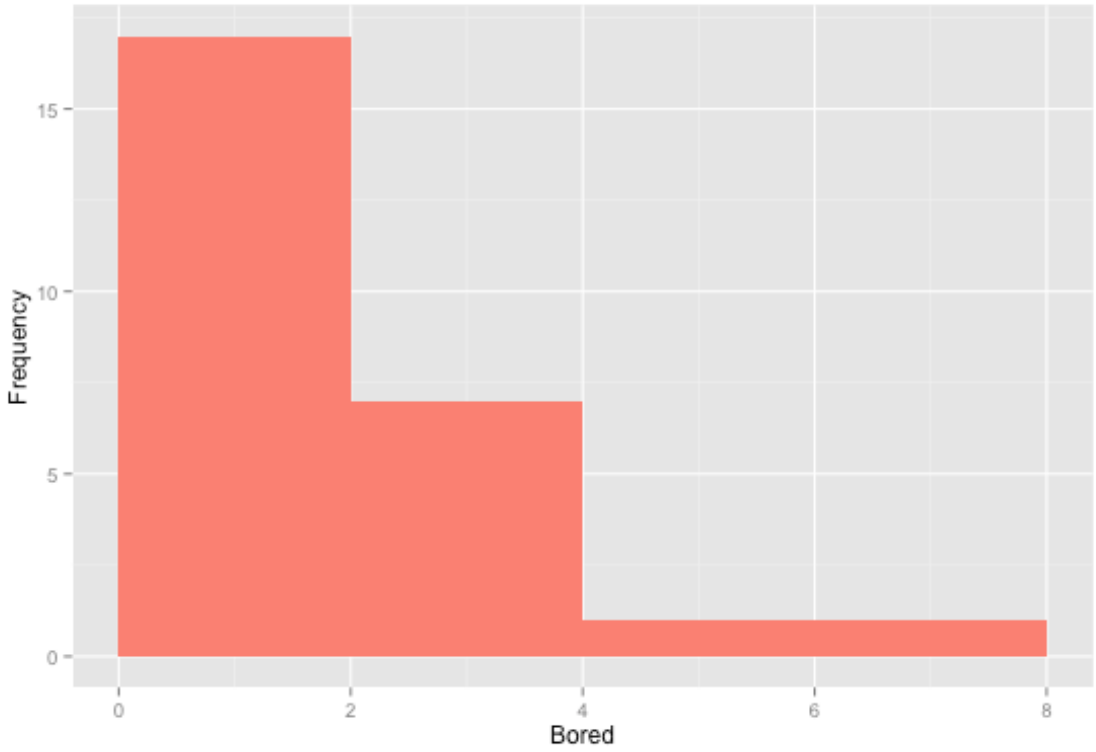
Row	Platform	Anonymous	ClassID	PlayerID	SessionID	EventID	Event	Time	Timestamp
0	iOS	0	57	5	X8984	1	Level Start	51:39.3	1384437099
1	iOS	0	57	5	X8984	6	Move	51:44.2	1384437104
2	iOS	0	57	5	X9073	1	Level Start	14:34.2	1384452874
3	Web	0	57	5	X9074	1	Level Start	14:50.5	1384452891
4	Web	0	57	5	X9074	6	Move	15:09.5	1384452910
5	Web	0	57	5	X9074	7	Rotate	15:12.4	1384452912
6	Web	0	57	5	X9074	6	Move	15:23.8	1384452924
7	Web	0	57	5	X9074	6	Move	15:49.3	1384452949
8	Web	0	57	5	X9074	6	Move	15:53.7	1384452954

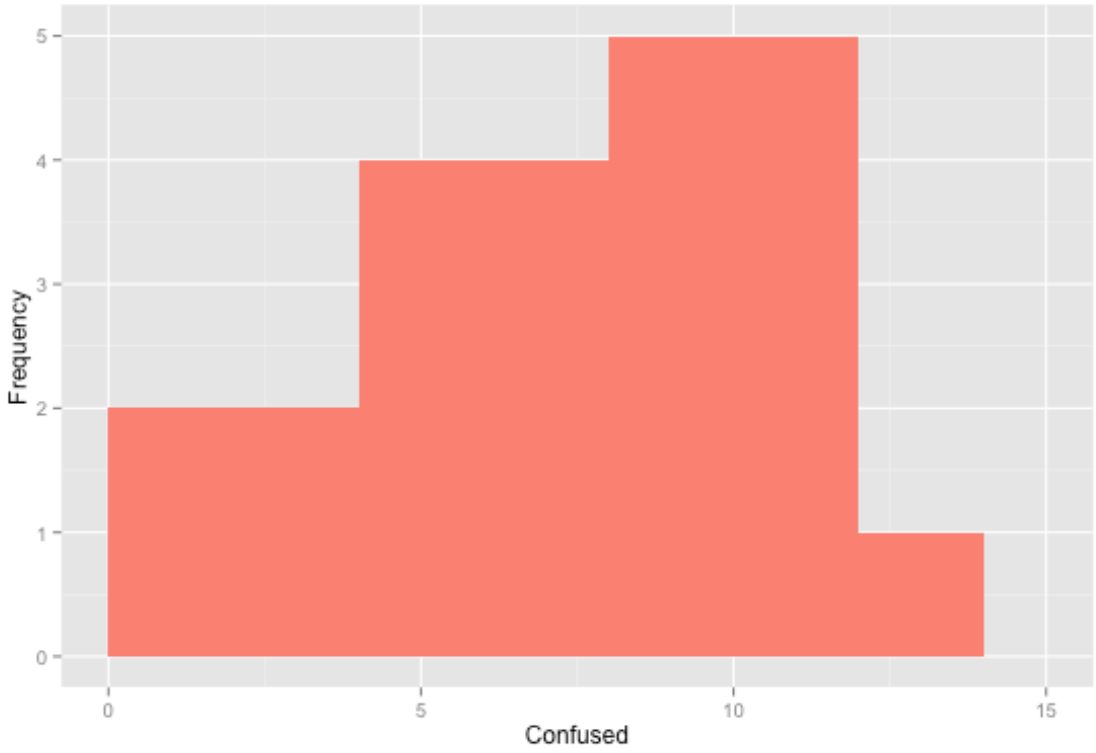
GameZone	Current_Game_Level	Tutorial	Total_Moves	Optimal_Moves	Good_Moves	Level_Outcome
3	9	1	.	9	17	Quit
3	9	1	.	9	17	Quit
3	9	1	.	9	17	Quit
3	9	1	11	9	17	Success
3	9	1	11	9	17	Success
3	9	1	11	9	17	Success
3	9	1	11	9	17	Success
3	9	1	11	9	17	Success
3	9	1	11	9	17	Success

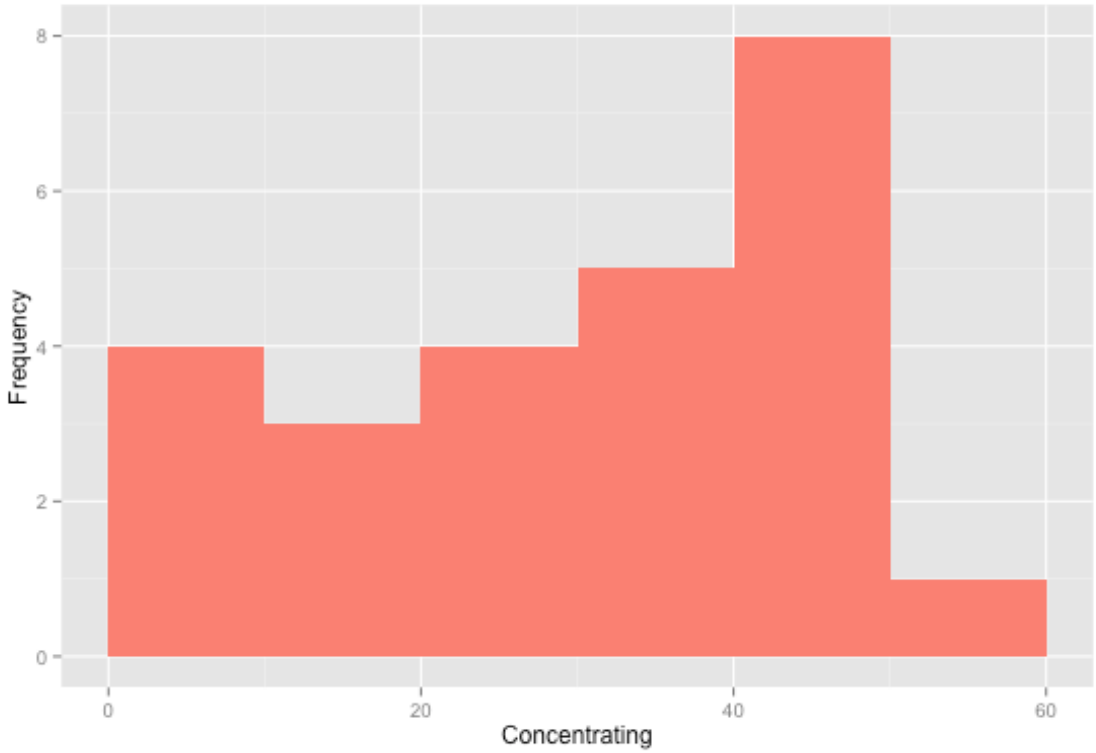
Note. Assembled from the information provided by EdGE at TERC in regards to their Quantum Spectre game.

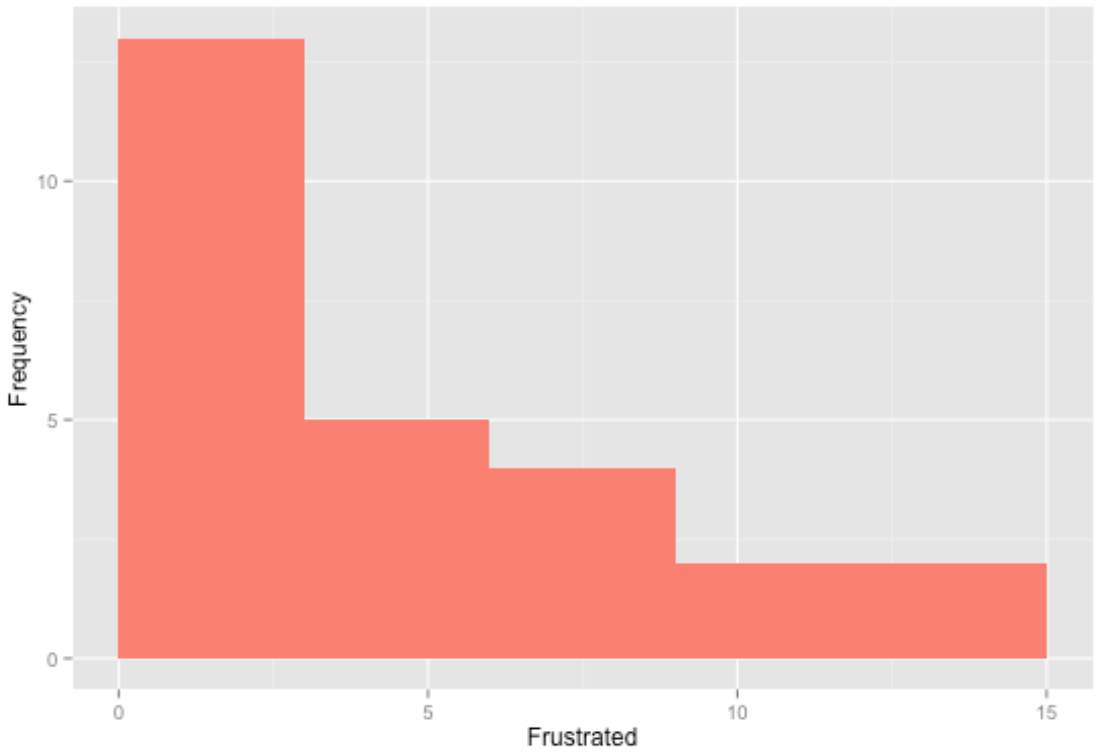
Appendix G. Classroom 1: Affect Variables' Distributions

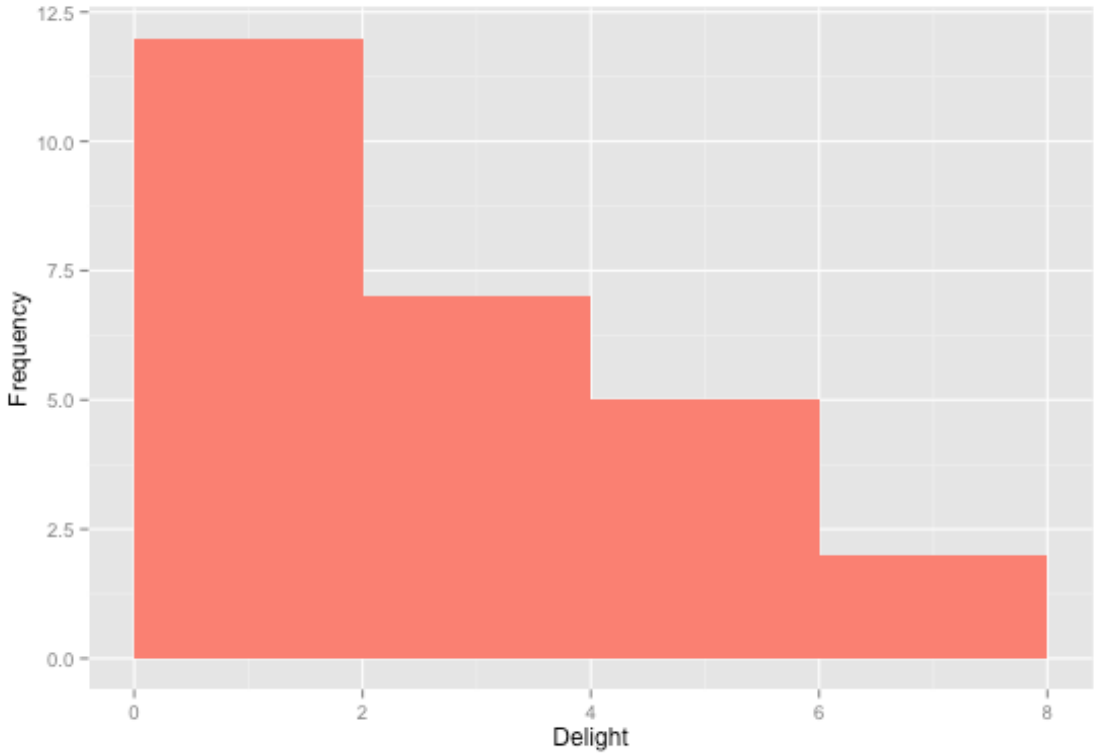
Below histograms represent the Classroom 1 and how many times each of the observed students experiences each of the affect categories. These histograms depict the occurrence of each of the categories in a certain range of values. Y-axis represents the students in the classroom and the area of the block represents the sum of the frequencies that are within that particular range of values for student affect. For example, first histogram showcases that 13 students in Classroom 1 were bored between 0 to 2 times throughout their observations, while 1 student was bored between 6 to 8 times.

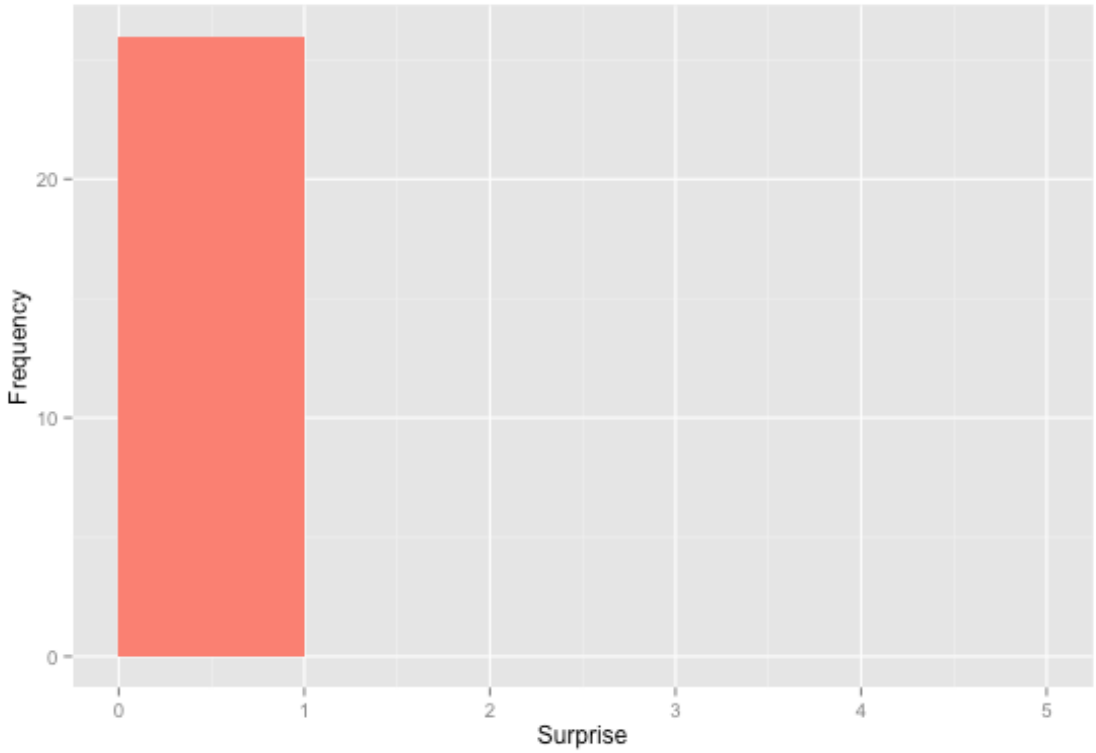


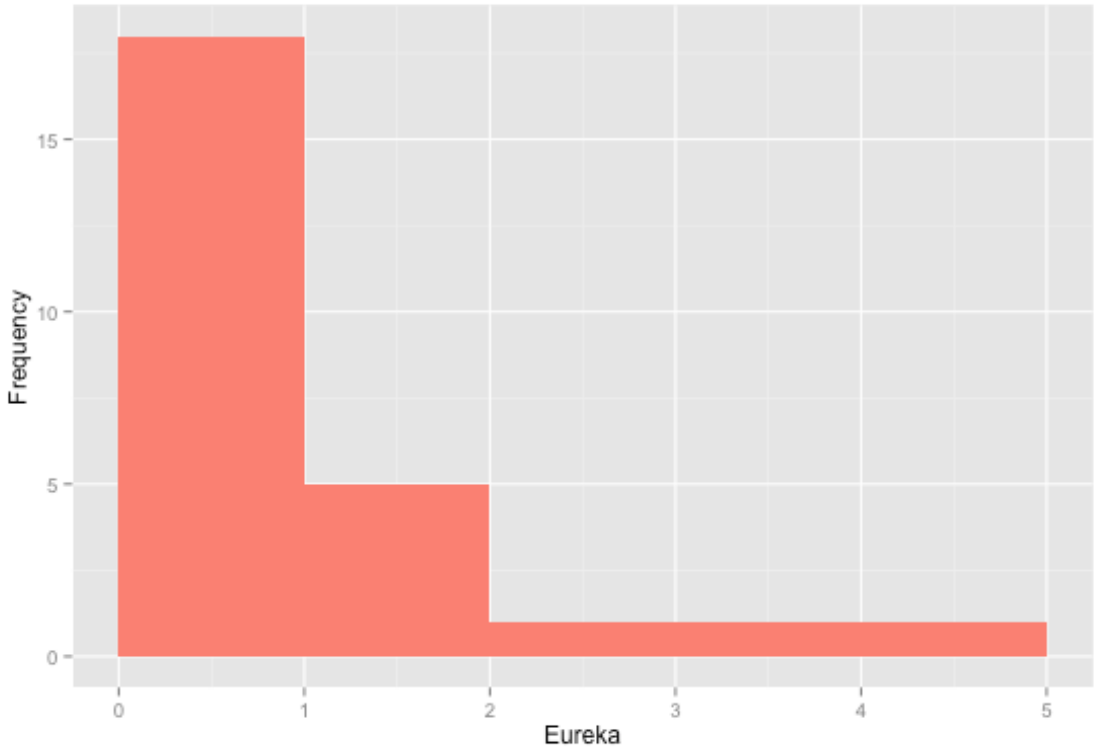

Frequency
5
4
3
2
1
0
0
5
10
15
Confused


8
6
4
2
0
Frequency
0
20
40
60
Concentrating


Frequency
10
5
0
0
5
10
15
Frustrated

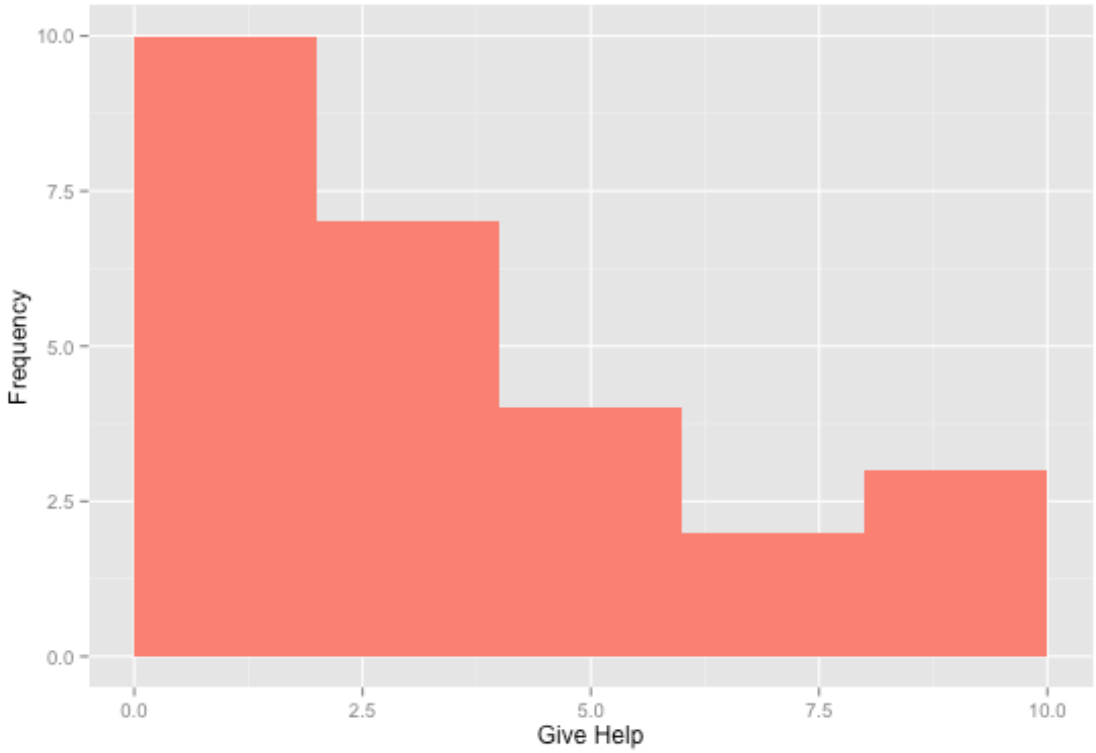

12.5
10.0
7.5
5.0
2.5
0.0
Frequency
0
2
4
6
8
Delight

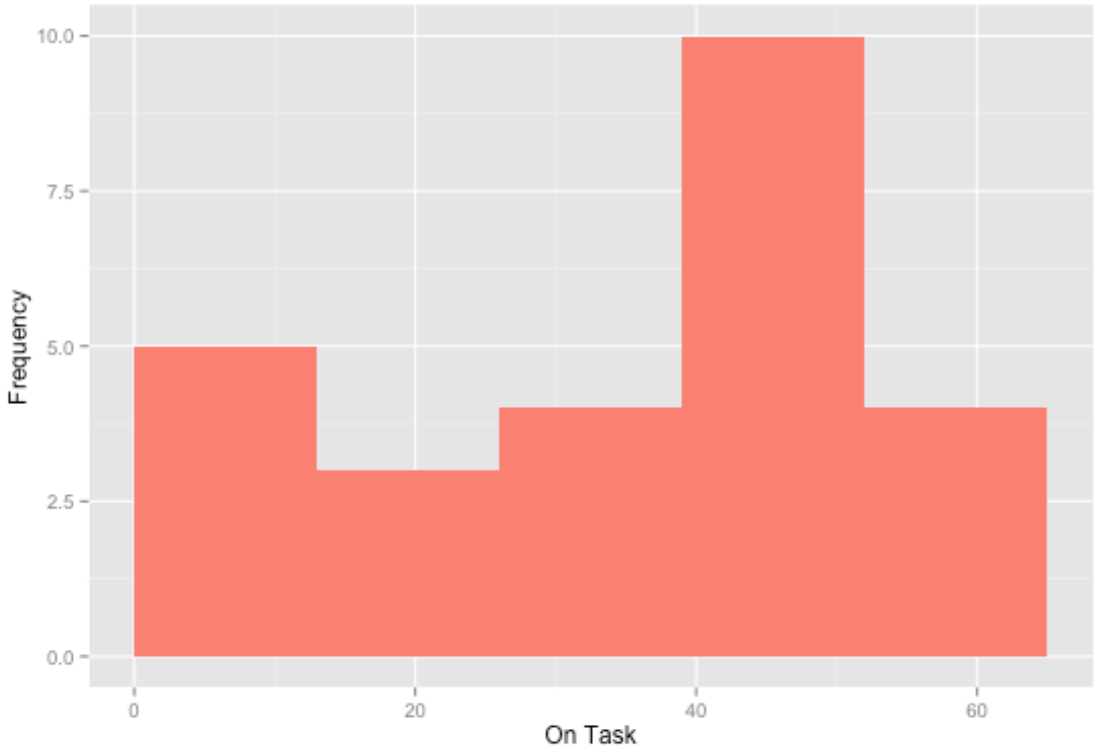

Frequency
20
10
0
0
1
2
3
4
5
Surprise

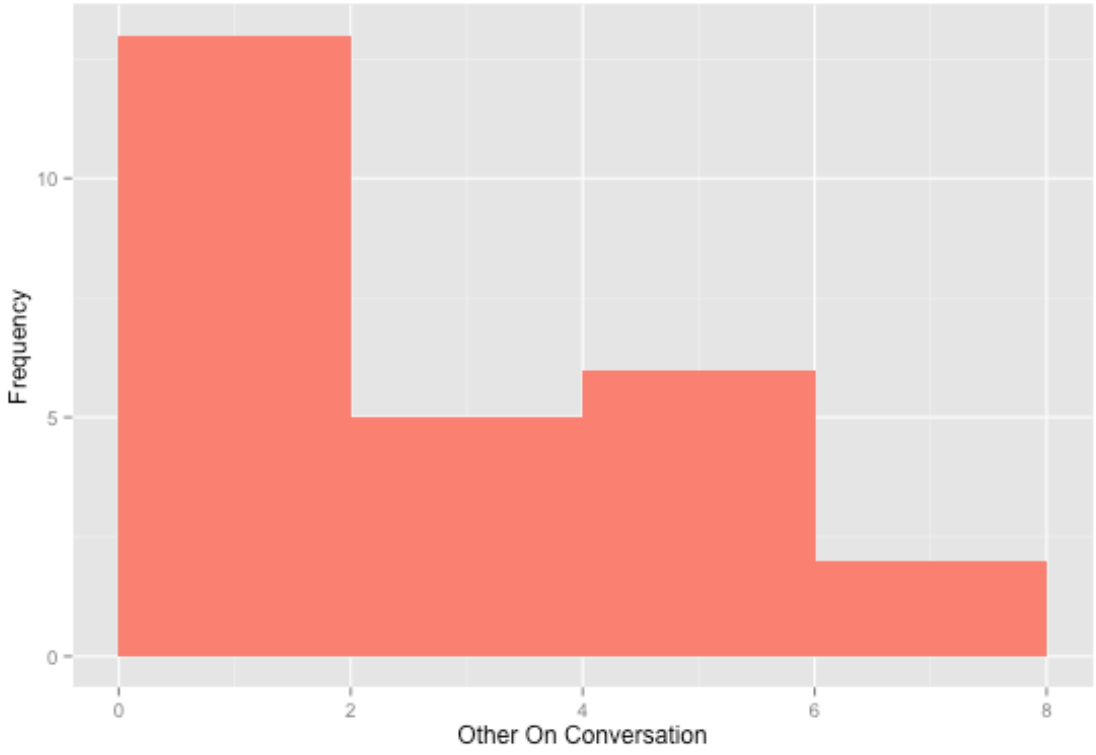

Frequency
15
10
5
0
0
1
2
3
4
5
Eureka

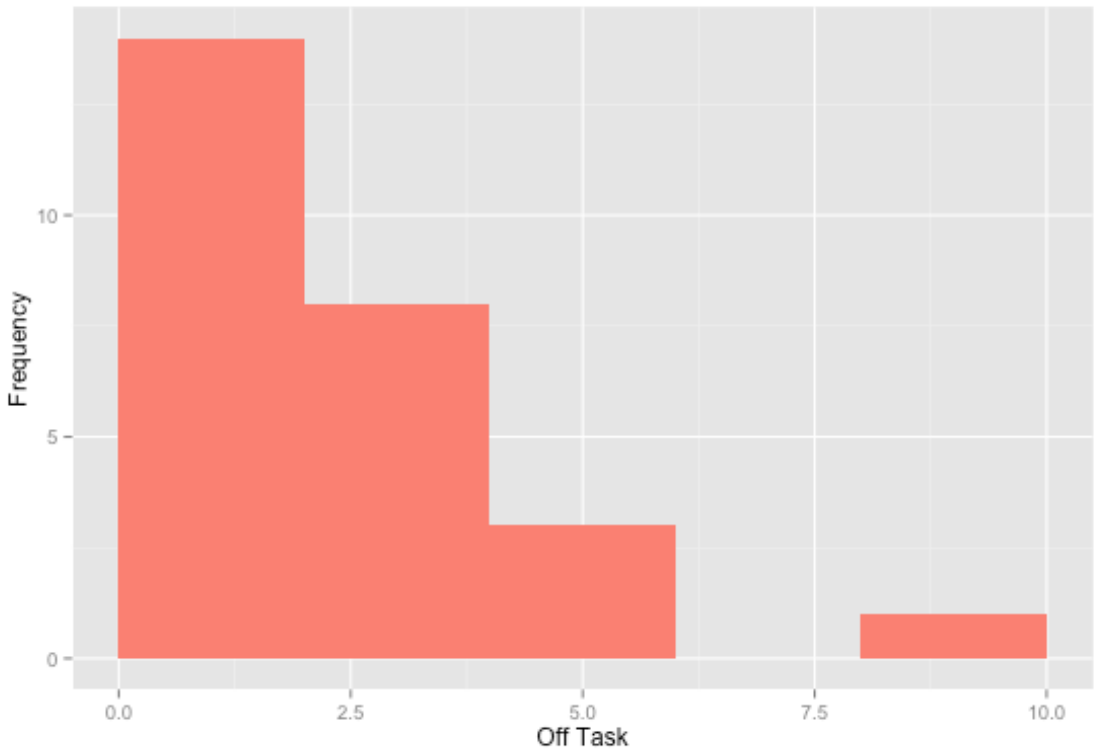
Appendix H. Classroom 1: Behavior Variables' Distributions

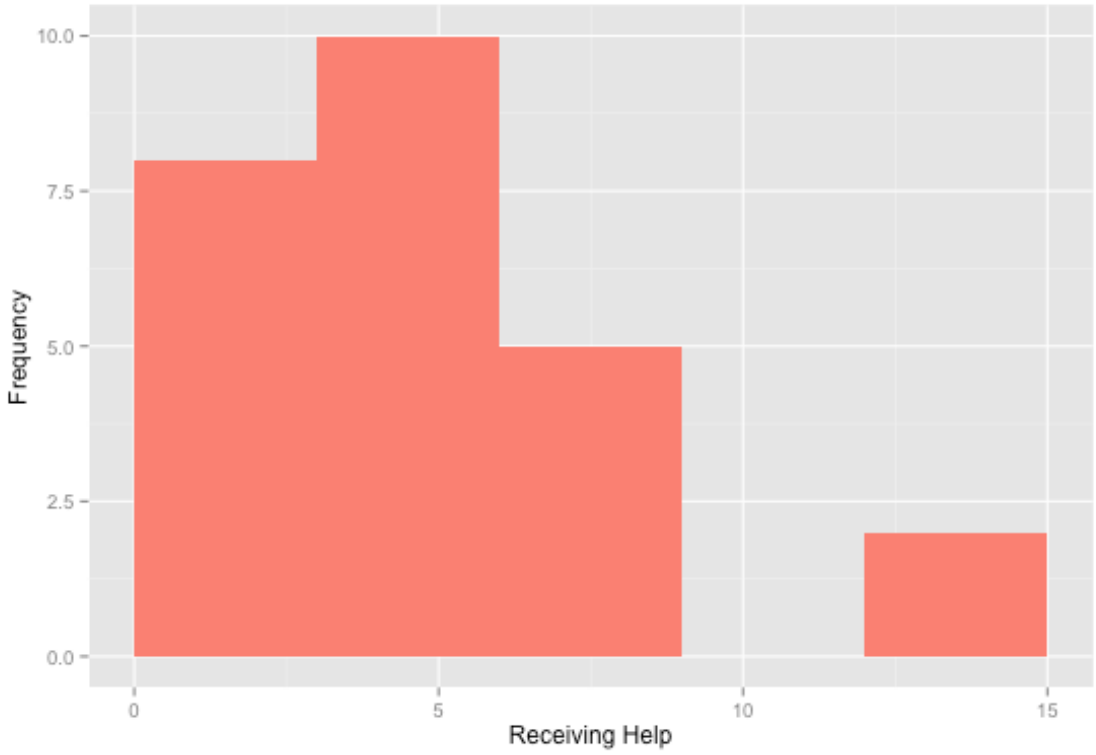
Below histograms represent the Classroom 1 and how many times each of the observed students experiences each of the behavior categories. Once again, Y-axis represents the students in the classroom and the area of the block represents the sum of the frequencies that are within that particular range of values for student behavior. For example, first histogram showcases that 10 students in Classroom 1 were giving help to their peers between 0 to 2 times throughout their observations, while 2 student were giving help between 6 to 8 times.




10.0
7.5
5.0
2.5
0.0
Frequency
0
20
40
60
On Task

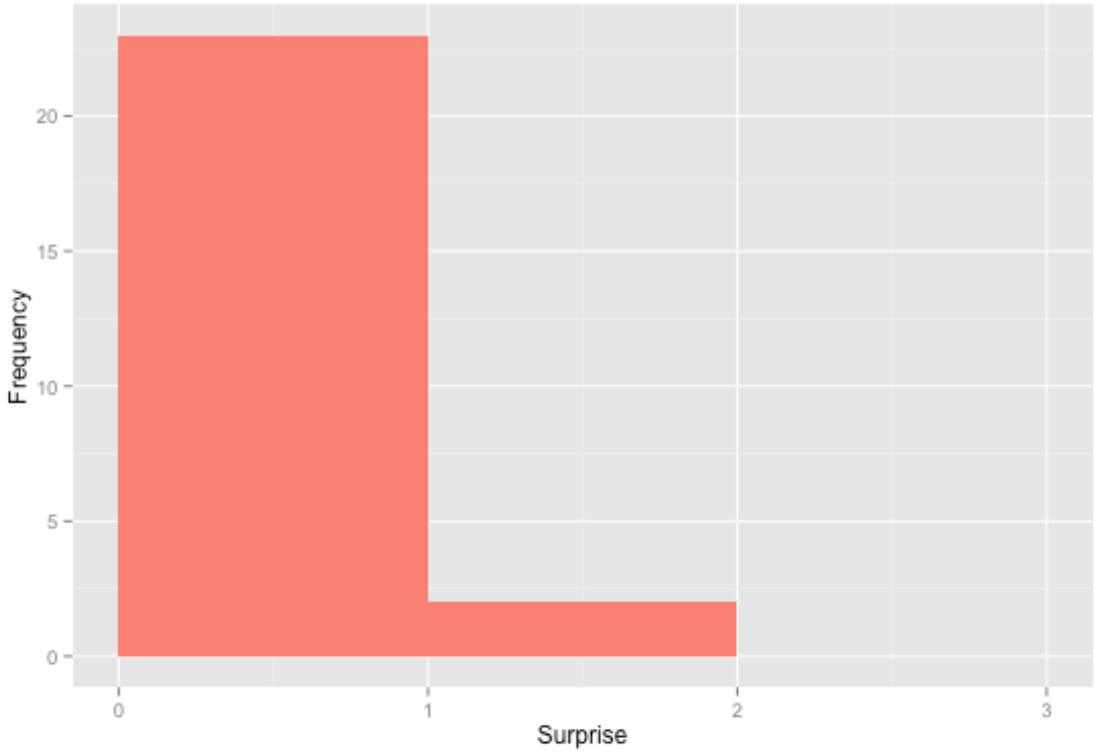

10
5
0
Frequency
0
2
4
6
8
Other On Conversation

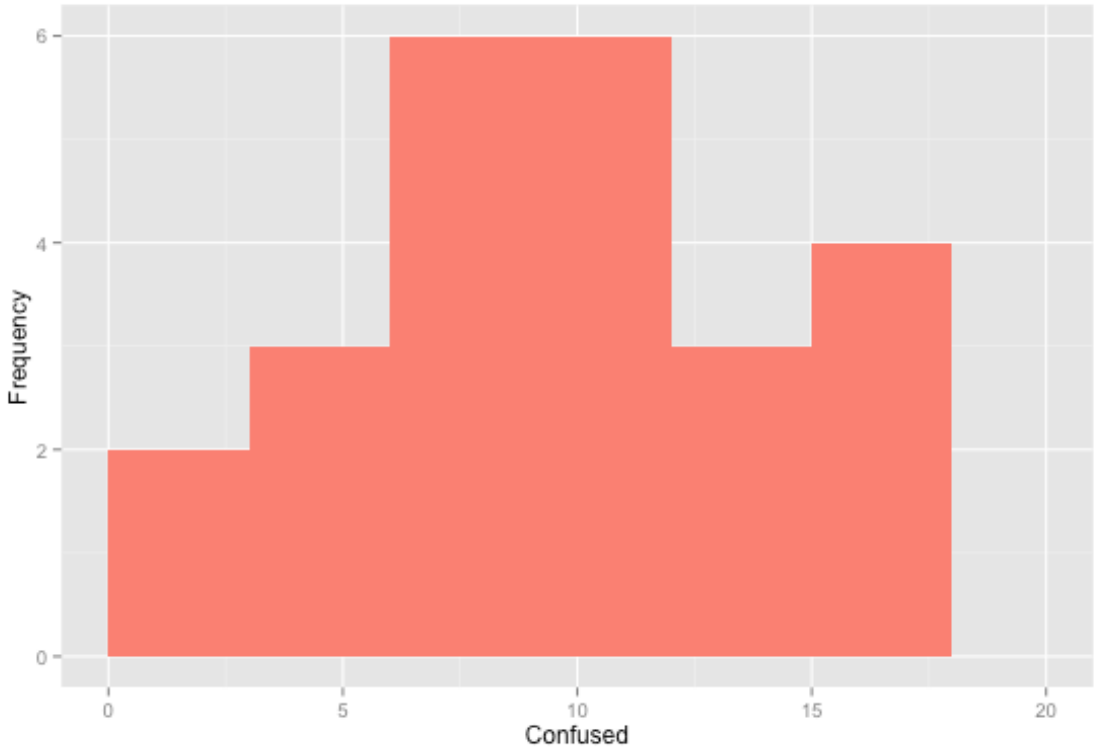

Frequency
10
5
0
0.0
2.5
5.0
7.5
10.0
Off Task

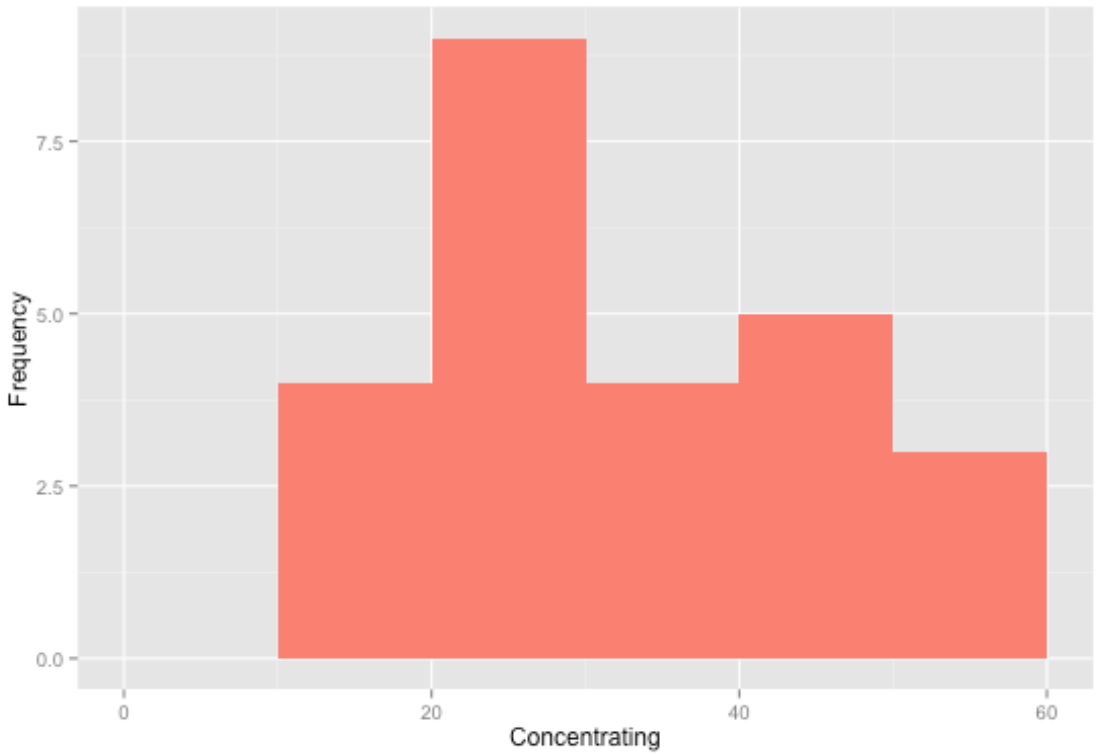

10.0
7.5
5.0
2.5
0.0
Frequency
0
5
10
15
Receiving Help

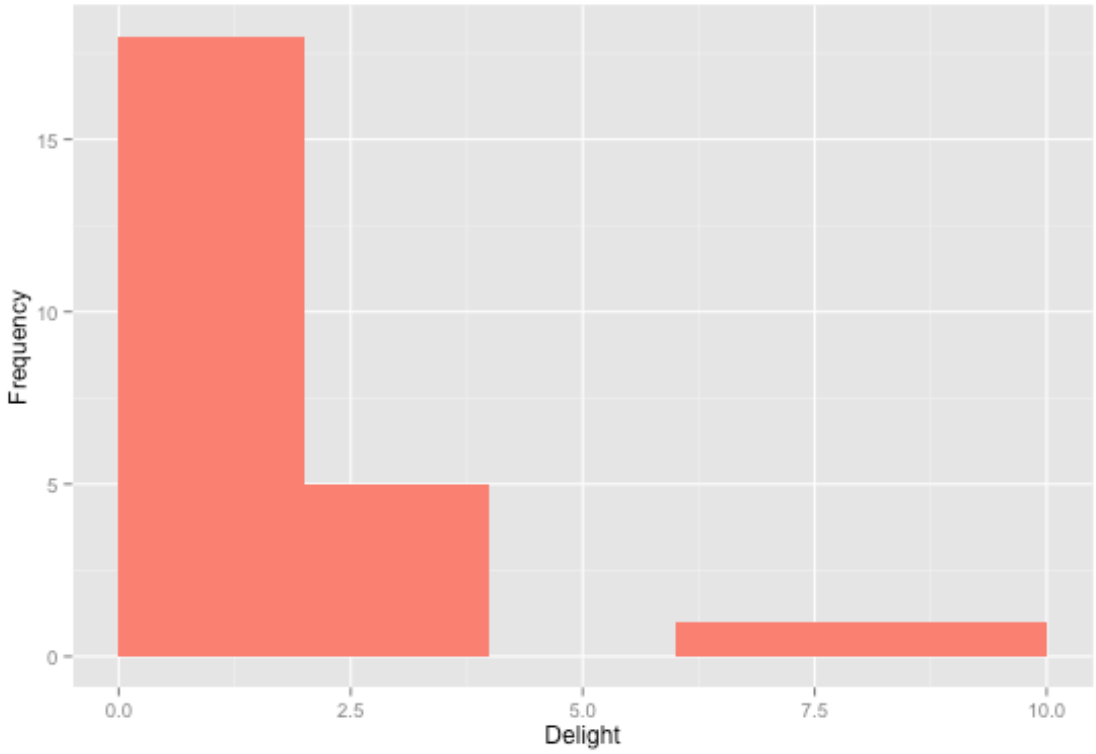
Appendix I. Classroom 2: Affect Variables' Distributions

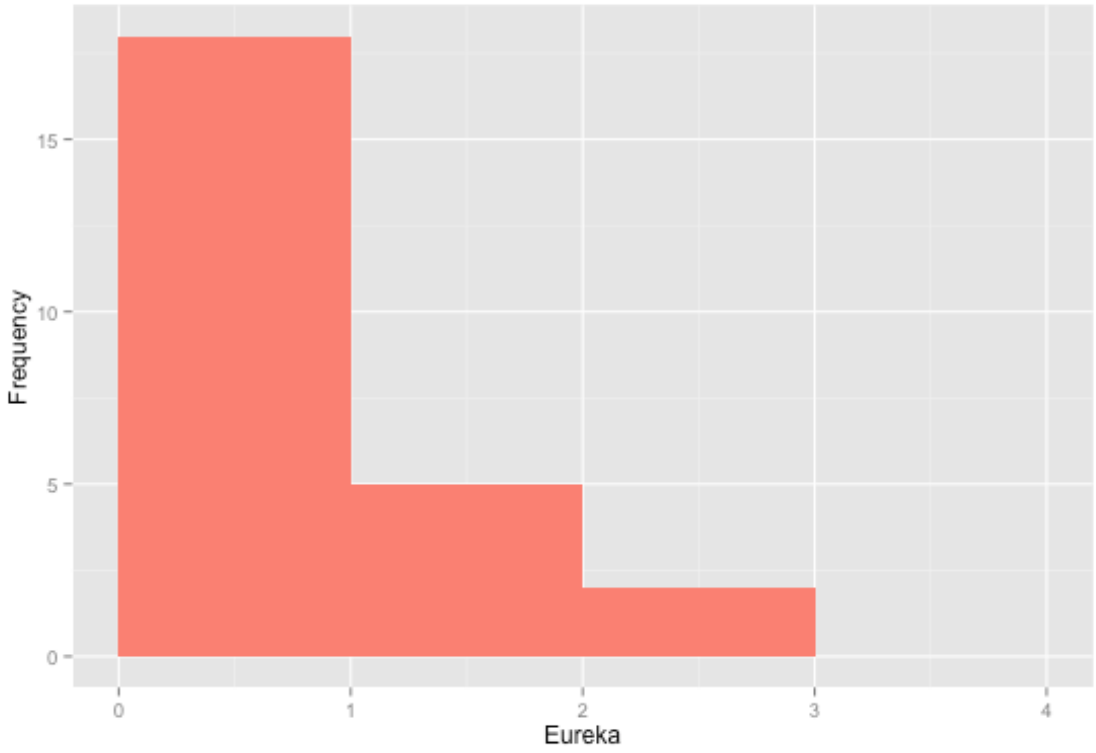
Below histograms represent the Classroom 2 and how many times each of the observed students experiences each of the affect categories. These histograms depict the occurrence of each of the categories in a certain range of values. Y-axis represents the students in the classroom and the area of the block represents the sum of the frequencies that are within that particular range of values. For example, first histogram showcases that 23 out of 25 students in Classroom 2 did not experience surprise throughout their observation periods. Only 2 students have experienced surprises during their observation time.

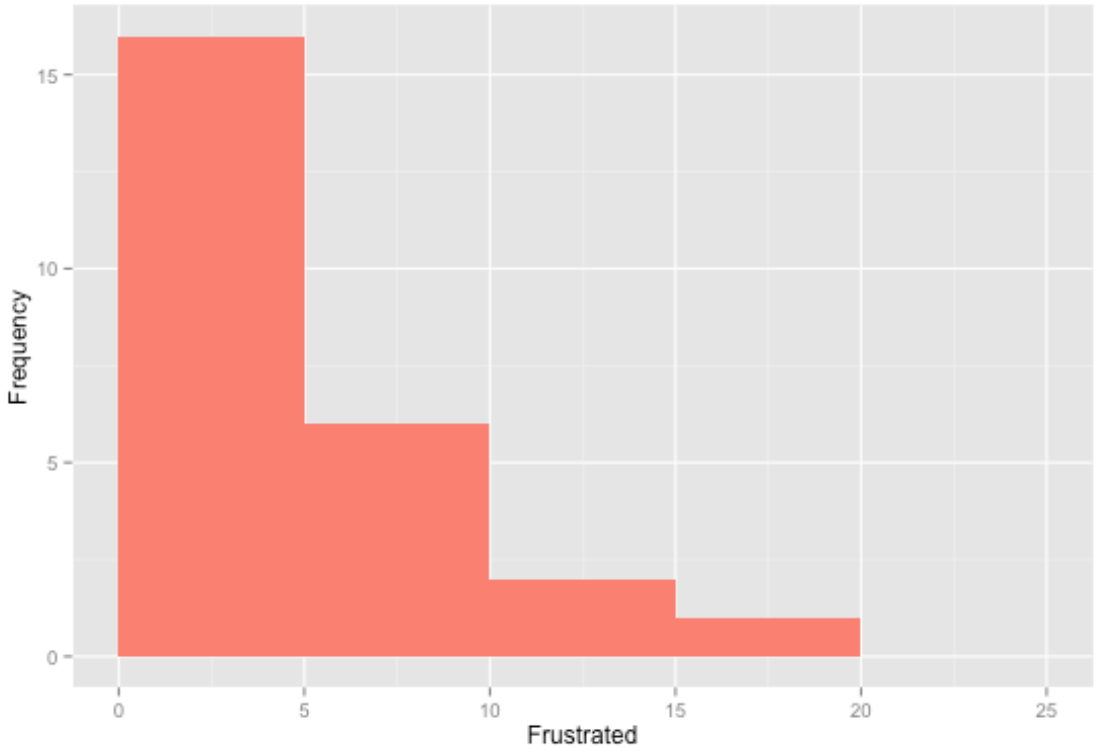


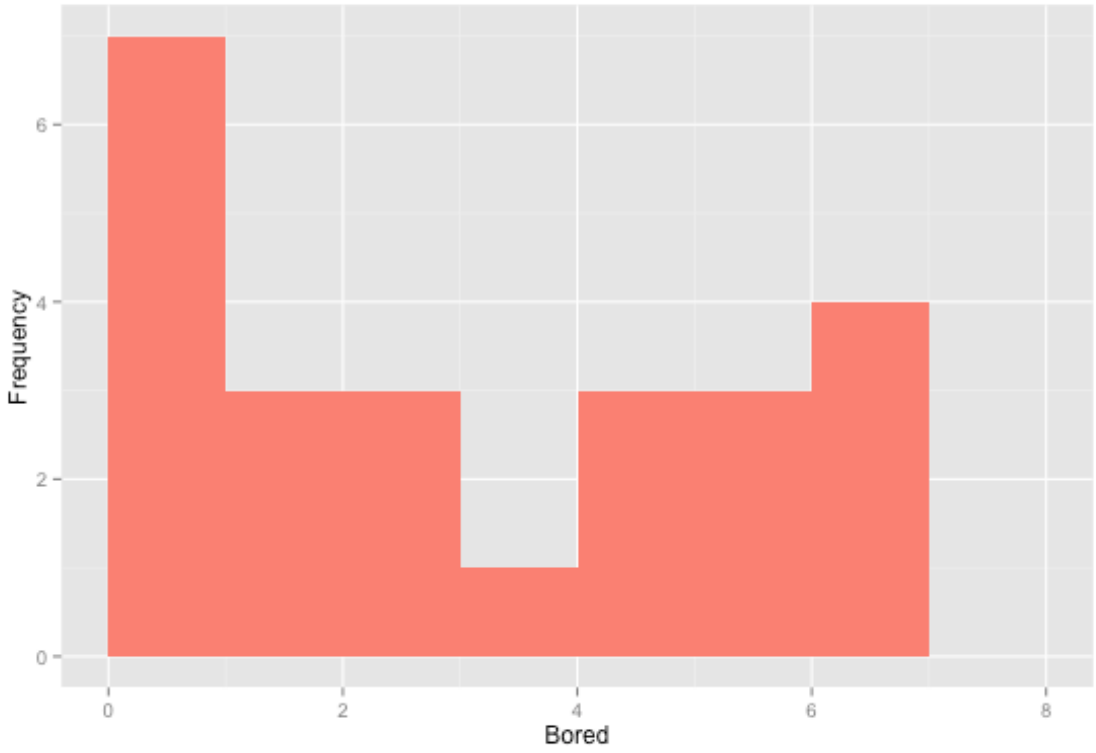

6
4
2
0
Frequency
0
5
10
15
20
Confused


7.5
5.0
2.5
0.0
Frequency
0
20
40
60
Concentrating


Frequency
15
10
5
0
0.0
2.5
5.0
7.5
10.0
Delight

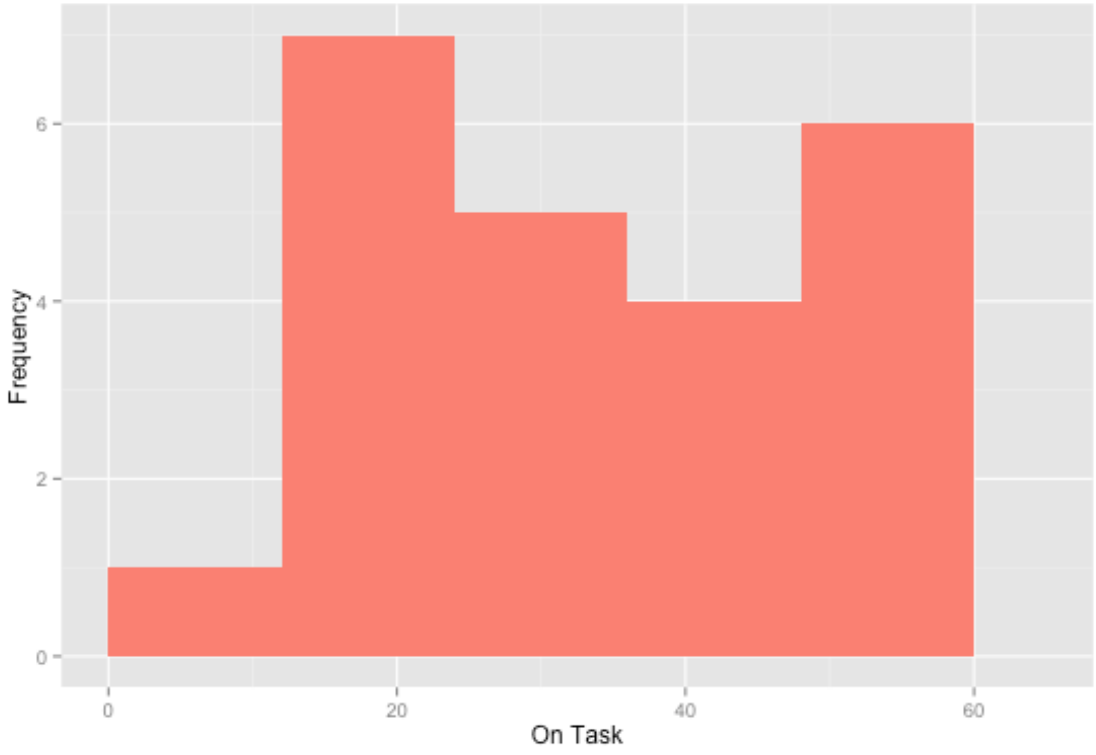

Frequency
15
10
5
0
0
1
2
3
4
Eureka

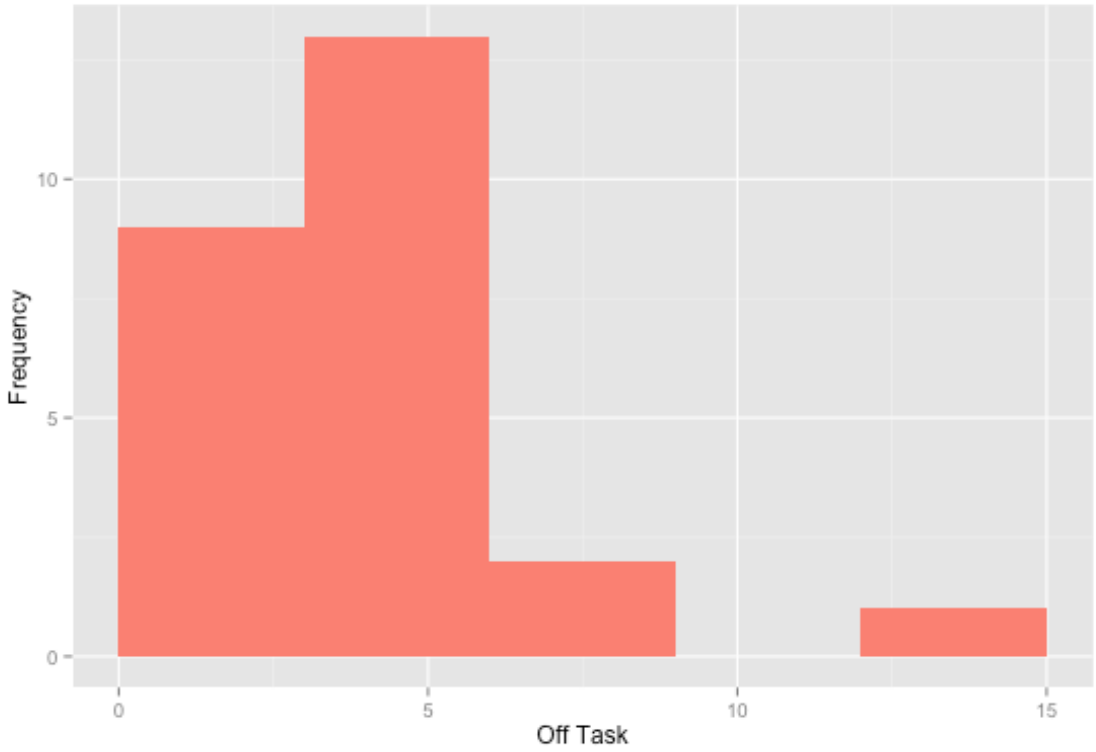

15
10
5
0
Frequency
0
5
10
15
20
25
Frustrated

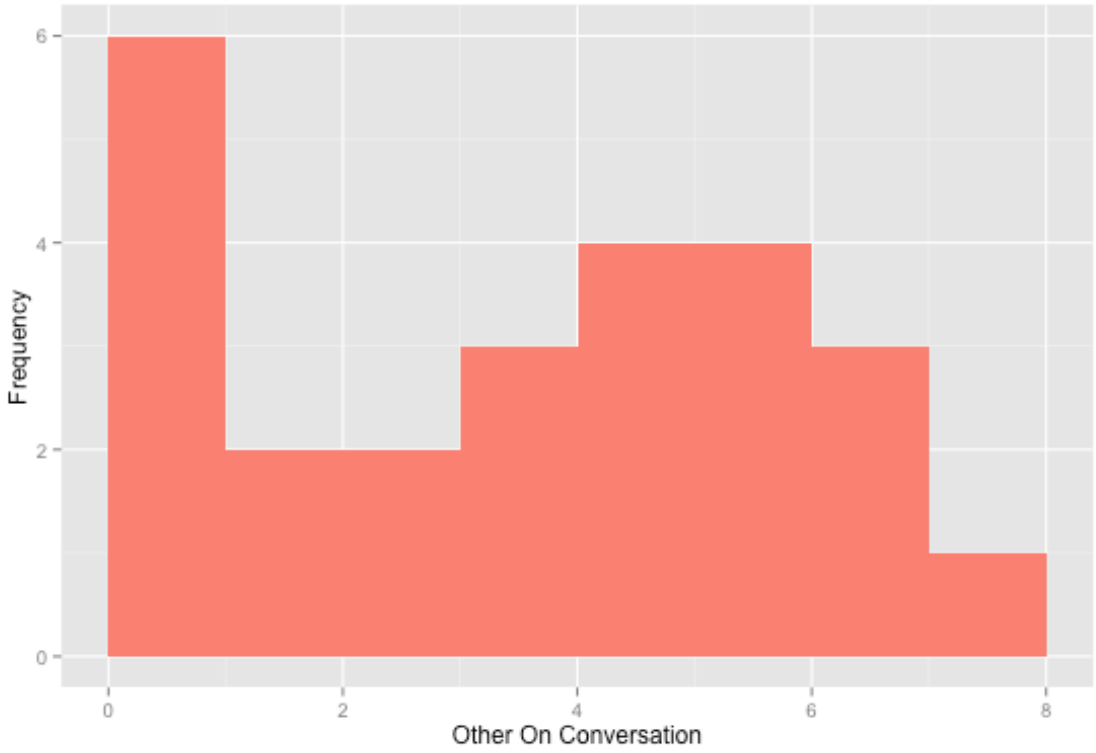

Frequency
6
4
2
0
0
2
4
6
8
Bored

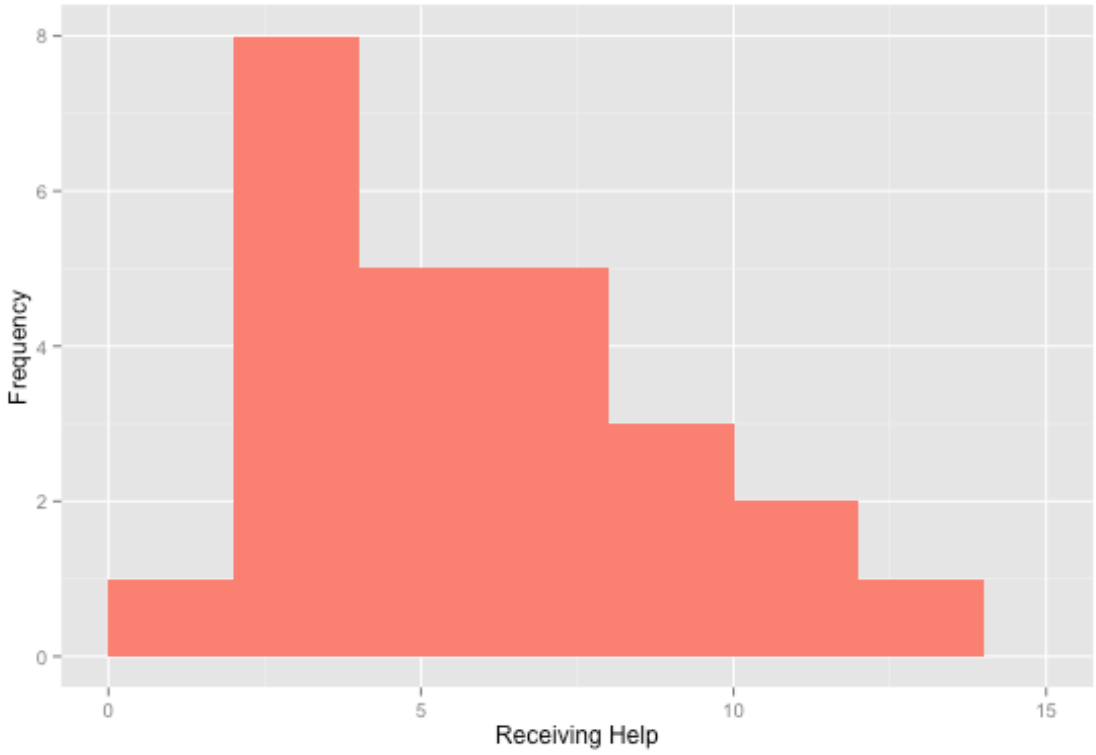
Appendix J. Classroom 2: Behavior Variables' Distributions

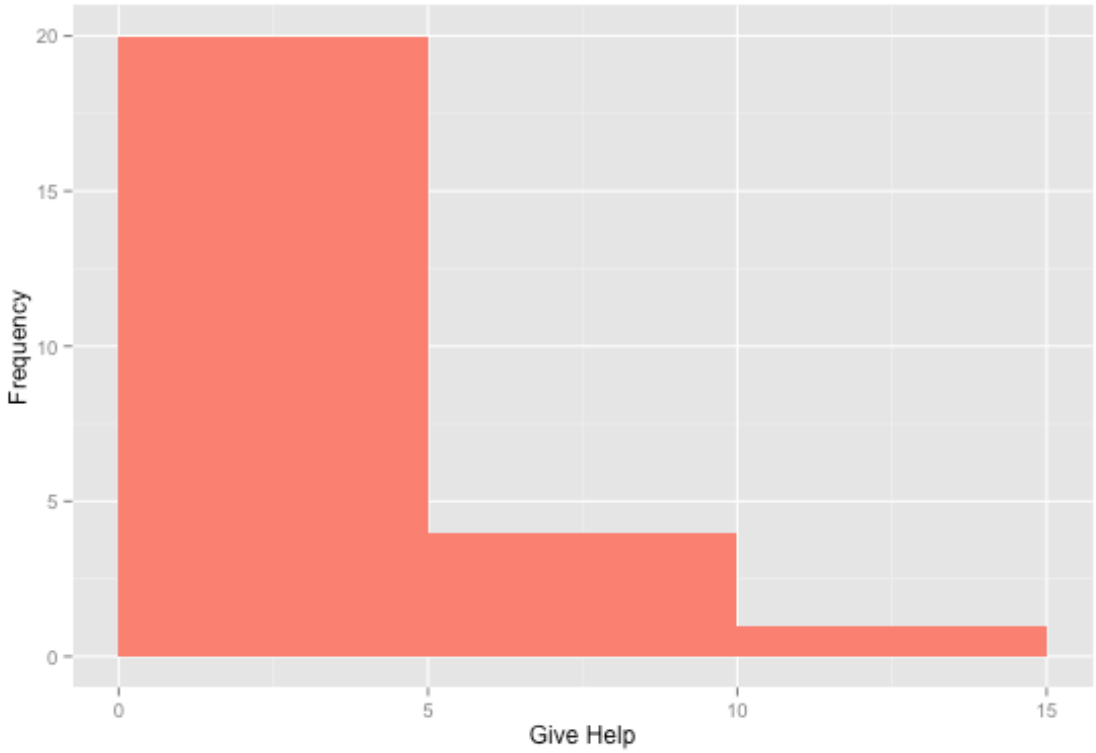
Below histograms represent the Classroom 2 and how many times each of the observed students experiences each and every behavior categories. Again, the area for each block corresponds to the sum of the frequencies that are within that particular range of values. For example, first histogram showcases that there is about 7 occurrences of on task because in the range of 12 and 21. This means that 7 students have demonstrated on task behavior with frequencies that fall between 12 and 21. Data shows that throughout the entire study in Classroom 2 some students had have recordings of 16, others 17, still others 20 on task behaviors.




Frequency
10
5
0
0
5
10
15
Off Task


6
4
2
0
Frequency
0
2
4
6
8
Other On Conversation


8
6
4
2
0
Frequency
0
5
10
15
Receiving Help


20
15
10
5
0
Frequency
0
5
10
15
Give Help

Appendix K. Transformation Table

Affect-Behavior.1	Affect-Behavior.2	Count	Initial Aff-Beh	Ratio of AB1/AB2
C_OT	C_OT	550	986	0.558
C_OT	CF_OT	155	986	0.157
CF_OT	C_OT	105	328	0.320
CF_OT	CF_OT	73	328	0.223
C_RH	C_OT	65	167	0.389
C_OT	F_OT	51	986	0.052
C_GH	C_OT	48	119	0.403
C_OT	C_RH	44	986	0.045
D_OT	C_OT	44	76	0.579
C_OT	C_GH	39	986	0.040
C_RH	C_RH	39	167	0.234
C_OOC	C_OT	38	74	0.514
C_OT	C_OOC	38	986	0.039
CF_OT	F_OT	37	328	0.113
C_GH	C_GH	35	119	0.294
C_OT	D_OT	33	986	0.033
CF_OT	C_RH	28	328	0.085
F_OT	F_OT	26	148	0.176
F_OT	C_OT	25	148	0.169
F_OT	CF_OT	23	148	0.155
F_OT	C_RH	20	148	0.135
C_GH	CF_OT	17	119	0.143
C_OfT	C_OT	16	39	0.410
CF_OT	B_OfT	16	328	0.049
C_OT	C_OfT	15	986	0.015
CF_OT	C_OOC	14	328	0.043
B_OT	C_OT	13	31	0.419
C_OT	CF_OOC	13	986	0.013
C_RH	CF_OT	13	167	0.078
CF_OT	C_GH	13	328	0.040
D_OT	C_GH	11	76	0.145

Note. See Glossary for complete descriptions of these categories. In addition, due to the length of the original table, the original list was truncated and only most frequent rows were included in this Appendix.

VITA

ANI AGHABABYAN
UTAH STATE UNIVERSITY (USU)
INSTRUCTIONAL TECHNOLOGY AND LEARNING SCIENCES (ITLS)
2830 Old Main Hill, Logan, UT 84321-2830, (435) 754-6285
www.linkedin.com/in/aniaghababyan
Active Learning Lab
anie.aghababyan@gmail.com

EDUCATION

PhD. Utah State University
Degree expected (August 2014)
Instructional Technology and Learning Sciences
2014
Dissertation: "E^3: Emotions, Engagement, and Educational Digital Games"

M.S Utah State University
Management of Information Systems 2013

M.S. Utah State University
Business Administration
2009

B.S. French University in Armenia
Law and Political Science
2008

RESEARCH EXPERIENCE

Graduate Student Researcher
2012 - Present
Utah State University, ITLS Department
Supervisor: **Taylor Martin, Ph.D.**

Contributing to several grants such as Gates Foundation & Darpa projects (Engage: Refraction), iPRO (Programming Standing Up), Virtual Manipulatives (collaborating grad student).

Mining and analyzing educational game back end data in order to detect patterns to inform algorithms for future behavior classifications and for sequential pattern mining. Employing visualization and data mining techniques in the analysis of the large-scale data that is being collected from the game log. The ultimate goal is to discover optimal learning pathways for students and eventually be able to adjust the learning process individually for each student.

Extracting student engagement data, synchronizing it with student gameplay log information (moves, success, failure, time, etc.) in order to find patterns that will allows us to classify certain behaviors and affective state for further automatic

detection. Conducting affect recording studies in order to look at temporal dimension of affect data along with its sequential patterns. Working with predictive statistical tools and techniques: applying algorithms such as inter-sequence distance with optimal matching, HiddenMarkov, SPAdes frequent sequencing etc. Using data mining and data science tools and techniques: python and ruby on rails for data manipulation and cleaning, used R packages for traditional and alternative statistical analyses.

Collaborating with the Department of Psychology to develop, design and conduct studies that use Near Infrared Spectroscopy brain imaging machine to analyze students' brain activity while playing Refraction game.

Graduate Student Researcher
2012 - 2013
Utah State University, Math and Statistics Department
Supervisor: **Juergen Symanzik, Ph.D.**

Developing a new prototype for graphically visualizing Refraction game's log data from each student's gameplay. Extracted data, analyzed log data for underlying patterns, and created visualizations using R and Ruby on Rails. Comparing demographic and learning data to understand student gameplay through exploratory analysis with implications for student learning.

Graduate Student Researcher 2012 - Present
Utah State University, ITLS Department
Supervisor: **Deborah Fields, Ph.D.**

Worked on the Scratch grant writing project: Macro Data for Micro Learning: Developing FUN! API for Automated Assessment of Computational Thinking in Scratch. Researched possible learning analytics tools and measures to help identify learning trajectories in the data collected from Scratch programs. Contributed to literature review section for the grant proposal. Facilitated Scratch workshops held at USU ITLS department.

Working on data management, data analysis and literature review on Computational thinking and Learning Analytics methods that deal with such data as Scratch code blocks. Reviewing and experimenting with methods used in the field for analysing code snapshots and inferring about student programming skills (e.g., Needleman-Wunsch algorithm, Abstract Syntax Tree, Hidden Markov Model etc.)

Graduate Student Researcher
2011 - 2012
Utah State University, ITLS Department
Supervisor: **Victor Lee, Ph.D.**

Worked on PAD (Physical Actvity Data) projects conducting interviews with several schools' 5th graders in order to learn about the relationship between use of

physical activity devices (e.g. pedometers, heart rate monitors, bike computers, accelerometers) and student ability to understand data representations especially in connection with mathematical concepts. I worked on designing classroom activities where students use their own physical activity data for mathematical and scientific investigations.

Graduate Student Researcher
2011
Utah State University, Management Department
Supervisor: **Kenneth Bartkus, Ph.D.**

Worked on "Learn and Earn" Institutions' research project funded by Gates Foundation. Investigated the attendance rate between four-year universities and two-year college programs.

Graduate Student Researcher
2009 - 2011
Utah State University, Management Department
Supervisor: **Kenneth Bartkus, Ph.D.**

Collected preliminary evidence of the reporting of coefficient alpha in the scholarly literature. Systematic review of relevant literature for the reporting of coefficient alpha. Compared the field standard. Collected a database of results and their deviation from the accepted number and prepared report based on my findings.

TEACHING EXPERIENCE

Teaching Assistantship
Utah State University, MIS Department
Professor: Nicole Forsgren Velasquez, Ph.D.
Course: MIS 3860 - Fall 2013

Bid Data Analytics is an undergraduate course that provides an introduction to business intelligence and analytics. The course includes the use of data, statistical and quantitative analysis, exploratory and predictive models, and evidence-based methods to inform business decisions and actions. Contributed to the design and implementation of the course: preparation of teaching modules, assignments, syllabus, grading rubrics, and in class instruction. Worked with students to implement analyses such as recommender systems, tableau data visualizations, exploratory, and predictive models (e.g., clustering, classification, decision trees, random forest and other algorithms). Graded all of the student projects throughout the course.

Teaching Assistantship
Utah State University, ITLS Department
Professor: Taylor Martin, Ph.D.
Course: ITLS 6870/7870 - Spring 2013 and Spring 2014)

Instruction and the Data Deluge is a graduate course that covers questions on the use of the massive quantity of existing data to improve learning – particularly online learning in environments like games. Examples are the recommender systems, data mining methods used by Facebook and Google and learning analytics used in Educational Data Mining field. Contributed to the design and development of the course, along with content selection, lab session facilitation, in class teaching and out of class support for students. Co-prepared the class content for data mining analyses such as data analysis in RapidMiner, Python IDLE, Weka, RStudio etc. Tutored students on data mining tools (e.g., R, Python IDLE, WEKA, SPSS, SQL) and techniques (e.g., Association Rule mining, Sequence Pattern mining, Inter-sequence Distance with Optimal Matching etc.)

SELECTED PUBLICATIONS & PRESENTATIONS

Aghababyan, A., Dai, X., & Martin, T. (2014, July). Supervised Learning Methods for Predicting Student Success in a Digital Game Environment. Paper submitted to Educational Data Mining Conference, London, UK.

Martin, T., Forsgren Velasquez, N., **Aghababyan, A.,** & Maughan, J. (2014, July). Microgenetic Designs for Educational Data Mining Research. Paper submitted to Educational Data Mining Conference, London, UK.

Aghababyan, A., Baker, J., & Martin, T. (2014, April). Students' neurological response patterns while playing math games. Paper to be presented at the American Educational Research Association Annual Conference, Philadelphia, PA.

Aghababyan, A., Martin, T., & Harris-Brasiel, S. (2014, April). Understanding how frustration and confusion manifest in educational games. Paper to be presented at the American Educational Research Association Annual Conference, Philadelphia, PA.

Martin, T., Petrick Smith, C., Forsgren Velasquez, N., **Aghababyan**, A., Janisiewicz, P., & Baker, S.. Learning fractions by splitting: Using learning analytics to illuminate the development of mathematical understanding. Journal of the Learning Sciences. *Under Review.*

Aghababyan, A., Martin, T., Forsgren Velasquez, N., and Janisiewicz, P. (2013). Educational data mining: Illuminating student learning pathways in an online fraction game. *Proceedings of the Sixth International Conference on Educational Data Mining,* Memphis, TN.

Aghababyan, A., Symanzik, J., and Martin, T. (2013). Visualization of "States" in online educational games. *Proceedings of the 59th World Statistics Congress.* Hong Kong, China.

Baker, S, Petrick-Smith, C., Martin, T., **Aghababyan, A**., Popovic, Z., Andersen, E., and Liu, Y. (April, 2013). *Learning fractions through splitting in an online game.* Paper

presented at the American Educational Research Association Annual Conference, San Francisco, CA.

Martin, T., **Aghababyan, A.,** Petrick-Smith, C., Olsen, J., Pfaffman, J., Phillips, R. Baker, S., & Janisiewicz, P. (2013). Nanogenetic learning analytics: Illuminating student learning pathways in an online fraction game. *Proceedings of the Third Conference on Learning Analytics and Knowledge, Leuven, Belgium*.

Martin, T., Petrick Smith, C., Forsgren Velasquez, N., **Aghababyan, A.,** Janisiewicz, P., & Baker, S. (2013). *Learning fractions by splitting: Using Learning Analytics to illuminate the development of mathematical understanding.* Manuscript submitted for publication.

Martin, T., Symanzik, J., **Aghababyan, A.,** & Janisiewicz, P. (2013, January). *The development of students' understanding of fractions in a splitting game: A microgenetic learning analytics approach.* Working Group lead at the Alpine Rendez-Vous 2013 Panel "It's About Time: Addressing the Many Challenges of Analyzing Multi-Scale Temporal Data," Vercors, France.

Philips, R., Smith, C., Martin, T., Horstman, T., Janisiewicz, P., & **Aghababyan, A.** (2013). Maximizing the use of human coders and automated techniques to study learning in educational games. *Proceedings of the 15th Biennial Conference of the European Association for Research in Learning and Instruction*. Munich, Germany.

Aghababyan, A. & Bartkus, K. (April, 2011). Is Coefficient Alpha Correctly Reported? Preliminary Evidence From The Scholarly Literature. Paper presented at the annual meeting of Western Decision Sciences Institute, Portland, OR.

ACADEMIC HONORS, AWARDS, SCHOLARSHIPS & CERTIFICATES

Dissertation Scholarship from the USU Center for Women and Gender
2014

Dissertation Scholarship from the Center for Open & Sustainable Learning
2013

TA for visualizations' week on Ryan Baker's EDM course on Coursera
2013

LearnLab Completion at Pittsburgh Science of Learning Center
2013

Outstanding Graduate Assistant
2013

Honorary Mention at Graduate Research Symposium
2013

Best Women's Team category at Women's Hackathon by WLC
2013

Baker-Rodrigo Observation Method Protocol (BROMP 1.0) Certificate
2012

Jon M. Huntsman Master's Full Academic Scholarship for MBA
2008–2009

WORK EXPERIENCE/PROFESSIONAL SERVICE

Web Developer/Programmer 2011 – 2013

Utah State University

Information Technology Department

- Provided personalized programming services and web development solutions: created web sites based on requested designs, provided training for the client's personnel
- Trained and supported clients with EZplug content management system
- Developed web templates and plugins upon clients' requests using cold fusion application development platform
- Project management: negotiated the necessary time for both the designed and coding process of new websites

Jon M. Huntsman School of Business

- Co- developed two new websites for the Jon M. Huntsman School of Business
- Designed and coded personalized profile accounts for all current and prospective students
- Researched and co-implemented one-login secure system across all business school domains and plugins

Instructional Technology Department

- Provided web support with content creation and management of the departmental website
- Redesigned, maintained and updated the website pages, plugins and populated the content
- Analyzed and reported the data on web traffic over time using Google Analytics tools
- Improved departmental website's incoming traffic rate by incorporating search engine optimization techniques (SEO)
- Monitored and ensured web servers uninterrupted functionality
- Worked closely with the marketing and social media team to enhance Department's marketability

Center for e-Commerce and Business Analytics

- Negotiated a website design project with Scandia Amusement parks in California
- Designed the website (using Photoshop and Illustrator)
- Programmed the website (using SQL Server, ASP.Net, and HTML/CSS)

Coursera

- Was the official assistant for the Educational Data Mining course on Coursera.com: data visualizations' week

Wishopy.com

- Created electronic business-project: HTML 5 based web site project with a business plan and marketing strategy

PROFESSIONAL SKILLS/QUALIFICATIONS

Programming Languages & Frameworks

Visual Basic	SQL Server	MySQL	PL/SQL
HTML (XHTML)	CSS	Cold Fusion	Java Script
jQuery	Ruby on Rails	Python	D3
Hadoop (HDFS)	Hadoop Pig (PigLatin)	Unix	R (RStudio)

Programming Tools & Other Computer Skills

MySQL Workbench	Visual Studio	RStudio	Virtual Box
VMware Fusion	Android SDK	WEKA	RapidMiner
Database Administration	Data Science	Data Mining	Big Data Analytics
Photoshop (CS5)	MS Office	MS Visio	MS Publisher
Tableau	Mondrian	CMS (e.g., Drupal, Joomla, EZplug)	
GGobi	MS Windows XP/Vista/7	Mac OS	Illustrator (CS5)
After Effect	Agile Scrum	Web Analytics	
Basecamp	Trello	Yammer SDLC	Zoho
QTP	SublimeText	GitHub	Version Control
Source Tree	TextWrangler	Aptana Studio 3	

Language Skills

English – fluent
Russian – fluent
French – advanced
Armenian – native