

## ABSTRACT

SHAREK, DAVID JAMES. Investigating Real-time Predictors of Engagement: Implications For Adaptive Video Games and Online Training. (Under the direction of Dr. Eric Wiebe.)

Engagement is a worthwhile psychological construct to examine in the context of online training and video games. In this context, previous research suggests that the more engaged a person is, the more likely they are to experience overall positive affect while performing at a high level. This research builds on theories of engagement, Flow Theory, and Cognitive Load Theory, to operationalize engagement in terms of cognitive load, affect, and performance. Studies One and Two examined the development of a puzzle-based video game as a platform for measuring engagement. Results of these studies demonstrated that: cognitive load could be measured using a secondary task, affect could be measured using a novel game-clock mechanism, and performance could be captured incrementally throughout gameplay. Study Three compared traditional linear gameplay, choice-based gameplay, and adaptive gameplay that used a real-time measure of engagement to select game levels. Results showed that those in the Adaptive condition performed higher compared to those in the other two conditions without any degradation of overall affect or self-report of engagement.

Investigating Real-time Predictors of Engagement: Implications For Adaptive Video  
Games and Online Training

by  
David J. Sharek

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

Psychology

Raleigh, North Carolina

2012

APPROVED BY:

---

Eric Wiebe, PhD.  
Committee Co-Chair

---

Christopher Mayhorn, PhD.  
Committee Co-Chair

---

Anne McLaughlin, PhD.

---

Doug Gillan, PhD.

## DEDICATION

For my father. Thanks for looking out for me. I miss you.

## BIOGRAPHY

I spent my formative years in France, Switzerland, and the United Kingdom. After moving back to the States for university, I studied architecture, art, graphic design, and journalism. Along the way, I taught myself various programming and multimedia development tools, with Flash as my weapon of choice. After graduating, I worked as a multimedia manager, programmer, and designer in the eLearning sector. During my career I co-authored a book on 3D Flash design, and this experience helped me realize how much I enjoyed researching and writing. After my seventh year of working in the industry, I decided to satisfy my intellectual passions by returning to university to pursue a doctorate in the field of human factors psychology. I'm immensely happy I made this decision and am looking forward to the next chapter.

## ACKNOWLEDGEMENTS

First and foremost I want to thank my advisor. Dr. Wiebe has been, without a doubt, the best advisor I could have hoped for. He pushed me when I needed to be pushed, gave me room when I needed room to explore, and was always available when I needed guidance. It has been an honor working with Dr. Wiebe and I consider him, not only a mentor, but also a friend.

I also would like to thank my committee members for their invaluable help with this dissertation. Dr. Doug Gillan has been instrumental in my overall Human Factors education. Through the classes he taught and the discussions we have had I know I am well prepared for a future in Human Factors. Dr. Anne McLaughlin has provided valuable help with the capturing and presenting of my data for both my thesis and this dissertation. Early on in the program, Dr. Chris Mayhorn taught me the importance of adhering to APA style and succinct writing, something I have striven for ever since.

Most importantly, my amazing wife Lisa has been a beacon of light throughout this process. I couldn't have done it without her love, support, and fun-loving personality. I feel like we did this together; she made this experience fun. My mother's encouragement, enthusiasm and earnest willingness to discuss ideas has helped me countless times. She taught me to never be afraid to ask difficult questions and I can't thank her enough for all she has done for me. Finally, I wish to thank my best friend and

big brother, Zach. His doctoral pursuits blazed the path for my own academic endeavors.  
His genuine interest and support in my research is greatly appreciated.

## TABLE OF CONTENTS

LIST OF TABLES .....	ix
LIST OF FIGURES.....	x
Introduction.....	1
Purpose of Study.....	5
Video Games .....	6
Designing Effective Video Games .....	8
Flow Theory.....	14
Cognitive Load Theory.....	21
Measuring Cognitive Load .....	24
Linking Flow Theory and CLT with Engagement .....	25
Engagement .....	29
Measuring Engagement .....	31
Summary.....	33
Study One.....	35
Independent Variables .....	37
Individual Game Level Performance Dependent Variables .....	37
Subjective Ratings Dependent Variables.....	38
Study One Method .....	39
Participants .....	39
Materials and Apparatus.....	39
GridBlocker .....	40
Procedure .....	42
Study One Results.....	43
Study One Discussion .....	46
Study Two .....	47
Independent Variable.....	48
Real-time Engagement Dependent Variables.....	49
Gameplay Performance Dependent Variables.....	49

Summative Self-Report Dependent Variables .....	50
Overall Engagement .....	50
Cognitive Load .....	51
Study Two Method.....	51
Participants .....	51
Materials and Apparatus .....	51
Procedure .....	52
Study Two Hypotheses .....	53
Confirmation of Game Performance Metric.....	53
Diagnosticity of the Real-time Measures .....	53
Study Two Results .....	54
Study Two Discussion.....	58
Study Three .....	62
Study Three Design.....	63
Independent Variables .....	63
Real-time Engagement Dependent Variables .....	67
Gameplay Performance Dependent Variables .....	67
Summative Self-Report Dependent Variables .....	68
Study Three Method.....	68
Participants .....	68
Materials and Apparatus .....	69
Procedure .....	69
Hypothesis – Summative Measures .....	70
Hypothesis – Performance .....	71
Results .....	72
Summative Measures .....	72
Performance .....	76
Study Three Discussion.....	84
Summative Measures .....	85
Real-Time Measures .....	86

General Discussion.....	90
Application .....	96
Limitations and Future Research .....	97
Conclusion .....	98
REFERENCES.....	99
APPENDICES .....	111
Appendix A: Study One Screenshots .....	112
Appendix B: Study One Informed Consent Form .....	114
Appendix C: Study One Performance and Self-report Distributions .....	115
Appendix D: Game Engagement Questionnaire (GEQ) Items.....	117
Appendix E: Intrinsic Motivation Inventory .....	118
Appendix F: Study Two Self-report Distributions .....	119
Appendix G: User Engagement Scale .....	121

## LIST OF TABLES

Table 1. <i>Research Overview</i> .....	6
Table 2. <i>The Dimensions of Flow Theory</i> .....	15
Table 3. <i>Mean Scores for all 108 levels</i> .....	43
Table 4. <i>Bivariate correlations</i> .....	44
Table 5. <i>Hierarchical Regression Results Predicting Self-report Difficulty</i> .....	45
Table 6. <i>Bivariate Correlation Table</i> .....	55
Table 7. <i>Summary of the Three Conditions Used in Study Three</i> .....	63
Table 8. <i>UES Descriptive Statistics</i> .....	72
Table 9. <i>Affect Descriptive Statistics</i> .....	73
Table 10. <i>Summative Cognitive-Load Descriptive Statistics</i> .....	74
Table 11. <i>Adaptive Algorithm's Difficulty Level Selection Descriptive Statistics</i> .....	75
Table 12. <i>Overtime Descriptive Statistics</i> .....	76
Table 13. <i>Difficulty, Levels, and Difficulty/level Ratio Descriptive Statistics</i> .....	77

## LIST OF FIGURES

<i>Figure 1.</i> A Non-linear increase in Challenge and Skill.....	10
<i>Figure 2.</i> Flow Theory’s Interaction between Challenge and Skill .....	16
<i>Figure 3.</i> The Four Channel Flow Model .....	18
<i>Figure 4.</i> The Eight Channel Flow Model .....	19
<i>Figure 5.</i> Adapted Model of Flow used in the Current Study .....	21
<i>Figure 6.</i> Cognitive Load Theory and Flow Theory.....	24
<i>Figure 7.</i> The Engagement / Flow Cycle .....	27
<i>Figure 8.</i> Moving the Block Around the Game Board in the GridBlocker Game.....	41
<i>Figure 9.</i> Mean Game-clock Checks as a Function of Difficulty .....	57
<i>Figure 10.</i> Adaptive Algorithm Model .....	66
<i>Figure 11.</i> Difficulty, Levels, and Difficulty/level Ratio .....	79
<i>Figure 12.</i> Overlay of Individual Player Paths - Linear Condition.....	80
<i>Figure 13.</i> Overlay of Individual Player Paths - Choice Condition.....	80
<i>Figure 14.</i> Overlay of Individual Player Paths - Adaptive Condition .....	81
<i>Figure 15.</i> Composite Overlay of Mean Difficulty by Level Order .....	82
<i>Figure 16.</i> Mean Time to Complete Individual Levels.....	83
<i>Figure 17.</i> Linear Condition Overlaid on the Flow Theory Model .....	91
<i>Figure 18.</i> Choice Condition Overlaid on the Flow Theory Model .....	93
<i>Figure 19.</i> Adaptive Condition Overlaid on the Flow Theory Model .....	95

## Introduction

The purpose of this research is to explore, define, and test the efficacy of the psychological construct of engagement as a predictor of performance and affect in learning and video games. Though engagement has been researched for many years, the psychological literature is still fraught with multiple and vague definitions of this construct. This is due, in part, to the disparate application contexts in which engagement has been researched. For example, engagement has been examined in areas ranging from online-shopping (O'Brien & Toms, 2008), to online-learning (Chen, Lambert, & Guidry, 2010), to employee engagement (Macey & Schneider, 2008). This point is not made in order to assert that only one definition of engagement should eventually prevail. Instead, the scope in which engagement is investigated should be clearly stated and not generalized to other untested areas. Despite these variations, researchers agree that engagement is an overall positive state and a worthwhile psychological construct to examine (e.g., Brockmyer et al., 2009; Whitton, 2011). Rather than attempt to create a single, overarching definition of engagement the research reported here narrows the scope of engagement to its theoretical importance and application in the synergistic area of online training and video games—an area where high engagement is thought to be of the utmost importance (McGinnis, Bustard, Black, & Charles, 2008; Richardson & Newby, 2006; Whitton, 2011).

## REAL-TIME PREDICTORS OF ENGAGEMENT

Online Training is important to corporations because it has the potential to provide scalable training to large numbers of employees at much lower costs compared to traditional classroom implementations (Choi, Kim, & Kim, 2007; Fisher, Wasserman, & Orvis, 2010; Welsh, Wanberg, Brown, & Simmering, 2003). Additionally, corporations have experienced positive benefits, such as increased employee retention (Cross, 2004a), when investing in online training that cultivates the wellbeing of current employees. Online training is also important to lifelong learners (Koper & Tattersall, 2004) who otherwise may not have the opportunity to learn new skills, concepts, or procedures without the practical and financial support typically offered in corporate or academic environments.

Corporations recognize these benefits, yet despite their total US expenditure on employee training and development of almost \$126 billion in 2009 (A.S.T.D., 2010), the effectiveness of online training has often failed to meet the expectations of its sponsors (Fisher, et al., 2010). Perhaps a reason for this is that despite its perceived benefits, participating in online training has been characterized as a boring, rigid, or unengaging experience (Adams & Granic, 2009; Cross, 2004b; Edwards, 2010; Sansone, Fraughton, Zachary, Butner, & Heiner, 2011). This characterization may be exacerbated by the common corporate view that online training is somewhat equivalent to online information delivery, where static content simply needs to be uploaded to a Web server for employees

## REAL-TIME PREDICTORS OF ENGAGEMENT

to download (Welsh, et al., 2003). This type of content delivery may make information more accessible, but it is not considered to be optimal online training since it does not include any form of pedagogical strategy or function to actively engage users.

Though a positive relationship has been found between the use of online training technologies with strong pedagogical foundations and student engagement (Chen, et al., 2010), attrition is still a concern for online training sponsors (Frankola, 2001). This may be caused, in part, to the difficulties in ensuring that a person will remain actively engaged in online training material when the content is delivered in a typically unmonitored, self-regulated environment (Sansone, et al., 2011). For example, it is common for employees to simply skip to the end of an online training course and attempt to guess the answers to the required assessment questions or, if they can get away with it, they may even choose to not complete the course at all (Frankola, 2001; Wiebe, Branoff, & Shreve, 2011).

A potential reason for this lack of motivation is that online training content does not always *adapt* to the user's cognitive (Chen & Macredie, 2002) and affective (Fisher, et al., 2010; Frankola, 2001) needs, resulting in a lack of engagement. This is an important issue to investigate because the more engaging online training is, the more likely a user will experience intrinsic motivation resulting in positive academic

## REAL-TIME PREDICTORS OF ENGAGEMENT

performance (Charlton & Birkett, 1999). Unfortunately though, online training does not always engage its users (Frankola, 2001).

Engagement is considered to be a positive behavior (Charlton & Birkett, 1999) and something to be pursued in video games and online training environments. Video games may provide useful insight into the development of adaptive and engaging online training. Research has already shown that games in general can be used to enhance learning (Cordova & Lepper, 1996; Huizinga, 1944) and video games, specifically, lend themselves to being both engaging and adaptive (Gilleade & Dix, 2004; Gilleade, Dix, & Allanson, 2005). In fact, video games can be so engaging (Garris, Ahlers, & Driskell, 2002) that research has found that some people exhibit behavioral addictions (Charlton & Danforth, 2007) when immersed in certain types of video games. Addictions notwithstanding, incorporating the affective and challenging qualities of video games may provide a means to decrease attrition and increase online training engagement. For example, video game research has shown that providing the appropriate level of challenge a person wishes to face (Champion, 2011; Malone, 1980) is essential to maintaining user interest (Aponte, Levieux, & Natkin, 2009; Rabin, 2005).

Challenge is a central component of gaming and game-like environments (Aponte, et al., 2009) and has been determined to be an important factor in engagement (Gilleade, et al., 2005; O'Brien & Toms, 2008). Challenge can be operationally defined as

## REAL-TIME PREDICTORS OF ENGAGEMENT

the amount of cognitive load a person experiences during a given task. Ideally, in a gaming environment, the challenge a user faces should dynamically adjust depending on the player's cognitive and affective state. However, cognitive load has been traditionally measured using post-task, self-report measurement tools that do not lend themselves well to real-time adaptive applications. If cognitive load can be measured in real-time with confidence, the balancing of a person's perceived level of challenge may result in higher levels of engagement. However, in order to improve the flexibility and attractiveness of a training environment it is important to integrate a degree of adaptability so that both the immediate cognitive and affective needs of the user can be met (Adams & Granic, 2009). Coupling what is known about balancing challenge with what is known about adaptive systems may make it possible to develop more engaging online training systems.

### *Purpose of Study*

The goal of this present study was to investigate an empirically-based approach to increasing user engagement through the development and application of a real-time cognitive load measure and intermediate measure of affect. The measure could potentially be used to have game-based environments automatically adapt to the degree of perceived challenge a person willingly seeks to experience while playing.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Table 1. *Research Overview*

Medium	Video Games & Serious Games
Theoretical Frameworks	Flow Theory Cognitive Load Theory Engagement
Application	Adaptive video game engine
Goal	Measure engagement and increase performance
Future Directions	Application of findings towards increasing engagement in video games and online training

As can be seen in Table 1, ideas from video game design, Flow Theory, Cognitive Load Theory, and theories of Engagement were used to develop an adaptive video game engine with the goal of increasing user engagement.

### *Video Games*

For the purposes of this study, a game is defined as “a rule-based formal system with a variable and quantifiable outcome, where different outcomes are assigned different values, the player exerts effort in order to influence the outcome, the player feels attached to the outcome, and the consequences of the activity are optional and negotiable” (Juul, 2003, p. 35). A video game can be thought of as a digital extension of this definition.

Video games are quickly becoming integrated into many aspects of our daily lives. In 2009, the combined sales for the video game and computer game industry

## REAL-TIME PREDICTORS OF ENGAGEMENT

reached \$10.5 billion in the US (ESA, 2010) and almost \$11 billion in the European market (ISFE, 2010). To put these numbers in perspective, the total 2009 revenue for the US and Canadian film industry was a comparable \$10.6 billion dollars (MPAA, 2010). Sixty-seven percent of American households play video games (ESA, 2010) and in 2010, the mean age of US gamers was 34 years old, with 26% of gamers over the age of 50 (ESA, 2010). Additionally, females made up 40% of the entire gaming population, with those over 18 years old representing 33% of the gaming community (ESA, 2010). These statistics highlight the broad popularity of video games among people of similar demographics to those in the workforce who are likely to be required to take online training.

Serious games are a subset of the video game industry and represent the use of games for training and educational purposes. The term serious games was coined by Clark Abt (1970) and was used to describe games that were designed with explicit educational (or training) goals. At the time, Abt focused on all types of games, not just electronic-based games and he emphasized that, though primarily educational in purpose, serious games must retain the playfulness and entertainment value found in traditional games.

Despite Abt's investigation into how games could improve the educational experience over 40 years ago, the formal research of serious games has only recently

## REAL-TIME PREDICTORS OF ENGAGEMENT

begun to gain headway. In 2002 the Serious Games Initiative (SGI, 2011) was founded to further the research and support of integrating video game frameworks into educational tools. In 2010, the United States White House began a games-based initiative to boost student competencies in the fields of science, technology, engineering, and math (STEM) (ESA, 2009). Video games were identified as a way to develop effective learning environments that could increase student motivation, engagement, and learning (ESA, 2009).

Video game design is also a rapidly growing area of research for human factors psychologists interested in training, simulation, and engagement (Adelson, 1992; Ballew & Jones, 2006; Malone, 1984; Pavlas, Heyne, Bedwell, Lazzara, & Salas, 2010). This research on video games as a paradigm for training has often hinged on the contention that video game environments are engaging due to their unique combination of narrative and play (Admiraal, Huizenga, Akkerman, & Dam, 2011; Burgos, Tattersall, & Koper, 2007; Ciavarro, Dobson, & Goodman, 2008) and provide simulated experiences that are often quickly embraced by the user despite the amount of complex information they may contain (Herrington, Oliver, & Reeves, 2003).

### *Designing Effective Video Games*

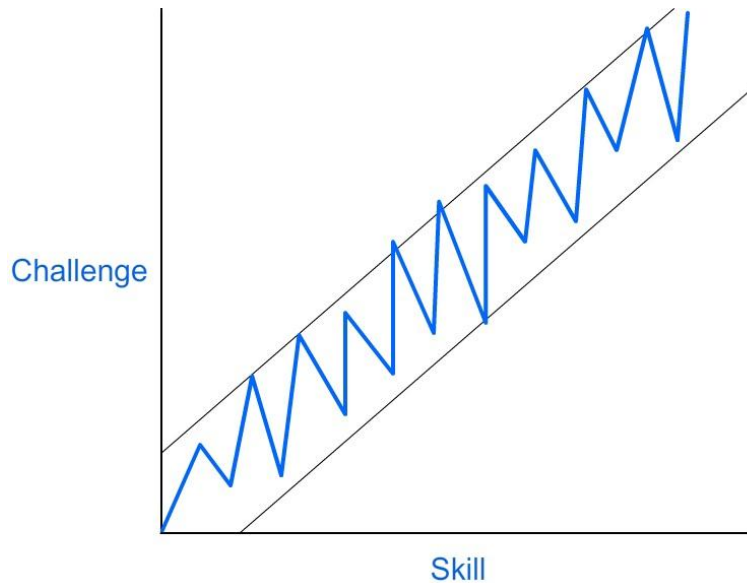
When discussing effective video game design methods, it is important to consider both the playability of the game and the gameplay (player experience). Playability

## REAL-TIME PREDICTORS OF ENGAGEMENT

describes the usability of the game. If a player is unable to interact with a game's interface and controls then their experience will be hampered. In recent years, efforts to formalize video game evaluation techniques have led game designers to adopt human factors methods such as heuristic evaluations. (Desurvire, Caplan, & Toth, 2004; Nacke et al., 2009). Though without good playability, a game will not be able to realize its full gameplay potential, the scope of this manuscript focuses on gameplay rather than on the game's usability.

Gameplay describes the interaction between the player and the game (Nacke, et al., 2009). The ideal type of gameplay was elegantly summarized by the founder of Atari, Inc., Nolan Bushnell, who in 1971 said, "A good game should be easy to learn, but difficult to master" (as cited in Malone, 1984). One of the most important aspects of good gameplay is the way the game balances challenge with the player's current skill level (Malone, 1980, 1984). Simply put, games that are too easy can become boring, and games that are too difficult can be frustrating (Hunicke, 2005).

## REAL-TIME PREDICTORS OF ENGAGEMENT



*Figure 1.* A Non-linear increase in Challenge and Skill

As depicted in Figure 1, this challenge/skill balance should typically be nonlinear to avoid a player's ability to predict the next challenge's degree of difficulty (Rabin, 2005). By mixing easier tasks with increasingly more difficult tasks, a game may be able to maintain player curiosity and interest while also allowing them to decompress after a particularly challenging task. For example, if a player becomes cognitively overloaded, an easier task will allow them to maintain gameplay while releasing cognitive resources so they can face the next difficult challenge. It is important to design an appropriate difficulty curve that challenges users enough to keep them engaged, but not too much so

## REAL-TIME PREDICTORS OF ENGAGEMENT

as to produce frustration (Aponte, et al., 2009); something Hunicke (2005) called a comfort policy.

Video games attempt to keep users in a comfort zone by manipulating difficulty using one of several traditional gameplay approaches: static, user-controlled, and adaptive (Gilleade & Dix, 2004; Gilleade, et al., 2005; Novak, 2008; Rabin, 2005). Similarly, one of the defining characteristics of electronic content delivery compared to traditional classroom training is the degree of control a user has over their training experience (Chen & Macredie, 2002; Lawless & Brown, 1997). Either the training program does not provide any control (static), or it provides internal control (user-controlled) where the user is in complete control of navigating and choosing their learning path, or it provides external control (adaptive) where the training program is in complete control of the content delivery (Milheim & Azbell, 1988).

Static gameplay uses a linear progression of difficulty based on the expected progress of an idealized player (Gilleade & Dix, 2004). This can be very effective for simple games, but complex games that require large amounts of schema development can quickly alienate users who do not advance as linearly as the developer-defined prototypical player.

## REAL-TIME PREDICTORS OF ENGAGEMENT

User-controlled gameplay can take several different forms. The most basic form allows users to select a level of difficulty (e.g., easy, medium, hard) at the onset of the game (choice). After selecting the difficulty level, the game falls back into a linear, static progression of difficulty as would be seen in a predetermined gameplay design. This is still a relatively inflexible type of gameplay, and if the user does not select the correct level of difficulty, they may become disengaged. However, user-controlled gameplay can also take the form of a series of convexities (Rabin, 2005). A convexity is a decision point within a game where a user can choose one of several paths. These decision points are non-trivial (Novak, 2008) and typically influence the level of difficulty a player will experience. Some of the paths may lead to easier or more difficult tasks, thus indirectly providing the user some degree of control over their cognitive load. With a convexity design, instead of a single choice at the beginning of the game, users are able to adjust difficulty at multiple points over the course of gameplay. In user controlled systems, higher engagement through perceived sense of control might lead to higher performance and potentially even increased intrinsic motivation (Cordova & Lepper, 1996). On the downside, if the user doesn't select the correct level of challenge, they may perform sub-optimally.

Adaptive gameplay relies on computational algorithms to determine the appropriate level of difficulty a player should experience. This can occur either between

## REAL-TIME PREDICTORS OF ENGAGEMENT

levels, or in more complex instances, during gameplay. Adaptive game balancing starts with the system's ability to recognize when a player is either over-challenged or under-challenged. This is typically carried out by analyzing a player's performance. For example, if a player is unable to get past a particular point in a game, certain elements may be manipulated to decrease the level of difficulty, such as reducing the number of spawning enemies. Alternatively, if the game is comprised of discrete levels, such as in puzzle-based games, the next level presented could simply be an easier level. It's important to note that adaptive gameplay should be designed carefully so as to prevent players from predicting how the adaptive algorithm will behave (Schell, 2008). If a player is aware of how the adaptive algorithm works, they could either "game" the system, or worse, become disengaged due to a lack of immersion once the reality of the world is spoiled (Schell, 2008). In adaptive systems, optimized challenge levels may lead to higher engagement when the game mechanics retain a degree of program control based on the user's real-time interactions with the system. (Gilleade & Dix, 2004). However according to Flow Theory (Csikszentmihalyi, 1988), the loss of user control may contribute to lower levels of engagement which, in turn, may lead to lower performance.

It is likely that adaptive gameplay will provide the greatest opportunity for player engagement because it can seamlessly be integrated into the game engine without disrupting the player's focus on gameplay. Predetermined, static gameplay is unlikely to

## REAL-TIME PREDICTORS OF ENGAGEMENT

result in engagement because it is rigid and does not adjust to the unique needs of each player. Similarly, user-controlled gameplay may provide a more refined gaming experience but it does put the burden of choice on the player. For some games, these choices may be designed into the gameplay to provide a more natural and seamless experience. However, for other games such as puzzle games which are more closely related to online training, asking the user to select a new degree of difficulty after each level may unnecessarily increase cognitive load and create a point of disengagement.

To create an adaptive gaming environment with the primary goal of increasing and sustaining engagement, it is essential to monitor engagement during gameplay. In the past, adaptive systems have been built around real-time performance measures. When a player is doing well, the adaptive engine might provide a more difficult game level; spawning more enemies, for example, or creating more complex puzzles to complete. However, there are very few empirical research studies on measuring real-time engagement from a cognitive standpoint and consequently little is known about how to effectively sustain engagement during gameplay.

### *Flow Theory*

Flow Theory's (Csikszentmihalyi, 1988) principles have high face-validity and have proven to be one of more thorough and practical frameworks from which to begin the study of video game behavior and engagement (O'Brien & Toms, 2008; Pavlas, et al.,

## REAL-TIME PREDICTORS OF ENGAGEMENT

2010). Being in the flow can be thought of as an optimal experience that includes feelings of exhilaration and deep enjoyment (Csikszentmihalyi, 1990). People in the flow are intrinsically motivated and commonly report: focused concentration, feelings of control, and a lack of awareness of time (Csikszentmihalyi, 1988). As can be seen in Table 2, some of the dimensions of the Flow construct include two task-oriented requirements and seven user-oriented experiences.

Table 2. *The Dimensions of Flow Theory*

Dimension	Reflected in the...
Clear goal	Task
Timely and unambiguous feedback	Task
Balance between perceived skill and perceived task challenges	User
Sense of control over the task outcome	User
Complete focused concentration on task	User
Loss of awareness of self and the outside world	User
Distorted sense of time	User

## REAL-TIME PREDICTORS OF ENGAGEMENT

Flow is more likely to occur when a person is challenged by tasks that they judge worthwhile to overcome. These experiences are commonly found in video games (Weibel, Wissmath, Habegger, Steiner, & Groner, 2008) which typically provide a progression of challenging but obtainable experiences (Rabin, 2005; Schell, 2008) that can be balanced with individual skill levels.

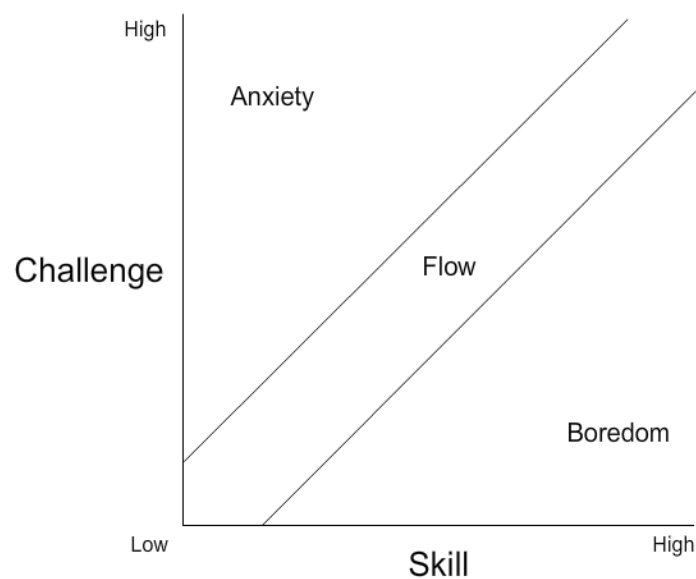


Figure 2. Flow Theory's Interaction between Challenge and Skill

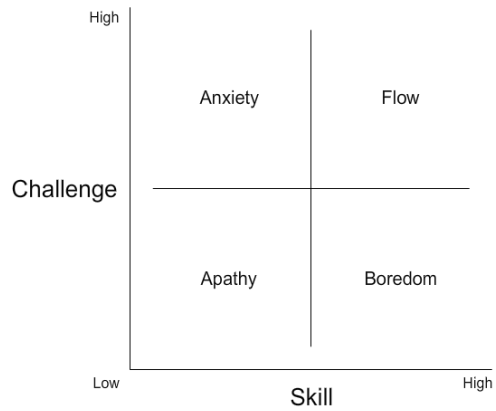
Figure 2 shows the original Flow model (Csikszentmihalyi, 1975). In this model, flow can occur whenever a balance between skill and challenge is achieved. A person can fall into the *Anxiety Area* if their skill level is not matched with a comparable difficulty

## REAL-TIME PREDICTORS OF ENGAGEMENT

level while participating in a task such as a video game. People usually begin a task with a low set of skills therefore their task should match their skill level with an appropriately low set of challenges. According to Flow Theory, as a person progresses through a task, their flow will more likely be maintained if their task difficulty manageably increases. Task difficulty can be managed through algorithmic adaptive methods, user-controlled methods, or static progressions based on a “prototypical” user. One of the goals of this research is to determine the most effective method to manage task difficulty.

Csikszentmihalyi later adapted his original Flow model to include a state of apathy (Csikszentmihalyi & Csikszentmihalyi, 1988). This new four-channel Flow model (Figure 3) shows that a balance between skill and challenge is no longer sufficient to allow a person to achieve a state of flow. This further refinement is based on data from Csikszentmihalyi’s experience sampling method studies where it was found that people in a low skill/low challenge situation were not achieving a state of flow in their day-to-day lives. He found that, in order to achieve flow, a person must be in a situation where they experience both a level of skill and challenge that are *higher* than their normal daily activities.

## REAL-TIME PREDICTORS OF ENGAGEMENT



*Figure 3.* The Four Channel Flow Model

Csikszentmihalyi and Nakamura (1989) further suggested that investigators should experiment with more than four channels and, as Figure 4 shows, an eight-channel model has emerged (Ellis & Voelkl, 1994).

## REAL-TIME PREDICTORS OF ENGAGEMENT

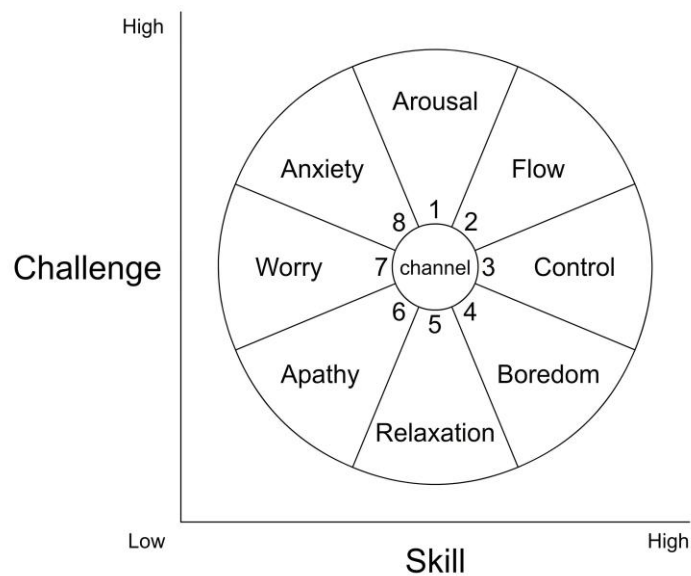


Figure 4. The Eight Channel Flow Model

These additions to the flow model are certainly useful when investigating how people respond to common day-to-day tasks, however, for novel tasks presented in video games and online training programs, these more-recent models are less useful. For the purpose of this research, the original Flow model was adapted since it includes the possibility of a person experiencing Flow in a low-skill, low-challenge situation. Rather than use Flow Theory's description of this state as *Apathy*, defined as "a lack of feeling or emotion" (Gove, 1963, p. 40), the more appropriate term *Effortlessness* will be used.

## REAL-TIME PREDICTORS OF ENGAGEMENT

A review of Flow literature shows that definitions of Flow are not always the same among researchers (Novak, Hoffman, & Yung, 2000) which may be due, in part, to the multidisciplinary nature and application of Flow Theory. For the purposes of this paper, Flow is operationally defined as a cognitive state that is achieved through experiencing a balance between skill and challenge while participating in a challenging *novel* task. A novel task in this context is defined as any task that is perceived to be challenging and requires a degree of skill that can be learned as a function of experience with the task. Based on this definition, an adaptation of Csikszentmihalyi's original Flow model will be used because the nature of participating in a novel task removes the need to compare levels of skill and challenge with daily activities. Furthermore, based on the use of Flow Theory for game design (Rabin, 2005), *Anxiety*, defined as a "painful or apprehensive uneasiness of mind over an impending or anticipated ill." (Gove, 1963, p. 40) was replaced with *Frustration* which could result in cognitive overload and negative affect when a person feels "a deep chronic sense or state of insecurity and dissatisfaction arising from unresolved problems" (Gove, 1963, p. 336). The resulting Flow model for this research can be seen in Figure 5.

## REAL-TIME PREDICTORS OF ENGAGEMENT

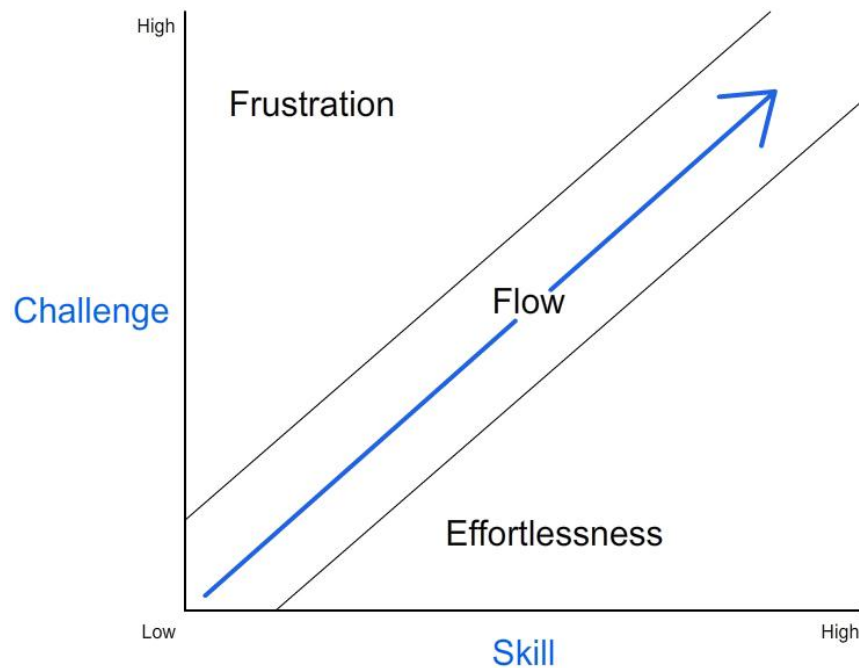


Figure 5. Adapted Model of Flow used in the Current Study

### *Cognitive Load Theory*

As described above, challenge can be thought of as the level of cognitive load a person experiences during a task. Cognitive Load Theory (CLT) provides a framework for understanding the psychological relationship of challenge and skill. Similar to the predictions of Flow Theory, CLT also hypothesizes that an optimal cognitive state for learning will exist when the learner's ability (skill) and the complexity (challenge) of their learning task are appropriately balanced (Moreno & Park, 2010). CLT has been used

## REAL-TIME PREDICTORS OF ENGAGEMENT

to guide the design and selection of multimedia and instructional elements that will be more likely to maintain an optimal amount of cognitive load during training (Kalyuga, Chandler, & Sweller, 2004).

Research involving CLT focuses on three types of load: extraneous, intrinsic, and germane, and assumes that all three loads are additive (Paas, Renkl, & Sweller, 2003). Intrinsic load is of most interest to this study because it represents the load generated due to the spread between a person's current skill and the task challenges. Intrinsic load is likely to be relatively uniform across users when facing a novel task, for example, when beginning a new video game. Since intrinsic load depends both on the learner's initial expertise and their developing schemata, it is important to remember that the relative challenge, and thus their cognitive load, will change as they develop expertise. Based on CLT, intrinsic load is likely to be highest when the ratio of challenge to skill is greatest and lowest where the ratio has reversed. Based on Flow Theory (Csikszentmihalyi, 1990), it is hypothesized that there would be an optimal balance of challenge to skill in a learning or training situation that would represent a middle range of intrinsic load. By extension, it is believed that as novices gain experience in a new game, the challenge level will need to be raised to keep individuals in this optimal range. Outside this middle range, learning would not be optimized because either cognitive capacity has been

## REAL-TIME PREDICTORS OF ENGAGEMENT

overloaded (causing frustration) or underutilized (causing boredom, or a state of effortlessness).

Germane cognitive load (Paas & van Merriënboer, 1994) can be thought of as the amount of cognitive effort a person expends on a task. Another way of considering germane load is to think of it as a form of intrinsic motivation (Mayer & Moreno, 2010), and that being motivated toward task completion requires a higher degree of cognitive processing. Intrinsic load needs to be high enough to provide meaningful challenge in order to provide the intrinsic motivation to raise germane load. However, if intrinsic load is too high, it will remove the capacity for additional germane load. In keeping with Flow Theory's terminology, particularly high intrinsic load represents Frustration, and particularly low intrinsic load represents Effortlessness; neither of which is likely to spur high germane load. Therefore, germane load provides a link to the intrinsic motivation and positive affect required for a person to enter the flow, and thus become actively engaged (see Figure 6). Currently no well-established measure of germane load exists (Brunken, Paas, & Moreno, 2010).

## REAL-TIME PREDICTORS OF ENGAGEMENT

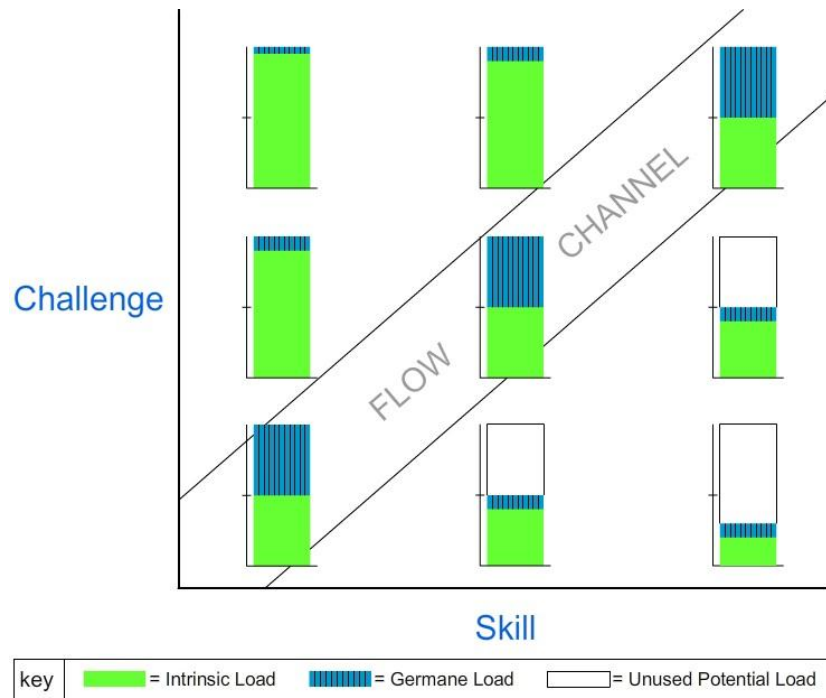


Figure 6. Cognitive Load Theory and Flow Theory

### *Measuring Cognitive Load*

Cognitive load is commonly measured using subjective measurement methods such as the Subjective Assessment Technique (SWAT) (Reid & Nygren, 1988), the Workload Profile (Tsang & Velazquez, 1996), and the NASA-TLX (Hart & Staveland, 1988). The NASA-TLX has been used extensively to measure overall mental workload, in addition to diagnosing individual affective and cognitive components of load (Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988; Wiebe, Roberts, & Behrend, 2010).

## REAL-TIME PREDICTORS OF ENGAGEMENT

Though such subjective measurements of cognitive load are typically easy to implement and provide straightforward data for analysis, they do have limitations. Some of these problems include an unclear relationship between mental effort and actual cognitive load (Brunken, Plass, & Leutner, 2003), and the fact that they must be delivered as a *post hoc* assessment prevents the identification of the cause of the reported cognitive load (Brunken, Seufert, & Paas, 2010). To overcome some of these limitations, a secondary task methodology for measuring cognitive load has been adopted by some researchers as a relatively reliable and acceptable cognitive load measurement method (Brunken, Paas, et al., 2010; Brunken, et al., 2003; Ogden, Levine, & Eisner, 1979).

### *Linking Flow Theory and CLT with Engagement*

The next step past Figure 6, characterizing the relationship between Flow Theory and CLT, is to describe how all three concepts—cognitive load, Flow, and engagement—are linked. Flow and engagement are commonly associated with each other in video game and online training research (Brockmyer, et al., 2009; Chapman, Sanjeebhan, & Jane, 1999). Though Flow has been measured in video game and online training research, this research will shift focus to the construct of engagement and use Flow Theory in a supporting role. However, to better understand the dynamic nature of engagement, it is important to first take a high-level look at the relationship between engagement and flow (Figure 7).

## REAL-TIME PREDICTORS OF ENGAGEMENT

In order to become engaged, a person must be motivated before they allocate their cognitive resources towards focusing on a task (Przybylski, Rigby, & Ryan, 2010; Webster, Trevino, & Ryan, 1993). Once engaged, a person has the opportunity to enter a flow state. The flow state can be thought of as a function of the active response a person has to the challenges of the task. In other words, to be in the Flow, a person must be actively participating and not mindlessly involved. At times during the task, a person can become disengaged (O'Brien & Toms, 2008). During these times, it follows that the person is also released from their flow state. After a period of disengagement, a person may have the opportunity to become reengaged in the task and subsequently, they may enter back into the flow.

When a person willingly participates in a task, it cannot be assumed that they are pursuing their optimal performance. This is an important point to consider when designing user-controlled game environments since players may not select a game level that provides the appropriate level of challenge. Some people have a low desire for control and therefore may select less-challenging levels, while those with a high desire for control may attempt levels that are too difficult (Burger, 1985). Flow, on the other hand, refers to a higher state of what is typically referred to as engagement where challenges and skills are sought to be balanced. In this state, a person is more likely to perform at their optimal level. Being in the flow creates an equilibrium where the

## REAL-TIME PREDICTORS OF ENGAGEMENT

interdependent skills and task challenges are actively adjusted to maintain a balance.

When in the flow, as a person's skill increases it is proposed that they will seek a more difficult, but obtainable, challenge to maintain this balance.

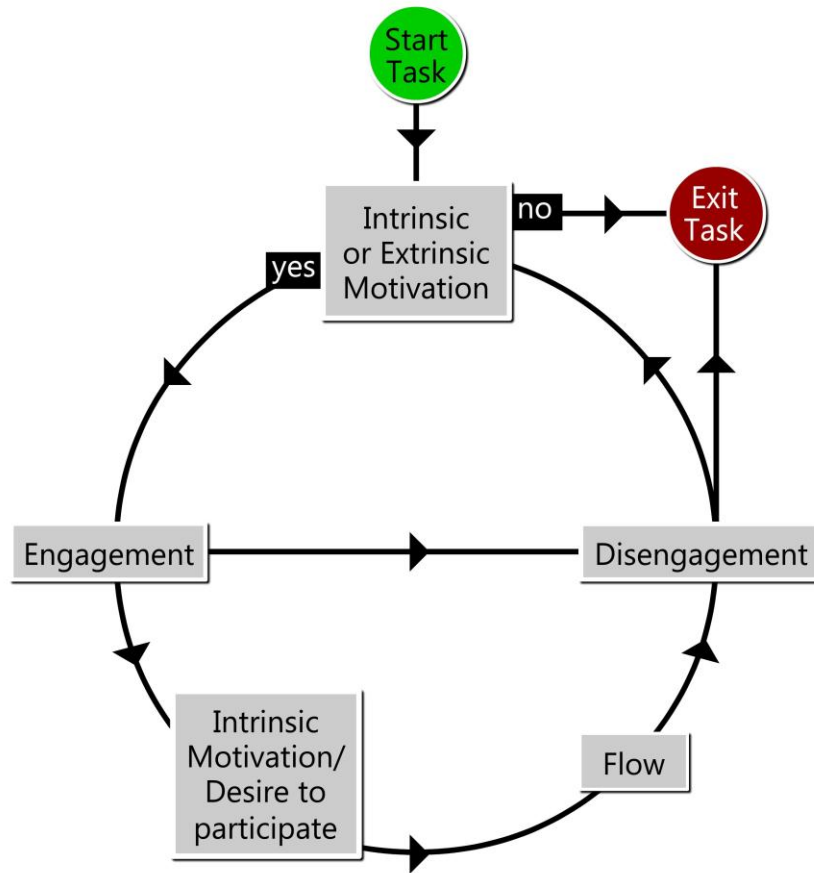


Figure 7. The Engagement / Flow Cycle

## REAL-TIME PREDICTORS OF ENGAGEMENT

As Figure 7 shows, the relationship between engagement and flow can be thought of as part of a continuous cycle. In order to be engaged in a task, a person must typically be either intrinsically or extrinsically motivated before they make the decision to allocate their cognitive resources towards focusing on the task. Once engaged, (i.e. motivated and focused on the task) a person may become intrinsically motivated (i.e., willing to participate) in the task and therefore the opportunity to enter into a higher level of engagement is manifested.

Many different definitions of engagement have been provided in video game and online training literature, making it difficult to compare results and design recommendations. For example, Barker and King (1993) investigated computer-based training materials, and found that engagement was one of the most influential factors when rating overall quality. Engagement was never defined, but an engaging product was described in the questionnaire as an interesting, or involving, product based on motivating, enjoyable or challenging factors. Indeed, the distinction between engagement and Flow was equally ambiguous, with authors seemingly using these words interchangeably. There are some exceptions though, for example, Webster and Ho (1997) discussed Flow and engagement in terms of multimedia presentations. They acknowledge that there is a large overlap between flow and engagement but they differentiate the two by saying that flow requires perceived control, while engagement does not.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Engagement*

For the purposes of this research, and building on a previous working definition (Sharek, 2010), engagement is defined as the active state of seeking out a challenge, where challenge is operationally defined as cognitive load. In other words, engagement is a function of the amount of cognitive load a person willingly desires. This *willingness to participate* is similar to Laurel's (1993) description of engagement which includes the willing suspension of disbelief when watching a film. However, a willingness to participate or interact is not enough for a person to be considered *engaged* in an activity that requires learning. For example, research has shown that people participating in what was thought to be a boring task were in fact experiencing positive affect, though their skills were not increasing (Sharek, 2010); they were content with not seeking out new challenges, therefore their skills stagnated.

If cognitively overloaded, a person may become frustrated and disengage, and if cognitively underloaded, a person may become bored and similarly disengage. The desire to be challenged—i.e., to experience a meaningful level of cognitive load—can fluctuate throughout a task depending on a person's level of intrinsic and germane cognitive load. In fact, no matter how engaged in a task a person may be, they are likely to eventually reach a point of disengagement (O'Brien & Toms, 2008). There are three main categories for why disengagement may occur: *external factors* (e.g., telephone ringing, or when a

## REAL-TIME PREDICTORS OF ENGAGEMENT

person becomes hungry), *internal factors* (e.g., inherent pauses or waiting times within a task), or *psychological factors* (e.g., cognitive overload or underload). This research focuses on the internal factors as they relate to the playability and mechanics of a video game, and the psychological factors as they relate to the affective and cognitive response a person has to a video game.

Based on the design of the game used in this research (and typical of many types of gameplay) during these transition periods of disengagement, a person will have the option of re-engaging in their task or exiting it completely. For designers to develop tasks that encourage users to remain engaged, they must understand what internal psychological factors affect the decisions made during the period of disengagement and what factors predict successful re-engagement. It is likely that increased engagement during gameplay creates a type of psychological momentum that propels individuals through natural points of disengagement back into engaged gameplay again.

The implications for understanding this momentum-through-engagement process can lead to: greater user interest and learning retention (Adelson, 1992; Webster & Ho, 1997), pleasure (Charlton & Danforth, 2007), and an increase in the overall direction and intensity of a person's learning experience—a characteristic called training motivation by Colquitt, LePine, and Noe (2000).

## REAL-TIME PREDICTORS OF ENGAGEMENT

When considering the measurement of engagement, a person's level of affect must be taken into account, as this provides insight into their willingness to participate.

Additionally, the degree of challenge a person seeks out is equally important to capture. It stands to reason that a method of measuring both affect and challenge (cognitive load) in real-time might provide a useful engagement measure.

### *Measuring Engagement*

Attempts have been made to measure Flow (Jackson & Marsh, 1996) and engagement using self-report questionnaires that are administered after a given task (e.g., Brockmyer, et al., 2009; O'Brien & Toms, 2008; O'Brien & Toms, 2010). While useful, these measures lack a degree of diagnosticity because they are not able to capture levels of engagement at different stages during a task. This lack of diagnosticity prevents the kind of actionable data that game designers need during play-testing (Davis, Steury, & Pagulayan, 2005; Dixit & Youngblood, 2008; Drachen & Canossa, 2009).

A possible solution may be found by measuring a person's awareness of the passage of time as type of a secondary task (McCarthy & Wright, 2004; O'Brien & Toms, 2008; Sharek, 2010; Skadberg & Kimmel, 2004). Skadberg and Kimmel (2004) found that people's self-reported sense of a distortion of time was an accurate measure of Flow. Flow Theory predicts that when a person engages in a task, they may enter a flow state and experience positive affect and a distorted sense of time. However, Flow Theory does

## REAL-TIME PREDICTORS OF ENGAGEMENT

not provide an explicit cognitive or affective mechanism for how or why this happens. When a person is not engaged in a task, they will be more likely to be aware of the passage of time. This awareness may be physically manifested in observable actions, such as a person looking at their watch when they become bored or lose interest in a task. When a person is in this state, it stands to reason that they are not engaged. This idea is further supported by the idea that a person's willingness to maintain task engagement is related to their perception of time, and that a person's sense of time can change depending on their level of engagement (McCarthy & Wright, 2004).

Previous research has been conducted on such a measure of the awareness of the passage of time with promising but mixed results (Sharek, 2010). A Game-clock mechanism was developed and imbedded in a video game. Before playing the game, participants were told that they must play for a minimum amount of time, but the amount of time was not revealed. They were then told that they could check the Game-clock, which would reveal whether or not the minimum amount of time had been met. The development of the Game-clock was based on Flow Theory and prospective memory research and it was initially proposed that the Game-clock would reveal a person's awareness of the passage of time, thus indicating their level of engagement in the video game. The more times they checked the Game-clock, the more aware they were of the passage of time and therefore the less they were engaged in the game. The results

## REAL-TIME PREDICTORS OF ENGAGEMENT

revealed that during times of disengagement (during game level intermissions) those who were in a frustrating game level condition (high challenge-low skill) would check the Game-clock significantly more times than during gameplay. It is likely that when frustrated players were cognitively overloaded during gameplay, they were unlikely to devote cognitive resources to check the Game-clock even though they wanted to exit the game. Once their cognitive resources were released during intermission, results showed they would check the Game-clock at a higher rate than those in the flow condition.

Rather than a measure of engagement, the Game-clock showed more promise as a measure (during intermissions) of affect and desire to play the game. Since one dimension of engagement is the willingness to participate, the Game-clock may still prove useful as part of a real-time measure of engagement. The other dimension of engagement is cognitive load and, as has been described in the Cognitive Load Theory section, could be potentially measured during gameplay via a secondary task. Together with a performance measure as a means to account for potential effortlessness, the combination of the Game-clock during intermission and a secondary task during gameplay may prove to be a viable real-time measure of engagement.

### *Summary*

Based on Cognitive Load Theory and Flow Theory, engagement has been defined as the willingness to participate in tasks that require an optimum level of cognitive load

## REAL-TIME PREDICTORS OF ENGAGEMENT

(not too easy, and not too difficult). The act of seeking out cognitively demanding tasks and mastering them can lead to positive affect and sustained task interest. When engaged, people will be more likely to learn new information as well as remain involved in a task. The investigation of possible real-time measures of positive affect, in the form of the willingness to seek out cognitively demanding tasks, may provide insight into both the development of video games and online training environments.

Methods for developing engaging video games may also provide possible solutions to increasing online training engagement through the use of the serious games design framework. Video game designers often have turned to Flow Theory as a guide for the gameplay development process. Specifically, Flow Theory's draw for game designers has been the proposition that a balance between skill and challenge will likely lead to optimal feelings of enjoyment and satisfaction. When considering learning, however, understanding how to design for this type of positive affect is only half of the equation. Learning is more likely to take place when a person is experiencing an optimal level of cognitive load. Thankfully, Flow Theory's model of challenge and skill fits this requirement: a person seeking the balance of skill and challenge described in Flow Theory will also experience an optimal level of cognitive load through a balance of intrinsic and germane load components, resulting in engagement. Both game design heuristics and Flow Theory speak to the important role of user control in the experience.

## REAL-TIME PREDICTORS OF ENGAGEMENT

When integrated with balancing skill and challenge, there are a number of possible combinations of user and system control of this balance that may produce an optimal experience of affect and cognitive load.

This research seeks to address the following research questions:

R1: Can engagement during gameplay, based on the combination of a willingness to participate and a desire to experience cognitive load, be characterized by measures of cognitive load and affect?

R2: Can engagement be measured in real-time? Specifically, can a person's desire to experience an optimum level of cognitive load be measured and assessed while playing a video game?

R3: Can a real-time measure of engagement be used to develop an adaptive video game?

R4: What are the performance differences between linear, user-controlled (choice), and adaptive, games?

### Study One

In a previous study, the use of Flow Theory was used to guide the development of BlockWalk (Sharek, 2009), a video game specifically designed as an experimental

## REAL-TIME PREDICTORS OF ENGAGEMENT

stimuli (Sharek & Wiebe, 2011) for the purpose of investigating the engagement construct. This current study focused on level building (Novak, 2008) to further refine the design of the previous video game in order to produce GridBlocker (Sharek, 2011), a more robust gaming platform that was used in Studies Two and Three.

The number of video game levels were increased from ~50 levels to 108 levels at varying difficulties to increase the range and granularity of difficulty a player could experience. Performance was captured for each newly designed level. These performance data were then used in a manipulation check against the self-report measures described below. Based on the previous study, it was expected that difficulty ratings would correlate with performance measures.

In addition to the performance measures, three self-report game level performance ratings were collected for each level (Difficulty, Challenging, Frustration). The primary rating of interest was Difficulty, but including the Frustration and Challenging measures may have produced additional affective data that could be used to influence the overall difficulty ratings of each level. Depending on the results, one or all of these data will be used along with the performance metrics to order each level in terms of difficulty. Additionally, a single item self-report measure of “fun” was collected as a potential functional measure of affect in gaming environments. Depending on the results, this measure of fun may be used to provide an affective rating schema for each level.

## REAL-TIME PREDICTORS OF ENGAGEMENT

While this study could be used to partially answer Research Question 1, this study is using a direct and overt, reflective self-report measure of both difficulty and fun (i.e., positive affect). However, the goal through Studies 2 and 3 was to develop an embedded measure that does not require direct responses to these states since real-time data collection methods—a goal of this work—may provide a more scalable solution to creating adaptive games.

### *Independent Variables*

The research design was a single-condition study. All participants were provided with the same video game environment for which to play 12 randomly assigned game levels.

### *Individual Game Level Performance Dependent Variables*

Based on a previous validation study of this game (Sharek & Wiebe, 2011) the following individual game level performance dependent variables were captured and used in a manipulation check against the self-report measures of difficulty, challenge, frustration, and fun (described below).

*Over-moves.* This measure was not included in the previous study. It represents the number of moves the player made over and beyond the theoretical minimum number of moves required to complete the level.

## REAL-TIME PREDICTORS OF ENGAGEMENT

*Directions.* The number of times the player changed the block's direction. For example, if the player moved the block to the left five spaces and then up for two spaces, the number of directions was recorded as two.

*Errors.* The number of times the player moved the block off of the game board (restarting the level).

*Time.* The amount of time the player spent playing each level. Timing will commence once each level loads and the timer will stop once the player has completed the level by successfully placing the block over the goal.

### *Subjective Ratings Dependent Variables*

Partially based on Schell's (2008) recommendations, players were asked to answer a short, four-item questionnaire regarding their gaming experience after each level. (**Appendix A**).

*Difficulty.* "How difficult was the level you just played?" (1 = Very Easy, 10 = Very Difficult)

*Frustration.* "How frustrating was the level you just played?" (1 = Not at all Frustrating, 10 = Very Frustrating)

## REAL-TIME PREDICTORS OF ENGAGEMENT

*Challenge.* “How challenging was the level you just played?” (1 = Not at all Challenging, 10 = Very Challenging)

*Fun.* “How fun was the level you just played?” (1 = Not at all Fun, 10 = Very Fun)

### Study One Method

#### *Participants*

A total of 308 people (54% male, 46% female; mean age = 30,  $SD = 11.19$ ) were recruited from Amazon’s Mechanical Turk (Amazon.com, 2011). The Mechanical Turk is an online crowdsourcing tool that connects people (workers) to tasks (HITs) that can be completed online. It has been shown to be a successful participant recruitment tool for online behavioral research (Behrend, Sharek, Meade, & Wiebe, 2011).

#### *Materials and Apparatus*

The experiment was hosted at [www.gridblocker.com/space/rateGame](http://www.gridblocker.com/space/rateGame), a Website maintained by the author developed specifically for hosting the GridBlocker game (Sharek, 2011). The Website uses a MySQL database and a PHP server-side scripting language to store participant responses. The experiment could be accessed by participants from any computer with Internet access. The computer requirements included a minimum

## REAL-TIME PREDICTORS OF ENGAGEMENT

computer monitor resolution of 1000x600 and an installed version of the Adobe Flash player v.10 or greater.

### *GridBlocker*

The goal of Gridblocker is to move a rectangular block around an isometric tile-based game board so that the block ends up standing over a goal. As Figure 8 shows, there are three main types of movements that change the way the block lies on the game board. Positioning the block around the game board so that it ends up standing over the goal can become tricky when the game board is laid out in a complex configuration. Sometimes only one sequence of moves exists to successfully traverse certain paths. This can become especially difficult when there appears to be multiple solutions. As expertise increases, most players learn certain algorithmic methods for sets of sub-moves that can be used for positioning the block in desired locations.

## REAL-TIME PREDICTORS OF ENGAGEMENT

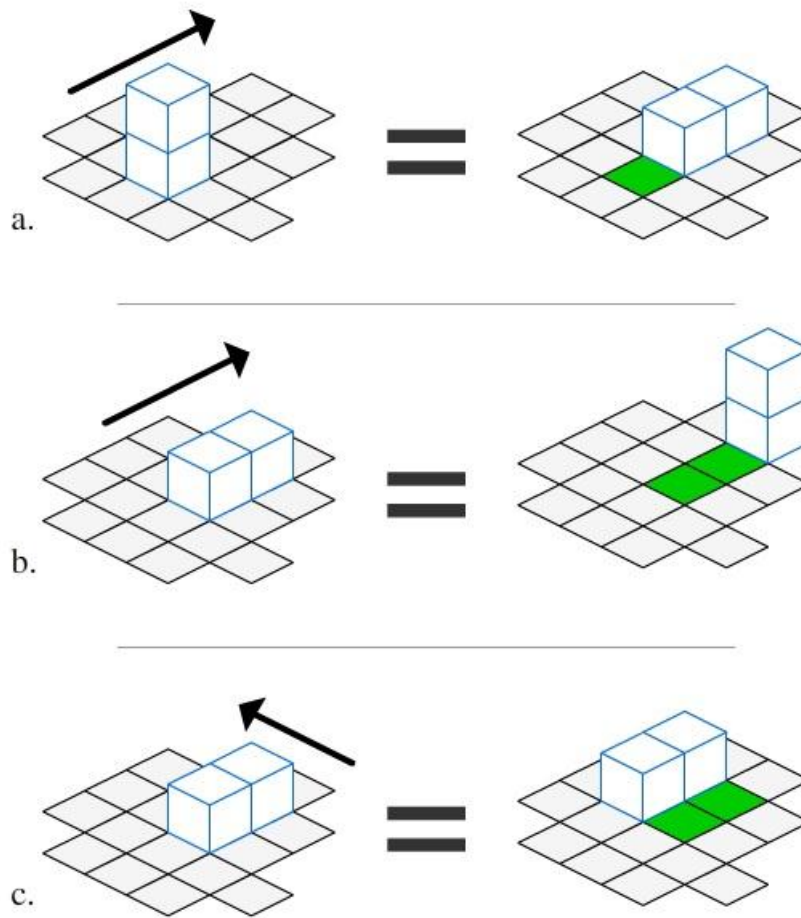


Figure 8. Moving the Block Around the Game Board in the GridBlocker Game

- a. When the block is standing up, any horizontal or vertical movement will cause it to fall down. Notice how the block no longer remains over the starting tile position (shown in green)
- b. When the block is laying down, it can be moved to a standing position if the movement is along the longer axis of the block. A combination of Move a. and Move b. will move the block three squares along an axis
- c. If the block is laying down and moved along block's shorter axis, it will continue to remain laying down and only move one square at a time. Combinations of Move a. and Move c. can move and offset the block.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Procedure*

After consenting to an online IRB Form (**Appendix B**) participants were asked to enter in a username. This was used to generate a secure 32 digit hexadecimal unique user ID (UID) using an MD5 cryptographic hash function. This UID is used to link individual participant performance and ratings throughout the experiment in the database tables.

Before beginning, participants were given instructions explaining how to play the game (Appendix A). The following story was included in these instructions to explain the premise of the game:

“You are a galactic aid worker on a floating cargo deck docked in deep space. For each level, your job is to position each cargo block upright on the transport tile. The blocks contain much needed supplies for the stranded people of Cubex-5. Normally, a cargo robot would be used to transport the blocks, but an asteroid shower has disabled the robot control tower. Since you are in deep space resources are limited, leaving you to move the blocks manually. Every time you move the block, you could potentially be damaging the supplies so be sure to move the block in as few moves as possible. Good luck!”

After reading the instructions and game story, participants were randomly assigned 12 of the 108 possible levels to play. After each level was played, participants were asked to rate their experience with the level. After all 12 levels were completed, participants were asked to complete a short demographics survey.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### Study One Results

Before any analyses were conducted, data from the first level each player played was treated as practice and removed from analysis. Table 3 shows descriptive statistics of all performance and self-report measures. Appendix C can be referred to for distributions.

Table 3. *Mean Scores for all 108 levels*

Measures	Mean	Median	SD	Min	Max
Time	96.11	68.93	94.04	3.55	404.76
Errors	3.25	2.37	3.40	0	16.43
Directions	42.54	33.19	41.22	1.97	192.74
Minimum moves	14.36	13.00	8.09	2.00	40.00
Over-moves	63.51	42.81	65.39	.50	255.00
Skipped	.10	.03	.15	.00	.83
Difficult	4.51	4.60	1.83	1.21	8.54
Frustrating	3.67	3.50	1.48	1.28	7.20
Challenging	4.85	5.01	1.71	1.34	8.54
Fun	5.56	5.60	.55	3.90	6.65

Pearson correlations were conducted to analyze the relationship between the performance and self-report data (Table 4).

## REAL-TIME PREDICTORS OF ENGAGEMENT

Table 4. *Bivariate correlations*

	Directions	Time	Errors	Over-moves	Difficulty	Frustrating	Challenging
Time	.98**						
Errors	.92**	.95**					
Over-moves	.96**	.97**	.90**				
Difficulty	.89**	.91**	.86**	.91**			
Frustrating	.91**	.94**	.89**	.93**	.98**		
Challenging	.87**	.89**	.84**	.89**	.99**	.97**	
Fun	.05	.04	.05	.03	.30*	.18	.33**

Note:  $n = 108$ , \* $p < .01$ , \*\* $p < .001$

The Frustrating and Challenging self-report measures were dropped from further analysis because they were highly correlated with Difficulty, the primary self-report measure of interest. The Fun self-report measure was dropped from analyses because it produced a small range of scores and lacked significant correlation with game performance metrics.

A hierarchical regression analysis was conducted to examine the extent to which three sets of performance predictors accounted for self-report difficulty. The first block included the amount of time it took player's to complete each level. The second block saw the addition of the number of errors players made during each level. The final block saw the addition of the number of moves players made over and beyond the optimal solution for each level.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Table 5. Hierarchical Regression Results Predicting Self-report Difficulty

Predictors	Block 1			Block 2			Block3		
	<i>B</i>	<i>SE</i>	$\beta$	<i>B</i>	<i>SE</i>	$\beta$	<i>B</i>	<i>SE</i>	$\beta$
Time	.02*	.00	.92	.02*	.00	1.07	.01*	.01	.59
Errors				-.08*	.08	-.17	-.03*	.08	-.06
Over-moves							.01*	.01	.40
R <sup>2</sup>	.83			.84			.85		
$\Delta R^2$				.00			.01		

Note. , n = 49, \* $p < .001$

As can be seen in Table 5, the first predictor block accounted for 84% of the variance in self report difficulty ratings. Time was positively and significantly related to self-report difficulty. Errors was added in block two and the amount of variance significantly ( $F(2,49) = 125.67, p < .001$ ) increased to 84%. Over-moves was added to block three and the amount of variance accounted for significantly ( $F(3,49) = 87.28, p < .001$ ) increased to 85%.

Turning to the regression estimates, time was the highest significant and positive predictor of self-report difficulty ratings, indicating that the more time a player spent playing a level indicated a higher perception of level difficulty. Both errors and over-moves were not uniquely predictive of self-report difficulty. While both errors and over-

## REAL-TIME PREDICTORS OF ENGAGEMENT

moves were found to be significant in the regression, neither added practical predictive value.

### Study One Discussion

Based on the small range of reported scores hovering around the neutral mark, the Fun self-report measure was dropped. This neutrality in rating the game as fun was likely due to multiple influences such as comparisons with other video games and personal preferences. This lack of predictive validity was also seen in the lack of correlation to any of the performance metrics.

The Frustrating and Challenging self-report measures were determined to be redundant and also dropped from analysis because they were highly correlated with the Difficulty measure and didn't add additional insight into perceived overall level difficulty.

The multiple regression analysis showed that time is the best predictor of self-report Difficulty. Based on instructions provided to participants, the goal of the game was to complete the puzzle in as quick of a time as possible, and therefore the data were consistent with expectations that the longer it takes a person to solve a level, the more difficult they would likely find it.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### Study Two

The goal for Study Two was to answer the first and second research questions: can engagement be measured in terms of cognitive load and affect, and can a person's engagement (desire to experience an optimum level of cognitive load) be accurately measured while playing a video game? Specifically, Study Two expanded the assessment of the criterion validity of the Game-clock as a potential embedded affective measure of engagement. In addition, a secondary task was introduced as a potential embedded measure of cognitive load. The predictive relationship of these real-time measures was modeled together and against game performance and level difficulty derived from Study One.

This study is exploratory in nature. The primary goal was to examine the relationships between the real-time measures of cognitive load and affect with calculated difficulty of a level in order to measure the criterion validity of the real-time measures. Since no other real-time embedded measure of engagement can practically be captured in parallel, additional data was used to triangulate with the real-time measures of cognitive load and affect. The analytic strategy to answer Research Questions 1 and 2 proceeds as follows:

## REAL-TIME PREDICTORS OF ENGAGEMENT

1. Game performance metrics were compared against the results of Study One. The expectation was that the secondary task will not add a significant decrement to gameplay performance.

2. Gameplay performance metrics were used in modeling the real-time engagement variables. That is, based on the derived Flow model in Figure 5 and the results of previous research (Sharek, 2010), it was expected that there would be a threshold of cognitive load (as measured by the secondary task) induced by game difficulty (measured by the gameplay performance metrics) after which desire to exit (as measured by Game-clock checks) would rise.

3. The summative cognitive and affective self-report measures at the end of the study were used to confirm that the overall experience of gameplay was within reasonable bounds of cognitive load and positive affect.

### *Independent Variable*

One game condition was designed using the levels tested in Study One.

**Design Condition 1 (Linear)** – This condition followed the design patterns of a video game that adheres to a typical linear increase in difficulty as players progress. The design was similar to the Flow condition used in previous research (Sharek, 2010). The arrangement of levels was predetermined using the level difficulty data from Study One.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Real-time Engagement Dependent Variables*

*Game-clock Checks.* The number of times a player checked the Game-clock during intermissions (points of disengagement) was recorded. Players were told they must play the game for 15 minutes and that they can check whether or not the minimum amount of time had been reached by pressing the spacebar.

*Secondary Task.* The secondary task was manifested as a measure of reaction time (Brunken, et al., 2003; Ogden, et al., 1979) to capture the relative cognitive load imposed by the primary task. Players were told to monitor a small radar screen at the bottom of the game. When asteroids enter the danger zone, players would need to press the **F** button to fire a missile and protect the loading deck.

This secondary task was not designed to suppress the player's ability to attend to the primary task of solving the puzzles. However, this secondary task was designed to be a simple, continuous monitoring task that requires the same resources in visual working memory.

### *Gameplay Performance Dependent Variables*

*Moves.* The number of times the player pressed the keyboard arrow keys to move the block for each level around the game board.

## REAL-TIME PREDICTORS OF ENGAGEMENT

*Over-moves.* The number of moves the player made over and beyond the minimum number of moves required to complete the level was recorded.

*Directions.* The number of times the player changed the block's direction was recorded. For example, if the player moved the block to the left five spaces and then up for two spaces, the number of directions would be recorded as two.

*Errors.* The number of times the player moved the block off of the game board.

*Total Time.* The amount of time the player spent playing the entire game.

*Overtime.* The amount of time the player spent playing the game after the minimum amount of time was met.

### *Summative Self-Report Dependent Variables*

*Overall Engagement.* The Game Engagement Questionnaire (GEQ) (Brockmyer, et al., 2009) was used as a summative rating of engagement. This 19-item measurement tool (Appendix D) was developed as a post-hoc self-report measure of a game player's level of engagement along the dimensions of absorption, Flow, presence, and immersion. This measure provided a post-hoc summative measure of the game-player's engagement experience.

## REAL-TIME PREDICTORS OF ENGAGEMENT

*Cognitive Load.* The NASA-TLX (Hart & Staveland, 1988) was used as a summative post-hoc self-report measure of cognitive load.

*Affect.* The Interest/Enjoyment items from the Intrinsic Motivation Inventory (Deci, 2009) (Appendix E) was used as a summative post-hoc self-report measure of affect.

### Study Two Method

#### *Participants*

A total of 101 people were recruited through Amazon's Mechanical Turk (Amazon.com, 2009). Data from fourteen participants were removed due to a lack of interaction with the radar screen secondary task; they did not press the **F** key at all during the game. To reduce the influence of practice effects, data from the first level everyone played was dropped. Of the remaining 87 participants, 47% were female and 53% were male (mean age = 32.99, *SD* = 10.43). Ninety-seven percent were from the United States. Thirty percent indicated that they had previous experience with a similar type of puzzle game.

#### *Materials and Apparatus*

The experiment was hosted at [www.gridblocker.com/space/game2](http://www.gridblocker.com/space/game2), a Website maintained by the author developed specifically for hosting the GridBlocker game.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Similar to Study One, this second version of GridBlocker was powered by a MySQL database and a PHP server-side scripting language. The experiment was accessible from any computer with Internet access. The computer requirements included a minimum computer monitor resolution of 1000x600 and an installed version of the Adobe Flash player v.10 or greater.

### *Procedure*

Once participants agreed to participate, they were asked to enter in a username of their choosing. This username was used to generate a secure 32 digit hexadecimal unique user ID (UID) using an MD5 cryptographic hash function. The purpose of this UID is to link individual participant performance and ratings throughout the experiment in the database tables.

Before beginning, participants were given instructions explaining how to play the game. The story described in Study One was used in order to explain the premise of the game, however the story was expanded to introduce the secondary task, thus:

Additionally, the asteroid showers have not completely left your sector and stray asteroids may damage the loading deck. Use the radar screen at the bottom left to monitor any stray asteroids. If an asteroid enters the red **danger zone**, press the **F** key to fire a missile and destroy the asteroid. Good luck!"

## REAL-TIME PREDICTORS OF ENGAGEMENT

After reading the instructions and story, participants began playing GridBlocker. They were required to play the game for a minimum time of 15 minutes. After playing the game, participants were asked to answer a short survey including demographic questions, the intrinsic/enjoyment questionnaire, NASA-TLX, and the GEQ. Upon completion of the survey, participants were debriefed and thanked. An experimental completion code was provided so participants could submit it back into the Mechanical Turk's HIT to receive payment.

### Study Two Hypotheses

#### *Confirmation of Game Performance Metric*

**H1:** Game performance metric data will follow the same trends as found in Study One.

#### *Diagnosticity of the Real-time Measures*

**H2:** Secondary task performance will correlate significantly and negatively with the individual level difficulty based on the game performance metric.

**H3:** Intermission Game-clock checks will negatively correlate with gameplay overtime.

## REAL-TIME PREDICTORS OF ENGAGEMENT

**H4:** Intermission Game-clock checks will begin to rise at a consistent inflection point when game level difficulty reaches a derived point found in exploratory analysis.

### *Summative Check of Gameplay Experience*

**H5:** Secondary task performance, across levels, will negatively correlate with summative self-report levels of cognitive load (NASA-TLX).

**H6:** Intermission Game-clock checks, across levels, will negatively correlate with summative self-report levels of positive affect using the interest/enjoyment items from the Intrinsic Motivation Inventory.

**H7:** Game-clock checks, both during intermissions and during gameplay will negatively correlate with summative self-report post-hoc engagement using the GEQ.

## Study Two Results

Player performance, self-reported engagement, self-reported subjective workload measures for the NASA-TLX, and self-reported affect are presented in this section. An alpha level of .05 was used for all analyses. Descriptive statistics were carried out on all dependent variables to check for data abnormalities such as outliers or floor/ceiling effects. Pearson product-moment correlations were used for hypotheses two through eight.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Because the game in Study Two presented levels in a linear increase in difficulty, only 45 individual levels were played. These levels were compared with their counterparts from Study One. As can be seen in Table 6, Hypothesis One was mostly supported. Study Two's performance data followed the same trends as found in Study One with the exception of no correlation being found between difficulty and both errors and time.

Table 6. *Bivariate Correlation Table*

	<i>M</i>	<i>SD</i>	Directions	Time	Errors	Over-moves
Study One						
Directions	52.97	34.85				
Time	124.37	85.71	.96**			
Errors	4.23	3.11	.84**	.91**		
Over-moves	85.54	62.82	.95**	.96**	.83**	
Difficulty	46.17	21.34	.91**	.94**	.88**	.90**
Study Two						
Directions	34.09	27.58				
Time	124.44	85.80	.59**			
Errors	1.43	1.03	.37*	-.08		
Over-moves	45.19	39.97	.94**	.54**	.38*	
Difficulty	45.87	21.85	.51**	.94**	-.07	.44**

Note:  $n = 45$ , \* $p < .05$ , \*\* $p < .001$

Hypothesis Two was found to be true. Secondary task performance ( $M=2750.60$ ,  $SD=1286.28$ ) was correlated significantly and negatively with the individual level

## REAL-TIME PREDICTORS OF ENGAGEMENT

difficulty ( $M=45.89$ ,  $SD=21.85$ ) based on the game performance metric  $r(43) = -.54$ ,  $p < .001$ .

Hypothesis Three was also found to be true. Intermission Game-clock checks ( $M=4.44$ ,  $SD=3.52$ ) were negatively correlated with gameplay overtime ( $M=93.06$ ,  $SD=324.41$ ),  $r(85) = -.23$ ,  $p = .03$ .

As can be seen in Figure 9, Hypothesis Four was supported. A hierarchical regression with the linear variance partialled out revealed a significant quadratic relationship between difficulty and the mean intermission Game-clock checks, ( $F(2,43) = 6.54$ ,  $p = .007$ , part correlation =  $-.62$ ); 40% of the variance was explained. Intermission Game-clock checks increased as players experienced more difficult levels until a difficulty level of around 27 was reached. At this point a decrease in Game-clock checks was found.

## REAL-TIME PREDICTORS OF ENGAGEMENT

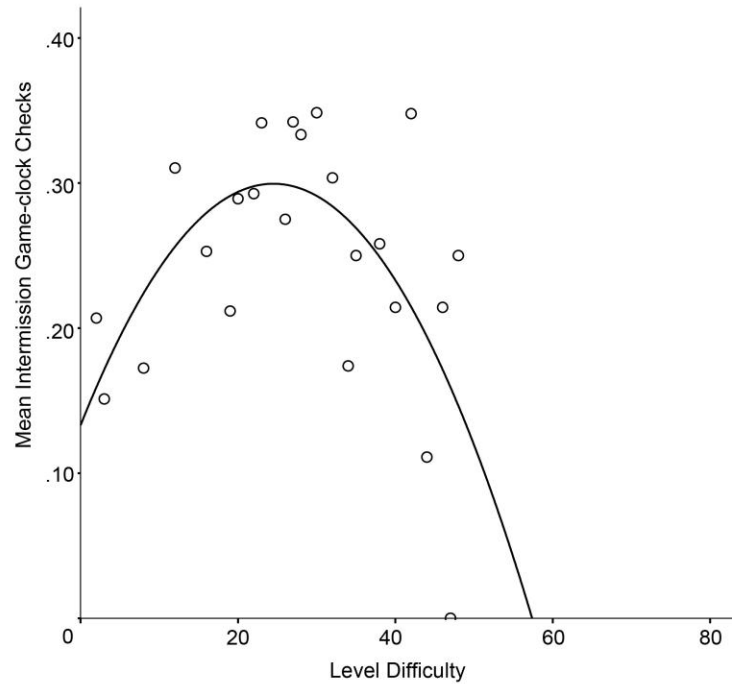


Figure 9. Mean Game-clock Checks as a Function of Difficulty

Hypothesis Five was not supported. Secondary task performance ( $M=2750.60$ ,  $SD=1286.28$ ) was negatively, but not significantly correlated with summative self-report levels of cognitive load using the NASA-TLX ( $M=60.65$ ,  $SD=14.72$ ).

Hypothesis Six was supported. Intermission Game-clock checks collapsed across levels ( $M=.13$ ,  $SD=.14$ ) were negatively correlated with summative self-report levels of positive affect ( $M=5.33$ ,  $SD=.57$ ) from the Intrinsic Motivation Inventory  $r(43) = -.78$ ,  $p < .001$ .

## REAL-TIME PREDICTORS OF ENGAGEMENT

Hypothesis Seven also was supported. Intermission Game-clock checks ( $M=.13$ ,  $SD=.14$ ) were negatively and significantly correlated with summative self-report post-hoc engagement using the GEQ ( $M=2.65$ ,  $SD=.09$ ),  $r(85) = -.25$ ,  $p<.05$ .

### Study Two Discussion

Data from the manipulation check supported Hypothesis One by providing evidence that the video game difficulty levels were designed as expected. Additionally, the introduction of the secondary task did not seem to add a significant decrement to gameplay performance.

The higher the difficulty rating a level had, the more moves on average a player made over and beyond the solution indicating that finding the optimal solution to levels became less clear as difficulty increased. As would be assumed, the more difficult a level was, the more directional changes a player made with the block. Additionally, as difficulty increased, so did the amount of time it took a player to solve the level. The only discrepancy with Study One was that difficulty and number of errors were not significantly correlated in Study Two. The number of errors made in Study Two were much lower than the number of errors made in Study One. This was possibly due to a scaffolding effect where the player's in Study Two were able to hone their skills by playing increasingly difficult levels, while those in Study One were given levels at

## REAL-TIME PREDICTORS OF ENGAGEMENT

random; some players were served very difficult levels from the moment they began playing.

Note that the difficulty ratings from Study One were derived from the individual subjective ratings using a 10-point rating scale (See Study One for more details). The difficulty ratings used in Study Two were derived by rank-ordering each of the 108 levels based on the subjective difficulty ratings from Study One.

Hypothesis Two was also found to be true, secondary task reaction time was negatively correlated with difficulty level. In other words, the more difficult a level was, the more time it took for a player to perceive and destroy an asteroid on the radar screen. This indicates that the primary task (moving the block towards the goal) in more difficult levels demanded greater amounts of cognitive effort compared to less difficult levels. This is a promising result because it shows that the secondary task can potentially be used to determine whether a level is cognitively under- or over-loading a player.

Hypothesis Three was found to be true; intermission Game-clock checks were negatively correlated with gameplay overtime. Players who stayed in the game after they were notified that they could exit clicked on the Game-clock significantly less than those who exited the game as soon as possible. This is an important validation that provides

## REAL-TIME PREDICTORS OF ENGAGEMENT

more evidence that intermission checks are measuring a player's desire to continue or leave the game.

Hypothesis Four was found to be true. Intermission Game-clock checks increased as players experienced more difficult levels until a difficulty level of around 40 was reached. It is likely that players who had reached this high level of difficulty were not interested in exiting the game therefore checking the Game-clock was not a priority. Around level 48, players on average had reached the minimum amount of required time and were notified that they could exit. This meant that checking the Game-clock was no longer required.

Hypothesis Five was not found to be true. Secondary task performance was negatively, but not significantly correlated with summative self-report levels of cognitive load using the NASA-TLX. The summative nature of the NASA-TLX may not provide the sensitivity required to determine exactly what aspects of the game influenced the summative rating scores. For example, it could be that a single, frustratingly difficult level resulted in a high self-report cognitive load score. The lack of a correlation between a player's NASA-TLX scores and their mean response times from the secondary task provides some evidence of a gulf between post-hoc and real-time data and the potential advantages of secondary task tools in game-based environments to measure real-time cognitive load. This is further explored in Study Three.

## REAL-TIME PREDICTORS OF ENGAGEMENT

On average, players rated their affect higher than five on a seven-point scale. Hypothesis Six was found to be true, intermission Game-clock checks, across levels, were negatively correlated with summative self-report levels of positive affect from the Intrinsic Motivation Inventory. The lower players rated their overall affect, the more they clicked on the Game-clock during intermissions.

Hypothesis Seven also was supported. Intermission Game-clock checks were negatively and significantly correlated with summative self-report post-hoc engagement using the GEQ. Though the correlation was significant, it was quite low. This could be due to the lack of variance in the GEQ scores. The average GEQ score was about average at 2.65 on a five-point scale with a standard deviation of only .09. The secondary task performance scores were more widely distributed and may provide more insight into an individual player's engagement compared to the GEQ. The fact that the GEQ provides a summative score could be the cause of this lack of strong correlation.

The primary goal of this exploratory study was to examine whether data from the summative questionnaires would be correlated with the real-time measures. A lack of strong correlation was found between the summative measures and the real-time measures. However, the real-time measures were correlated with the performance measures. This opens the question of whether or not the summative measures are sensitive enough to be able to accurately summarize the individual state a player

## REAL-TIME PREDICTORS OF ENGAGEMENT

experiences while playing the game. On average, the summative measures showed that players had a good experience with the game, yet the Game-clock and secondary task performance measures showed that players experienced difficulty and higher cognitive load at various points throughout the game. These real-time data may be used to adapt the game levels to a player's current state of engagement. The next study will be used to investigate the viability of the real-time measures to increase player performance without negatively affecting the summative outcomes.

### Study Three

Study Two revealed that the real-time measures will provide a usable method for assessing engagement in terms of cognitive load and affect. For Study Three, an Adaptive game condition was developed and tested against two other conditions: Linear and Choice. Table 7 provides a summary of the three conditions used in this study. The Linear condition was the same design used in Study Two. The Choice condition allows the player to choose whether an upcoming level will be easier or harder than the just-completed level, while the Adaptive condition uses an algorithm derived from the real-time measures examined in Study Two to determine the difficulty of the upcoming level. The goal of research question three is to find out if this real-time measure of engagement based on the combination of a willingness to participate and a desire to experience cognitive load can be used to develop a video game that adapts to a player's real-time

## REAL-TIME PREDICTORS OF ENGAGEMENT

level of engagement. Based on research that alludes to the potential compelling experiences that dynamic game difficulty balancing can provide (Hunicke, 2005), research question four seeks to uncover whether or not there are performance differences between adaptive game design and the more traditional approaches of either linear or choice gaming environments. This study was conducted to answer research questions three and four.

Table 7. *Summary of the Three Conditions Used in Study Three*

Condition	Static (Y/N)	Locus of Control (System/User)
1 Linear	Y	System
2 Choice	N	User
3 Adaptive	N	System

### Study Three Design

#### *Independent Variables*

Three design conditions were created using the results from Studies One and Two. As in Study Two, the individual game level ratings collected on Study One were used to create a rank order progression of difficulty for the game levels. This difficulty rating for each level was used in the creation of each of the three conditions.

## REAL-TIME PREDICTORS OF ENGAGEMENT

**Design Condition 1 (Linear)** - This condition followed the design patterns of a video game that adheres to a typical linear increase in difficulty as players progress. The design was similar to the Flow condition used in the Thesis study. The arrangement of levels was predetermined using the level difficulty data from Study One. This was also the design used in Study Two. Each new level increased by a difficulty of three. For example, players would begin at a difficulty of one, and then the next level's difficulty rating would be four, followed by seven and so on.

**Design Condition 2 (Choice)** – After each level, players were given the option to play an easier level (minus one in difficulty), a moderately more difficult level (plus three in difficulty just as in the Linear condition) or a much more difficult level (plus five in difficulty).

**Design Condition 3 (Adaptive)** - This condition was designed based on the results of Study Two. An algorithm was developed to predict a player's real-time level of engagement. Game levels were selected according to past game level performance, Game-clock checking behavior, and secondary task performance. When it was determined that the player was becoming bored (effortlessness), more difficult levels were provided. The options for these levels mirrored the Choice condition where either an increase in three for a moderately more difficult level, or an increase in five for a much more difficult level was determined using the adaptive algorithm. When it was

## REAL-TIME PREDICTORS OF ENGAGEMENT

determined that players were becoming frustrated, easier levels were served (minus one in difficulty). As with the Choice and Linear conditions, the options are biased towards more difficult levels based on the assumption that time-on-task is developing their skill level with the game regardless of performance outcome.

As can be seen in Figure 10, the adaptive algorithm incorporates two primary data inputs. The first input takes data from all previous game levels played and provides an overall impression of the player's engagement and performance. The second input takes data from the most recent level the player completed. This provides an immediate impression of the player's engagement and performance. Both data inputs consist of: the number of times a player checked the Game-clock during intermission, mean secondary-task response time, and the amount of time the player spent solving each level (i.e., player performance). This approach creates an algorithm that draws from a broad, but relevant set of data to inform the choice of next level. It takes into account both past and current performance along with both cognitive and affective real-time state measures.

# REAL-TIME PREDICTORS OF ENGAGEMENT

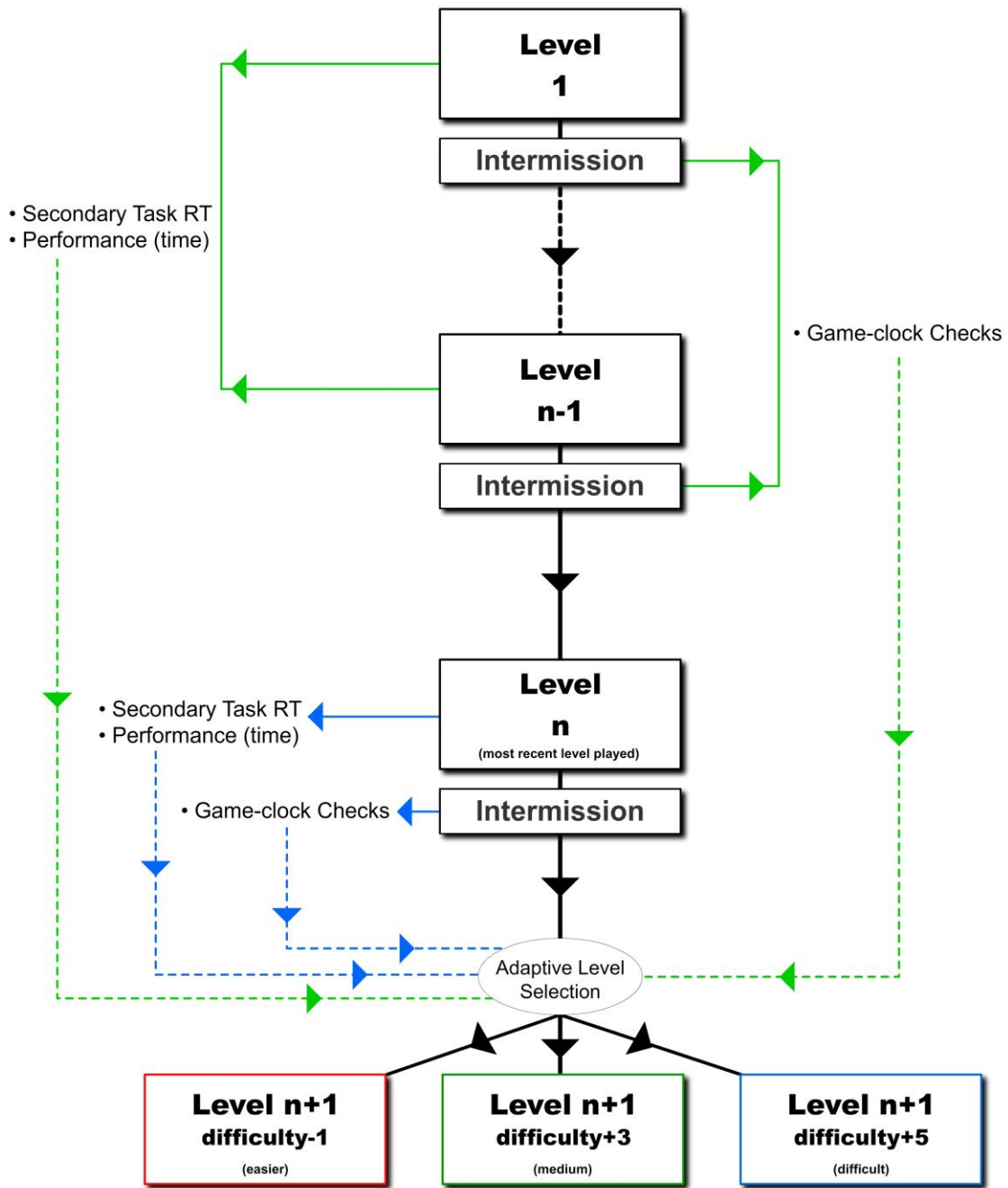


Figure 10. Adaptive Algorithm Model

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Real-time Engagement Dependent Variables*

*Game-clock Checks.* The number of times a player checked the Game-clock during intermissions (points of disengagement). Players were told they must play the game for 15 minutes and that they can check whether or not the minimum amount of time had been reached by pressing the spacebar.

*Secondary Task.* The secondary task was manifested as a measure of reaction time (Brunken, et al., 2003; Ogden, et al., 1979) to capture the relative cognitive load imposed by the primary task. Players were told to monitor a small radar screen at the bottom of the game. When asteroids enter the danger zone, players would need to press the **F** button to fire a missile and protect the loading deck.

This secondary task was not designed to suppress the player's ability to attend to the primary task of solving the puzzles. However, this secondary task was designed to be a simple, continuous monitoring task that requires the same resources in visual working memory as the primary task.

### *Gameplay Performance Dependent Variables*

*Total Time.* The amount of time the player spent playing the entire game.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Summative Self-Report Dependent Variables*

*Overall Engagement.* The User Engagement Scale (UES) (O'Brien & Toms, 2010) was used as a summative measure post-hoc ratings of engagement. This 31-item measurement tool (Appendix G) was developed as a *post hoc* self-report measure of a person's level of engagement along the dimensions of focused attention, perceived usability, aesthetics, endurance, novelty, and involvement. This measure will provide a post-hoc summative measure of the game-player's experience.

*Cognitive Load.* The NASA-TLX (Hart & Staveland, 1988) was used as a summative post-hoc self-report measure of cognitive load.

*Affect.* The Interest/Enjoyment items from the Intrinsic Motivation Inventory (Deci, 2009) (Appendix E) was used as a summative post-hoc self-report measure of affect. This measure was compared with Game-clock checks.

### Study Three Method

#### *Participants*

A total of 340 people were recruited through Amazon's Mechanical Turk (Amazon.com, 2009). Data from seventeen participants were removed due to a lack of interaction with the radar screen secondary task; they did not press the **F** key at all during the game. To reduce the influence of practice effects, data from the first three levels

## REAL-TIME PREDICTORS OF ENGAGEMENT

everyone played was dropped. Of the remaining 323 participants, 60% were female and 40% were male (mean age = 30.60,  $SD = 9.94$ ). Ninety-seven percent were from the United States. Twenty-nine percent indicated that they had previous experience with a similar type of puzzle game. Participants reported playing video games for an average of seven hours ( $SD = 9.70$ ) per week.

### *Materials and Apparatus*

The experiment was hosted at [www.gridblocker.com/space/game3](http://www.gridblocker.com/space/game3). All materials and apparatuses were the same as those described in Study Two.

### *Procedure*

Once participants agreed to participate, they were asked to enter in a username of their choosing. This username was used to generate a secure 32 digit hexadecimal unique user ID (UID) using an MD5 cryptographic hash function. The purpose of this UID was to link individual participant performance and ratings throughout the experiment in the database tables.

Once participants logged in to the game, they were randomly assigned to one of the three conditions: Linear, Choice, or Adaptive. Participants then began playing GridBlocker. After playing the game, participants were asked to answer a short survey including a few demographic questions, the NASA-TLX, and the UES. Upon completion

## REAL-TIME PREDICTORS OF ENGAGEMENT

of the survey, participants were debriefed and thanked. An experimental completion code was provided so participants could submit it back into the Mechanical Turk's HIT in order to receive payment.

### *Hypothesis – Summative Measures*

**H1:** Engagement - Those in the Adaptive and Choice conditions will report higher levels of engagement (UES summative measure) compared to those in the Linear condition.

$$H1: UES_{Adaptive} \geq UES_{Choice} > UES_{Linear}$$

**H2:** Affect - Those in the Adaptive and Choice conditions will report higher levels of personal affect (AFF) via the Intrinsic motivation/Enjoyment Inventory (Deci, 2009) compared to those in the Linear condition.

$$H2: AFF_{Adaptive} \geq AFF_{Choice} > AFF_{Linear}$$

**H3:** Cognitive Load - Those in the Adaptive and Choice conditions will report lower levels of cognitive load (CL) via the NASA-TLX compared to those in the Linear condition.

$$H3: CL_{Adaptive} \leq CL_{Choice} < CL_{Linear}$$

**H4:** Total Time - Those in the Adaptive and Choice conditions will play longer into overtime (OT) compared to those in the Linear condition.

## REAL-TIME PREDICTORS OF ENGAGEMENT

$$H4: OT_{Adaptive} \geq OT_{Choice} > OT_{Linear}$$

### *Hypothesis – Performance*

**H5:** Difficulty –Those in the Adaptive condition will require fewer levels to reach greater difficulty (Diff) followed by those in the Choice conditions and then those in the Linear condition. This metric was calculated as a ratio of the maximum difficulty achieved divided by number of levels played.

$$H5: Diff_{Adaptive} > Diff_{Choice} > Diff_{Linear}$$

**H6:** Level Time - Those in the Adaptive and Choice conditions will require the least amount of time to complete each level (LT) compared to those in the Linear condition.

$$H6: LT_{Adaptive} \leq LT_{Choice} < LT_{Linear}$$

**H7:** Reaction Time - Those in the Adaptive and Choice conditions will produce quicker reaction times (RT) in the secondary task compared to those in the Linear condition.

$$H7: RT_{Adaptive} \leq RT_{Choice} < RT_{Linear}$$

## REAL-TIME PREDICTORS OF ENGAGEMENT

### Results

#### *Summative Measures*

The first hypothesis, that those in the Adaptive and Choice conditions will report higher levels of overall engagement (UES) compared to those in the Linear condition, was investigated by analyzing the results from the UES using a one-way analysis of variance (ANOVA). Descriptive statistics can be found in Table 8. No significant differences were found between conditions,  $F(2,320) = 2.00, p = .14, \eta_p^2 = .01$ .

Table 8. *UES Descriptive Statistics*

Condition	<i>N</i>	<i>M</i>	<i>SD</i>
Linear	102	3.19	.67
Choice	113	3.30	.69
Adaptive	108	3.12	.66

The second hypothesis, that those in the Adaptive and Choice conditions will report higher levels of personal affect compared to those in the Linear condition, was investigated by analyzing the results from the Interest/Enjoyment items from the Intrinsic Motivation Inventory using a one-way analysis of variance (ANOVA). Descriptive

## REAL-TIME PREDICTORS OF ENGAGEMENT

statistics can be found in Table 9. No significant differences were found between conditions,  $F(2,320) = 1.84, p = .16, \eta_p^2 = .01$ .

Table 9. *Affect Descriptive Statistics*

Condition	<i>N</i>	<i>M</i>	<i>SD</i>
Linear	102	4.52	1.53
Choice	113	4.70	1.39
Adaptive	108	4.31	1.51

The third hypothesis, that those in the Adaptive and Choice conditions will report lower levels of cognitive load compared to those in the Linear condition, was investigated by analyzing the results from the NASA-TLX using a one-way analysis of variance (ANOVA). Descriptive statistics can be found in Table 10. No significant differences were found between conditions,  $F(2,320) = .34, p = .71, \eta_p^2 < .01$ .

## REAL-TIME PREDICTORS OF ENGAGEMENT

Table 10. *Summative Cognitive-Load Descriptive Statistics*

Condition	<i>N</i>	<i>M</i>	<i>SD</i>
Linear	102	60.45	14.15
Choice	113	59.07	14.95
Adaptive	108	60.46	14.01

To further explore the lack of significance from the three self-report measures, the difficulty level selection generated by the adaptive algorithm was used to analyze level selection for each condition. After each level, the adaptive algorithm calculated whether a player should be served an easier, a more difficult, or a very difficult level (see Table 11). Though the Linear and Choice conditions did not use this score, the calculation was still conducted for analytical purposes. The mean number of times an easier level was selected was analyzed using a one-way ANOVA and significant differences were found between the conditions,  $F(2,320) = 9.78, p < .001, \eta_p^2 = .06$ . A follow-up Bonferroni post-hoc test indicated significant differences between the Choice condition and both the Linear and Adaptive conditions. That is to say that the adaptive algorithm determined that participants in the Choice condition should have played an easier level the least amount of times (i.e., less than they chose to do so) compared to those in the other two conditions.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Similarly, the mean number of times the adaptive algorithm selected a very difficult level was analyzed using a one-way ANOVA and significant differences were found between the conditions,  $F(2,320) = 26.59, p < .001, \eta_p^2 = .14$ . A follow-up Bonferroni post-hoc test indicated significant differences across all three conditions. That is to say that the adaptive algorithm determined that those in the Choice condition should have played very difficult levels more times compared to those in the Adaptive condition, with those in the Linear condition required very difficult levels the least amount of times. No significant differences were found in the adaptive algorithm's medium difficulty selection.

Table 11. *Adaptive Algorithm's Difficulty Level Selection Descriptive Statistics*

	<i>N</i>	<i>M</i>	<i>SD</i>
Easier Level			
Linear Condition	102	1.08	1.57
Choice Condition	113	.37	.80
Adaptive Condition	108	.88	1.19
Medium-Difficult Level			
Linear Condition	102	5.04	3.29
Choice Condition	113	4.46	3.03
Adaptive Condition	108	5.02	3.55
Very-Difficult Level			
Linear Condition	102	5.03	4.66
Choice Condition	113	10.18	6.53
Adaptive Condition	108	7.10	4.05

## REAL-TIME PREDICTORS OF ENGAGEMENT

The fourth hypothesis, that those in the Adaptive and Choice conditions will play longer into overtime compared to those in the Linear condition, was investigated by analyzing overtime performance results (measured in seconds) using a one-way analysis of variance (ANOVA). Descriptive statistics can be found in Table 12. Despite those in the Adaptive condition producing a mean three-times larger than the other two conditions, the accompanying large standard deviation resulted in producing no significant differences between the conditions,  $F(2,320) = .56, p = .57, \eta_p^2 < .01$ .

Table 12. *Overtime Descriptive Statistics*

Condition	<i>N</i>	<i>M</i>	<i>SD</i>
Linear	102	30.79	86.36
Choice	113	31.53	180.03
Adaptive	108	97.39	898.17

### *Performance*

The fifth hypothesis, that those in the Adaptive and Choice conditions will require fewer levels to reach greater difficulty compared to those in the Linear condition was

## REAL-TIME PREDICTORS OF ENGAGEMENT

investigated using a one-way analysis of variance (ANOVA). Descriptive statistics can be found in Table 13.

Table 13. *Difficulty, Levels, and Difficulty/level Ratio Descriptive Statistics*

Condition	<i>N</i>	<i>M</i>	<i>SD</i>
Levels Played			
Linear	102	14.17	4.02
Choice	113	15.96	5.52
Adaptive	108	13.17	2.95
Difficulty Reached			
Linear	102	40.50	12.05
Choice	113	33.18	13.69
Adaptive	108	44.61	13.93
Difficulty/Level Ratio			
Linear	102	2.85	.05
Choice	113	2.24	.91
Adaptive	108	3.40	.69

## REAL-TIME PREDICTORS OF ENGAGEMENT

To answer the fifth hypothesis a difficulty/level ratio metric was calculated as a ratio of the maximum difficulty achieved to the number of levels played. A higher difficulty/level ratio indicates higher performance. A significant difference was found between the conditions,  $F(2,320) = 81.88, p < .001, \eta_p^2 = .34$ , and a follow-up Bonferroni post-hoc test indicated significant differences between all three conditions. As can be seen in Figure 11, those in the Adaptive condition produced a higher difficulty/level ratio compared to those in the Linear condition with those in the Choice condition producing the lowest difficulty/level ratio value. In other words, those in the Adaptive condition played fewer levels yet achieved greater difficulty compared to the other two conditions. The low Choice condition ratio resulted from the group playing the most number of levels, yet reaching the lowest level of overall difficulty.

## REAL-TIME PREDICTORS OF ENGAGEMENT

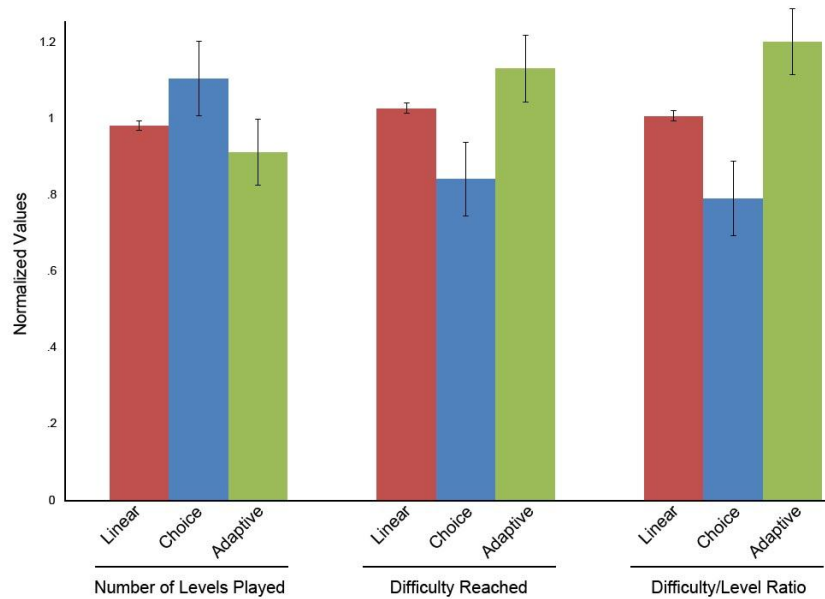


Figure 11. Difficulty, Levels, and Difficulty/level Ratio

To further understand the relationship between the order of levels played and the difficulty participants reached in each condition, individual player paths were visualized in Figures 12 – 14. Negatively sloping lines indicate players picked easier levels. As Figure 12 demonstrates, those in the Linear condition were only given levels that increased in difficulty. In contrast, Figure 13 shows that those in the Choice condition produced a wide range of difficulty selection behaviors. Figure 14 shows that, apart from one player, those in the Adaptive condition displayed performance characteristics that allowed the adaptive algorithm to pick increasingly difficult levels for them. Figure 14

## REAL-TIME PREDICTORS OF ENGAGEMENT

indicates that the Adaptive group appears more like the Linear condition, but with more variation in the exact trajectory that a player took.

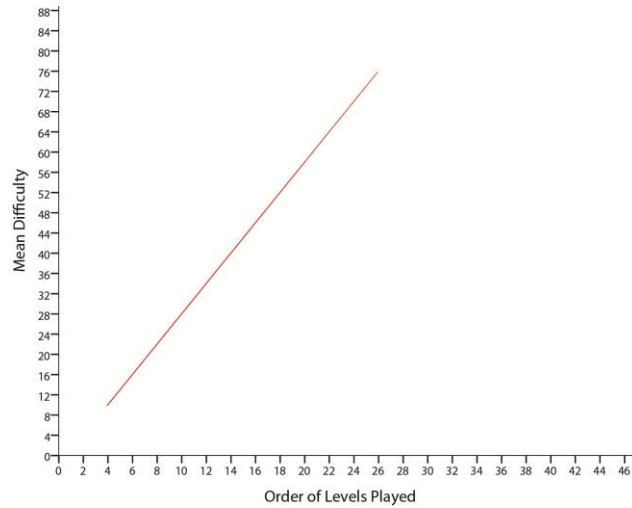


Figure 12. Overlay of Individual Player Paths - Linear Condition

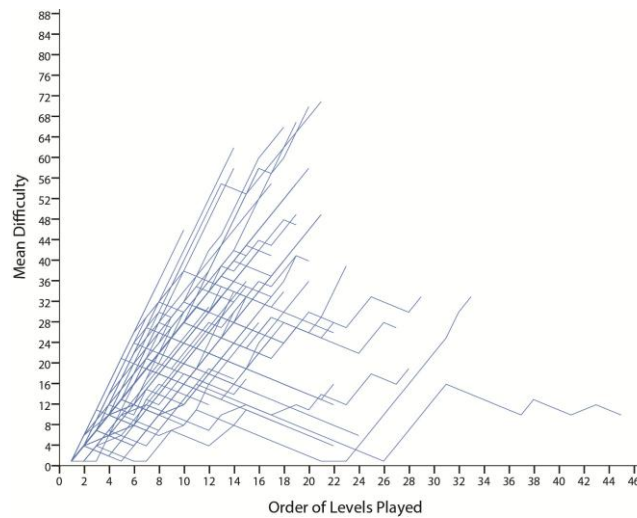
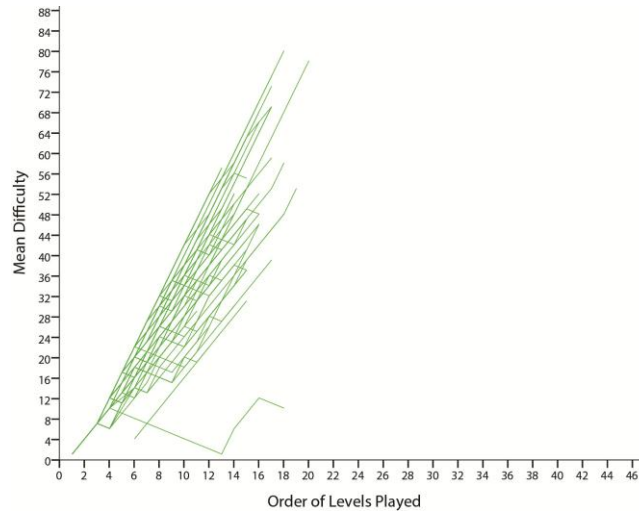


Figure 13. Overlay of Individual Player Paths - Choice Condition

## REAL-TIME PREDICTORS OF ENGAGEMENT



*Figure 14.* Overlay of Individual Player Paths - Adaptive Condition

Figure 15 provides a composite visualization of the mean order of levels played compared to the mean maximum difficulty reached for all three conditions. It can be seen that after level three, the slope became steeper for those in the Adaptive condition indicating that the adaptive algorithm determined many players' performance to be high enough to successfully play increasingly difficult levels. Additionally, those in the Choice condition who played up to 18 levels, on average, selected increasingly difficult levels, but those who played more than 18 levels never reached a high level of difficulty. After approximately 18 levels, players in the Choice condition increasingly chose easier levels to play.

## REAL-TIME PREDICTORS OF ENGAGEMENT

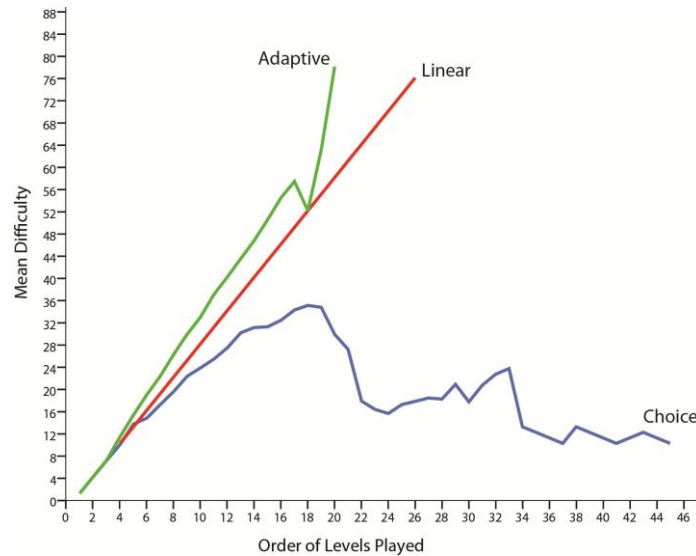


Figure 15. Composite Overlay of Mean Difficulty by Level Order

The sixth hypothesis, that those in the Adaptive and Choice conditions will require the least amount of time to complete each level compared to those in the Linear condition was investigated using a one-way analysis of variance (ANOVA). A significant difference was found between the conditions,  $F(2,320) = 17.35, p < .001, \eta_p^2 = .10$ , and a follow-up Bonferroni post-hoc test indicated that those in the Linear condition took a significantly longer time to complete each level ( $M = 56.61$ ) compared to those in the Choice ( $M = 40.83$ ) and Adaptive ( $M = 45.87$ ) conditions (Figure 16). No significant differences were found between the Choice and Adaptive conditions.

## REAL-TIME PREDICTORS OF ENGAGEMENT

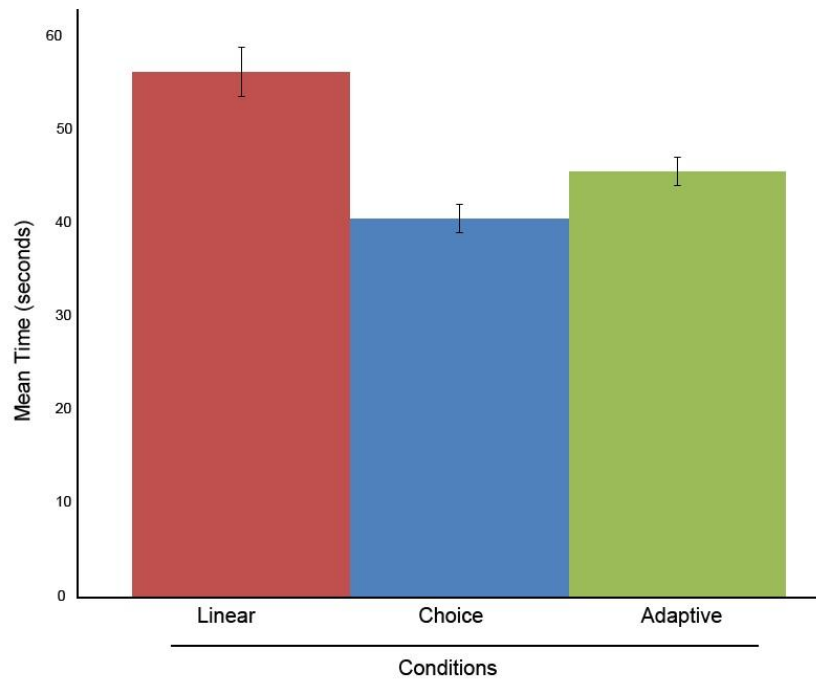


Figure 16. Mean Time to Complete Individual Levels

The seventh hypothesis, that those in the Adaptive and Choice conditions will produce quicker reaction times in the secondary task compared to those in the Linear condition was investigated using a one-way analysis of variance (ANOVA). A significant difference was found between the conditions,  $F(2,320) = 10.52, p < .001, \eta_p^2 = .06$ , and a follow-up Bonferroni post-hoc test indicated that those in the Linear condition produced significantly higher reaction times ( $M = 3116.73$ ) compared to those in the Choice ( $M = 2585.13$ ) and Adaptive ( $M = 2565.09$ ) conditions. No significant differences were found between the Choice and Adaptive conditions.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### Study Three Discussion

The Adaptive condition represents the operationalization of a psychological theory of engagement to a game-based learning environment. In contrast, the Linear condition represents the traditional, fixed approach many games and online-learning tools use. Comparing the Adaptive and Linear conditions provides evidence that managing a player's engagement can increase performance without degrading the gameplay experience. The Choice condition represents an alternative approach to the Adaptive condition. It has been thought that providing player choice and autonomy in learning and gameplay can increase satisfaction and ultimately lead to positive learning outcomes (Deci & Ryan, 2000), however the results of Study Three show that providing choice in the puzzle game did not maximize performance compared to the other two conditions.

Though the summative measures did not provide a clear distinction among the three conditions, the performance measures and plotted visualizations show that players in each condition experienced very different gaming experiences. These differences, and their implications, are discussed below. In addition, reasons why the summative measures did not reveal these distinctions will also be discussed.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Summative Measures*

Analysis of the three summative self-report measures revealed no significant differences between the three conditions. This may be due to the favorable ratings players gave: the UES engagement measure revealed above average engagement (3/5), the affect measure revealed above average affect (~4.5/7), and TLX scores revealed players felt only a moderate amount of cognitive load (60/100). These ratings could be attributed to the overall high-entertainment value of the game. However, the real-time measures (discussed below) identified differences that the summative measures did not reveal. This highlights a limitation in the diagnosticity of these cumulative self-report measures and strengthens the case for including real-time measures when possible.

Overtime can be thought of as a type of summative measure in the sense that the amount of time a player spends over and beyond the minimum amount of required time could be used to indicate intrinsic motivation. Despite the positive summative self-report ratings, on average, players only played about 30 seconds into overtime in all three conditions. This indicates that, though the game might have been considered to be an overall positive experience, it was still not intrinsically motivating enough to keep a player involved past the required time. In addition, the Mechanical Turk environment in which this game was played may have exerted an external contextual pull (i.e., “Time is money”) that discouraged time spent on the game beyond the minimum required. It is

## REAL-TIME PREDICTORS OF ENGAGEMENT

interesting to note that a few players in the Adaptive condition played the game well beyond the minimum time.

For analytical purposes, the adaptive algorithm ran for all three conditions, but only affected players in the Adaptive condition. An overall engagement score was recorded after each level played. This score was used to determine whether a player should receive an easier, moderately difficult, or very difficult next level. It was found, that on average, players in the Linear condition would have been given easier levels by the algorithm compared to those in the other two conditions. Additionally, those in the Choice condition should have been playing very difficult levels—according to the algorithm—more than those in the other two conditions. The fact that the adaptive algorithm’s summative score (but not the players’ self-report) revealed significant differences between the conditions provides some evidence that the self-report measures are unable to capture the subtle differences in engagement experienced across the three conditions. For example, at the surface level, the game interface and mechanics appeared to be the same between all three conditions yet the real-time measures proved to be more beneficial in measuring levels of engagement that players experienced.

### *Real-Time Measures*

Results from the sixth and seventh hypothesis seem to indicate that the static, linear design of the first condition created a level of challenge (via high mental workload)

## REAL-TIME PREDICTORS OF ENGAGEMENT

that exceeded that of both the Choice and Adaptive conditions. Those in the Linear condition took, on average, longer to complete their levels. In addition, the reaction times on the secondary task show that those in the Linear condition performed significantly worse compared to those in the other two conditions. This indicates that those in the Linear condition experienced greater levels of cognitive load. While one might assume that higher cognitive load would lead to higher performance, this was not, in fact, the case. It is reasonable to conclude that the adaptive algorithm produced a more useful scaffolding design by matching individual abilities using the real-time engagement measures.

As predicted, those in the Adaptive condition required fewer levels to achieve higher degrees of difficulty. Those in the Linear condition fell short of the performance of the Adaptive condition even though a number of measures showed higher mental workload on the primary task of solving the puzzle levels. It seems that the Adaptive algorithm was able to appropriately respond to both the performance and mental state of the player and offer a level that provided “achievable difficulty” without either overloading or disengaging the player.

On average, those in the Choice condition played the highest number of levels yet achieved the lowest degree of difficulty. Providing choice/control has often been considered to be a successful approach to gameplay and learning. These results, however,

## REAL-TIME PREDICTORS OF ENGAGEMENT

show that providing choice can sometimes lead to sub-optimal performance, perhaps due to a lack of motivation to pursue increasingly difficult challenges. Previous research (Sharek, 2010) led to the interpretation of this behavior as effortlessness or mindless engagement. In other words, players enjoy their experience, yet they lack the motivation or desire to pursue more difficult challenges. This behavior is perfectly acceptable for some purely enjoyment-based video games, but it is not ideal for serious games and other types of online training where learning outcomes are desired.

Those in the Choice condition, on average, selected easier levels to play compared to the levels their counterparts were served (based on their engagement) in the Adaptive condition. This indicates that giving players a choice in similar tasks might not be the best approach if optimal performance is required – such as in learning tasks. The Adaptive algorithm kept players in a fairly tight range of difficulty that sent them on a difficulty trajectory more challenging than the fixed Linear condition. Additionally, those in the Adaptive condition were served more difficult levels than what a linear design would have produced indicating that leveraging a linear increase in difficulty may not match a player's potential. That is, this non-linear trajectory many of those in the Adaptive condition followed resulted in lower cognitive load compared to those in the slightly less challenging Linear path. Going back to the summative measures, even though those in the Adaptive condition were served more difficult levels in a shorter

## REAL-TIME PREDICTORS OF ENGAGEMENT

period of time, they did not express feeling negative affect or rate their experience as less engaging compared to those in the other two conditions.

Based on this paper's working definition of engagement, the adaptive algorithm effectively took into account a player's cognitive load, their affect, and their performance. A reasonable explanation for why this algorithm was successful is that it managed a player's cognitive load so that they never experienced frustration or effortlessness for more than one level. This reflects the Flow Theory model (see Figure 5) by balancing skill and challenge while still allowing players to experience the fringes of the flow channel where players may experience easier or more difficult levels from time to time.

In summary, the real-time measure of engagement, as defined in this study, provided a successful framework from which to build an adaptive game algorithm. Players in the Adaptive condition performed better by achieving greater difficulty in a fewer number of levels. They performed well in the secondary task, indicating that they were not cognitively overloaded, and they completed individual levels in less time compared to those in the other two conditions. The fact that there were no significant differences between the Linear and Adaptive condition in regards to the self-report measures indicates that implementing an adaptive algorithm will likely produce better

## REAL-TIME PREDICTORS OF ENGAGEMENT

performance with little-to-no cost to the playability (player perceived experience) of the game.

### General Discussion

The goal of these three experiments was to further current theories of engagement and test the efficacy of an empirically-based approach to increasing player engagement through the development and application of a real-time cognitive load, affect, and performance measure. Results from Study Three show that this measure can be successfully operationalized in a game-based environment to automatically adjust game difficulty to the degree of perceived challenge a person willingly seeks to experience while playing. The phrase “willingly seeks” led the researcher to explore the comparison of having the player define the level of challenge (the Choice condition) to having a system-based algorithm calculate the level of challenge (the Adaptive condition). While it would seem to make sense to let the player make this decision, the results seem to indicate that players will make sub-optimal decisions when it comes to maximizing gameplay performance. It is likely that this sub-optimization may also carry over the game-based training environments.

Using the engagement measure as input, an adaptive algorithm, despite being somewhat rudimentary, produced a condition that promoted higher performance, and

## REAL-TIME PREDICTORS OF ENGAGEMENT

reduced cognitive load. Those in the Adaptive condition accessed greater difficulty in fewer levels compared to the other two conditions. It appears that the Adaptive condition was able to effectively ‘push’ players past the difficulty level they would normally choose to tackle without decreasing their overall engagement and cognitive load, as demonstrated by the summative self-report measures. This is a very positive result and begs the question: “why was the adaptive algorithm successful?” In order to answer this question, a comparison of the three conditions is provided below.

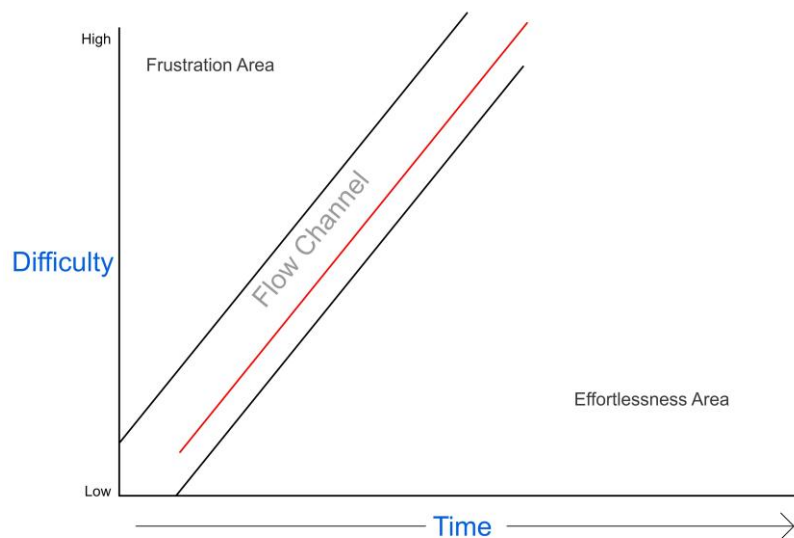


Figure 17. Linear Condition Overlaid on the Flow Theory Model

## REAL-TIME PREDICTORS OF ENGAGEMENT

Figure 17 shows the Linear condition overlaid on a conceptualized flow channel. Based on this diagram, it would appear that those in the Adaptive Condition never experienced frustration. This supports the assumption that, if designed well, a linear increase in difficulty can provide a scaffolding effect where people will be able to develop their skills in a gradual manner. What a linear increase design does not take into consideration is individual differences. In the case of the game used in this research, it appears that players were never overcome with frustration because they found the gradual increase in difficulty to be satisfactory. This is supported by their ratings of the summative engagement and affect measures. It is quite possible that those in the Linear condition experienced a degree of Flow. However, in a learning context, just being in the Flow may not be enough to push a person into situations where their full potential may be realized. This is where Flow Theory and engagement part ways. In the context of learning and serious games, being engaged requires a person to not only be in the Flow but also to actively seek out more difficult challenges rather than simply balance their skill with the challenges of the task.

## REAL-TIME PREDICTORS OF ENGAGEMENT

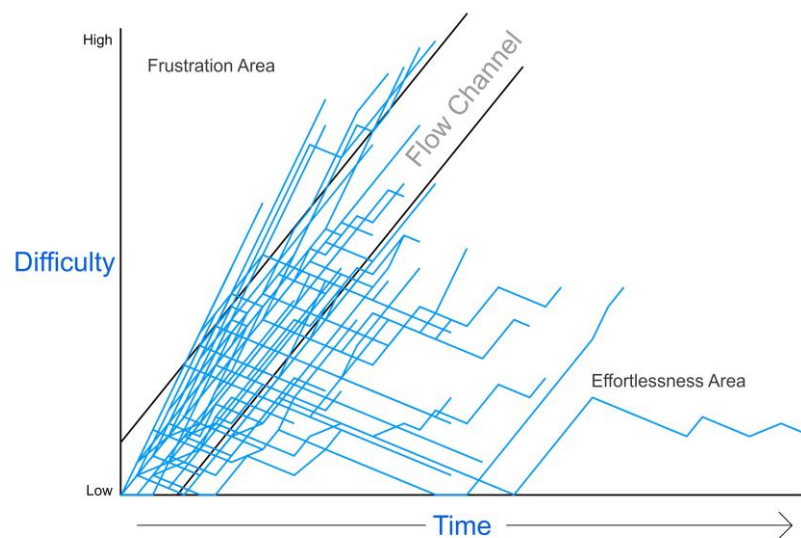


Figure 18. Choice Condition Overlaid on the Flow Theory Model

Figure 18 shows that many of those in the Choice condition were more inclined to enter a state of effortlessness rather than seek out more difficult challenges. It should be noted that there were quite a few players who indeed did seek out more difficult challenges and this variance in choice selection may be due to individual differences in regards to their desire for control (Burger, 1985; Burger & Cooper, 1979). This individual difference should be investigated in future studies. Past research has provided evidence that perceived autonomy (choice) can lead to an increase in intrinsic motivation and performance (Deci & Ryan, 2000; Reeve & Deci, 1996). However, in this game when players were given a choice to interact with a task effortlessly or seek out more difficult

## REAL-TIME PREDICTORS OF ENGAGEMENT

challenges, many of them opted for the less challenging option. Their intrinsic motivation was no different compared to those in the other conditions but their performance lagged behind. This may be due to their desire to succeed (Schüler, Sheldon, & Fröhlich, 2010) and since they were still receiving points in the game, they may not have felt the motivation to increase their skill past the minimum effort required to solve levels. This may also explain why those in the Choice condition played more levels compared to those in the other two conditions. The participant pool was sourced from the Mechanical Turk and so players were provided with an extrinsic reward (monetary payment) for playing a set amount of time, regardless of how well they performed. This contextual factor reflects real-world scenarios where employees are required to complete online training.

## REAL-TIME PREDICTORS OF ENGAGEMENT

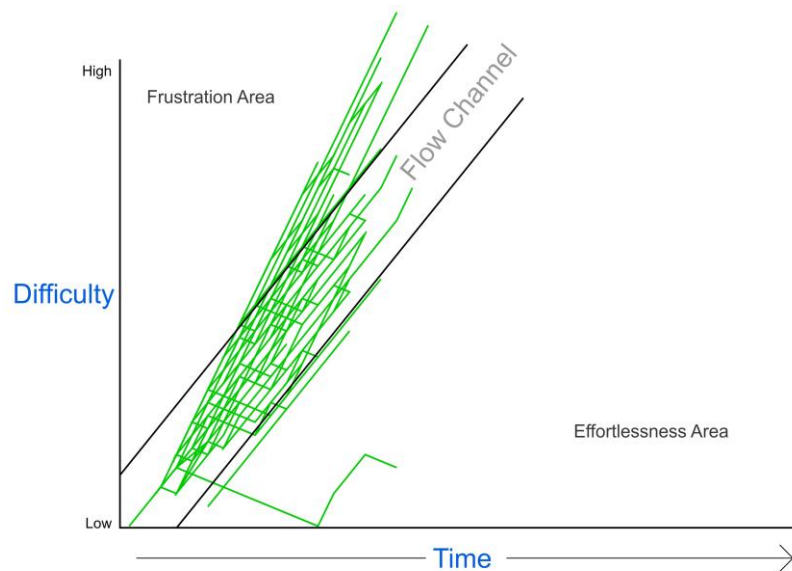


Figure 19. Adaptive Condition Overlaid on the Flow Theory Model

Figure 19 shows that those in the Adaptive condition were more likely to stay in the Flow channel. Some players exhibited greater skill as evidenced by the steeper paths. Based on a player's individual skill level, the adaptive algorithm was able to effectively shift the flow channel in order to keep a player in the flow. The application of the engagement measure allowed players to not only maintain a state of flow, but also to outperform those in the other two conditions.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Application*

There are two major advantages to pursuing the development of real-time engagement data collection in video games. The first advantage is that data collected during gameplay could be analyzed during playtesting to help developers identify parts of a game that players find to be too easy or too difficult. This actionable data could be used to quickly influence game development decisions as part of an iterative design methodology such as Microsoft's Rapid Iterative Testing and Evaluation (RITE) method.

The second advantage to collecting real-time engagement data lies in the use of the data as part of an adaptive algorithm. Discreet levels could be served to the players based on an algorithm using the real-time engagement data. For example, if a player is performing poorly in the secondary task (as indicated by higher response times), the algorithm could determine that the next level they play should be slightly less difficult. Incorporating performance data would provide an even clearer picture of the player's potential level of engagement. For example, if a player is performing poorly in the secondary task, but well in the primary task, the algorithm may determine that the level is maxing out the player's cognitive capacity. In this case, the next level served may need to be equally difficult, despite the poor performance in the secondary task.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### *Limitations and Future Research*

The limitations of this research begin with the application of the adaptive algorithm. It was essentially a rudimentary calculation based on an iterative test cycle. The weightings of the individual engagement measure inputs (affect, cognitive load, performance) could be further refined to produce even greater performance and potentially a more positive overall gameplay experience. Additionally, Game-clock affect measure requires a stopping point, such as the intermission between levels. Not all tasks are designed to have finite stopping points and the more fluid a task, the more difficult it would be with the current design to manage a person's engagement. Also, with this current design, the task would also need to contain multiple levels or learning objects of varying difficulties. This can be time consuming and, depending on the content, the sometimes subjective nature of what is difficult could be cumbersome in developing these ratings.

The data show that the use of a secondary task as a real-time measure of cognitive load during video gameplay is promising. Future studies should be conducted to determine the efficacy of data from the secondary task as input into an adaptive algorithm.

The Choice condition required players to select the level of difficulty they wished to play next. This action overlapped and interfered with the Game-clock checks making

## REAL-TIME PREDICTORS OF ENGAGEMENT

the affective measure potentially less useful. If choice was to be included in a game, an alternative method would need to be developed to allow for the Game-clock checks to be made in an uninhibited environment. Another factor potentially influencing the Choice condition (as well as the other conditions) was that all players were receiving an extrinsic monetary reward for playing the game for a set amount of time. This may have influenced level difficulty selection for those in the Choice condition as well as the willingness to play over and beyond the minimum time in any of the conditions.

Exploration of gameplay in alternative contextual settings may lead to different results.

### *Conclusion*

The efficacy of leveraging the psychological construct of engagement as a predictor of performance in the context of learning and video games was explored, defined, and tested. In this context, engagement was defined as the active state of seeking out a challenge. This definition was operationalized in terms of cognitive load, affect, and performance. All three of these variables are measurable and lend themselves well to the successful development and implementation of an adaptive algorithm that increased performance while maintaining a positive overall gaming experience. The results from this research are promising and provide a foundation for future engagement research.

## REAL-TIME PREDICTORS OF ENGAGEMENT

### REFERENCES

- A.S.T.D. (2010). State of the Industry Report: ASTD's Annual Review of Trends in Workplace Learning and Performance. Alexandria, VA.: American Society for Training & Development.
- Abt, C. C. (1970). *Serious Games* New York: Viking.
- Adams, R., & Granic, A. (2009). Cognitive Learning Approaches to the Design of Accessible e-Learning Systems. In C. Mourlas, N. Tsianos & P. Germanakos (Eds.), *Cognitive and Emotional Processes in Web-based Education: Integrating Human Factors and Personalization*. Athens, Greece: National & Kapodistrian University.
- Adelson, B. (1992). *Evocative Agents and Multi-Media Interface Design*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, Monterey, California, United States.
- Admiraal, W., Huizenga, J., Akkerman, S., & Dam, G. t. (2011). The Concept of Flow in Collaborative Game-based Learning. *Computers in Human Behavior*, 27(3), 1185-1194. doi: 10.1016/j.chb.2010.12.013
- Amazon.com. (2009). Mechanical Turk Retrieved May 2, 2009, from <https://www.mturk.com/mturk/welcome>
- Amazon.com. (2011). Mechanical Turk, from <https://www.mturk.com/mturk/>
- Aponte, M.-V., Levieux, G., & Natkin, S. (2009). Scaling the Level of Difficulty in Single Player Video Games Entertainment Computing – ICEC 2009. In S. Natkin & J. Dupire (Eds.), (Vol. 5709, pp. 24-35): Springer Berlin / Heidelberg.

## REAL-TIME PREDICTORS OF ENGAGEMENT

- Ballew, T. V., & Jones, K. S. (2006). Designing Enjoyable Videogames: Do Heuristics Differentiate Bad from Good. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 50, 646-650.
- Barker, P., & King, T. (1993). Evaluating interactive multimedia courseware--A methodology. *Computers & Education*, 21(4), 307-319.
- Behrend, T., Sharek, D., Meade, A., & Wiebe, E. (2011). The Viability of Crowdsourcing for Survey Research. *Behavior Research Methods*, 1-14. doi: 10.3758/s13428-011-0081-0
- Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The Development of the Game Engagement Questionnaire: A Measure of Engagement in Video Game-playing. *Journal of Experimental Social Psychology*, 45(4), 624-634. doi: 10.1016/j.jesp.2009.02.016
- Brunken, R., Paas, F., & Moreno, R. (2010). Current Issues and Open Questions in Cognitive Load Research. In J. Plass, R. Moreno & R. Brunken (Eds.), *Cognitive Load Theory* (pp. 253-272). New York: Cambridge.
- Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct Measurement of Cognitive Load in Multimedia Learning. [Article]. *Educational Psychologist*, 38(1), 53-61.
- Brunken, R., Seufert, T., & Paas, F. (2010). Measuring Cognitive Load. In J. Plass, R. Moreno & R. Brunken (Eds.), *Cognitive Load Theory* (pp. 181-202). New York: Cambridge.
- Burger, J. M. (1985). Desire for Control and Achievement-Related Behaviors. *Journal of Personality and Social Psychology*, 48(6), 1520-1533.
- Burger, J. M., & Cooper, H. M. (1979). The Desirability of Control. *Motivation and Emotion*, 3(4), 381-393.

## REAL-TIME PREDICTORS OF ENGAGEMENT

- Burgos, D., Tattersall, C., & Koper, R. (2007). Re-purposing existing generic games and simulations for e-learning. *Computers in Human Behavior*, 23(6), 2656-2667. doi: 10.1016/j.chb.2006.08.002
- Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary? In A. Mital (Ed.), *Advances in industrial ergonomics and safety*. London: Taylor & Francis.
- Champion, E. (2011). Game–Style Interaction Playing with the Past (pp. 83-128): Springer London.
- Chapman, P., Sanjeebhan, S., & Jane, W. (1999). *Engagement in Multimedia Training Systems*. Paper presented at the Proceedings of the 32nd Hawaii International Conference on System Sciences, Washington, DC.
- Charlton, J. P., & Birkett, P. E. (1999). An Integrative Model of Factors Related to Computing Course Performance. *Journal of Educational Computing Research*, 20(3), 237–257.
- Charlton, J. P., & Danforth, I. D. W. (2007). Distinguishing Addiction and High Engagement in the Context of Online Game Playing. *Computers in Human Behavior*, 23(3), 1531-1548.
- Chen, P., Lambert, A., & Guidry, K. (2010). Engaging Online Learners: The Impact of Web-based Learning Technology on College Student Engagement. *Computers & Education*, 54(4), 1222-1232. doi: 10.1016/j.compedu.2009.11.008
- Chen, S., & Macredie, R. (2002). Cognitive Styles and Hypermedia Navigation: Development of a Learning Model. *Journal of the American Society for Information Science and Technology*, 53(1), 3-15.

## REAL-TIME PREDICTORS OF ENGAGEMENT

- Choi, D. H., Kim, J., & Kim, S. H. (2007). ERP Training with a Web-Based Electronic Learning System: The Flow Theory Perspective. *International Journal of Human-Computer Studies*, 65(3), 223-243. doi: 10.1016/j.ijhcs.2006.10.002
- Ciavarro, C., Dobson, M., & Goodman, D. (2008). Implicit Learning as a Design Strategy for Learning Games: Alert Hockey. *Computers in Human Behavior*, 24(6), 2862-2872. doi: 10.1016/j.chb.2008.04.011
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an Integrative Theory of Training Motivation: A Meta-Analytic Path Analysis of 20 Years of Research. *Journal of Applied Psychology*, 85(5), 678-707.
- Cordova, D. I., & Lepper, M. R. (1996). Intrinsic Motivation and the Process of Learning: Beneficial Effects of Contextualization, Personalization, and Choice. *Journal of Educational Psychology*, 88(4), 715-730. doi: 10.1037/0022-0663.88.4.715
- Cross, J. (2004a). The Future of eLearning. *On the Horizon*, 12(4), 151-157.
- Cross, J. (2004b). An Informal History of eLearning. *On the Horizon*, 12(3), 103-110.
- Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*. San Francisco: Josey-Bass.
- Csikszentmihalyi, M. (1988). The Flow Experience and its Significance for Human Psychology. In M. Csikszentmihalyi (Ed.), *Optimal Experience: Psychological Studies of Flow in Consciousness* (pp. 15-35). Cambridge, UK: Cambridge University Press.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Csikszentmihalyi, M., & Csikszentmihalyi, I. (1988). Optimal Experience: Psychological Studies of Flow in Consciousness. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.). Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M., & Nakamura, J. (1989). The Dynamics of Intrinsic Motivation: A Study of Adolescents. In R. Ames & C. Ames (Eds.). New York: Academic Press.

Davis, J., Steury, K., & Pagulayan, R. (2005). A Survey Method for Assessing Perceptions of a Game: The Consumer Playtest in Game Design. *Game Studies*, 5(1).

Deci, E. (2009). Intrinsic Motivation Inventory Retrieved March, 30, 2009, from [http://www.psych.rochester.edu/SDT/measures/intrins\\_scl.html](http://www.psych.rochester.edu/SDT/measures/intrins_scl.html)

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227-268.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). *Using Heuristics to Evaluate the Playability of Games*. Paper presented at the CHI '04 extended abstracts on Human factors in computing systems, Vienna, Austria.

Dixit, P., & Youngblood, G. (2008). *Understanding Playtest Data Through Visual Data Mining in Interactive 3D Environments*. Paper presented at the 12th International Conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia and Serious Games, Louisville, USA

Drachen, A., & Canossa, A. (2009). *Analyzing User Behavior via Gameplay Metrics*. Paper presented at the Proceedings of the 2009 Conference on Future Play on @ GDC Canada, Vancouver, British Columbia, Canada.

Edwards, E. (2010). Why is Software Training so Boring? Retrieved from <http://info.alleninteractions.com/bid/34845/Why-is-software-training-so-boring>

## REAL-TIME PREDICTORS OF ENGAGEMENT

Ellis, G. D., & Voelkl, J. E. (1994). Measurement and Analysis Issues with Explanation of Variance in Daily Experience Using the Flow. *Journal of Leisure Research*, 26(4), 337.

ESA. (2009). ESA 2009 Annual Report: Entertainment Software Association.

ESA. (2010). Essential Facts About the Computer and Video Game Industry: Entertainment Software Association.

Fisher, S. L., Wasserman, M. E., & Orvis, K. A. (2010). Trainee Reactions to Learner Control: An Important Link in the eLearning Equation. *International Journal of Training and Development*, 14(3), 198-208. doi: 10.1111/j.1468-2419.2010.00352.x

Frankola, K. (2001). Why Online Learners Drop Out. *Workforce*, 80, 53-60.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming*, 33(4), 441-467. doi: 10.1177/1046878102238607

Gilleade, K., & Dix, A. (2004). *Using Frustration in the Design of Adaptive Videogames*. Paper presented at the Proceedings of Advances in Computer Entertainment Technology, Singapore.

Gilleade, K., Dix, A., & Allanson, J. (2005). *Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me*. Paper presented at the Proceedings of Digital Games Research Association, Vancouver, B. C., Canada.

Gove, P. (Ed.) (1963). Springfield, Massachusetts: G. & C. Merriam.

## REAL-TIME PREDICTORS OF ENGAGEMENT

- Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North Holland.
- Herrington, J., Oliver, R., & Reeves, T. C. (2003). Patterns of Engagement in Authentic Online Learning Environments. *Australian Journal of Educational Technology*(19(1)), 59–71.
- Huizinga, J. (1944). *Homo Ludens A Study of the Play-element in Culture*. London: Routledge.
- Hunicke, R. (2005). *The Case for Dynamic Difficulty Adjustment in Games*. Paper presented at the Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology, Valencia, Spain.
- ISFE. (2010). Video Gamers in Europe 2010: Interactive Software Federation of Europe.
- Jackson, S. A., & Marsh, H. W. (1996). Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. *Journal of Sport and Exercise Psychology*, 18, 17-35.
- Juul, J. (2003). The Game, the Player, the World: Looking for a Heart of Gameness. In M. Copier & J. Raessens (Eds.), *Level Up: Digital Games Research Conference Proceedings* (pp. 30–45). Utrecht, Holland: Utrecht University.
- Kalyuga, S., Chandler, P., & Sweller, J. (2004). When Redundant On-Screen Text in Multimedia Technical Instruction Can Interfere With Learning. *Human Factors*, 46(3), 567-581. doi: 10.1518/hfes.46.3.567.50405
- Koper, R., & Tattersall, C. (2004). New Directions for Lifelong Learning Using Network Technologies. *British Journal of Educational Technology*, 35(6), 689-700. doi: 10.1111/j.1467-8535.2004.00427.x

## REAL-TIME PREDICTORS OF ENGAGEMENT

Laurel, B. (1993). *Computers as Theatre*. Reading, MA: Addison-Wesley.

Lawless, K. A., & Brown, S. W. (1997). Multimedia Learning Environments: Issues of Learner Control and Navigation. *Instructional Science*, 25(2), 117-131. doi: 10.1023/a:1002919531780

Macey, W., & Schneider, B. (2008). The Meaning of Employee Engagement. *Industrial and Organizational Psychology*, 1(1), 3-30.

Malone, T. W. (1980). *What Makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games*. Paper presented at the Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems, Palo Alto, California, United States.

Malone, T. W. (1984). Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games *Human factors in computer systems* (pp. 1-12): Ablex Publishing Corp.

Mayer, R., & Moreno, R. (2010). Techniques that Reduce Extraneous Cognitive Load and Manage Intrinsic Cognitive Load during Multimedia Learning. In J. Plass, R. Moreno & R. Brunken (Eds.), *Cognitive Load Theory* (pp. 131-152). New York: Cambridge.

McCarthy, J., & Wright, P. (2004). *Technology as Experience*. Cambridge, Massachusetts: MIT Press.

McGinnis, T., Bustard, D. W., Black, M., & Charles, D. (2008). *Enhancing E-Learning Engagement Using Design Patterns from Computer Games*.

Milheim, W. D., & Azbell, J. W. (1988). *How Past Research on Learner Control can Aid in the Design of Interactive Video Materials*. Paper presented at the National Conference for the Association for Educational Communications and Technology, New Orleans, LA.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Moreno, R., & Park, B. (2010). Cognitive Load Theory: Historical Development and Relation to Other Theories. In J. Plass, R. Moreno & R. Brunken (Eds.), *Cognitive Load Theory* (pp. 9-28). New York: Cambridge.

MPAA. (2010). Theatrical Market Statistics: Motion Picture Association of America.

Nacke, L., Drachen, A., Kuikkaniemi, K., Niesenhaus, J., Korhonen, H., van den Hoogen, W., . . . de Kort, Y. (2009). *Playability and Player Experience Research*. Paper presented at the Proceedings of DiGRA 2009: Breaking New Ground: Innovation in Games, Play, Practice and Theory, London, UK.

Novak, J. (2008). *Game Development Essentials* (Second ed.). Clifton Park, NY: Delmar Cengage Learning.

Novak, T. P., Hoffman, D. L., & Yung, Y.-F. (2000). Measuring the Customer Experience in Online Environments: A Structural Modeling Approach. *Marketing Science*, 19(1), 22-42.

O'Brien, H. L., & Toms, E. G. (2008). What is User Engagement? A Conceptual Framework for Defining User Engagement with Technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938-955.

O'Brien, H. L., & Toms, E. G. (2010). The Development and Evaluation of a Survey to Measure User Engagement. *Journal of the American Society for Information Science and Technology*, 61(1), 50-69. doi: 10.1002/asi.21229

Ogden, G., Levine, J., & Eisner, E. (1979). Measurement of Workload by Secondary Tasks. *Human Factors*, 21, 529-548.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38(1), 1-4.

## REAL-TIME PREDICTORS OF ENGAGEMENT

Paas, F., & van Merriënboer, J. (1994). Variability of Worked Examples and Transfer of Geometrical Problem-solving Skills: A Cognitive-load Approach. *Journal of Educational Psychology*, 86(1), 122-133.

Pavlas, D., Heyne, K., Bedwell, W., Lazzara, E., & Salas, E. (2010). *Game-based Learning: The Impact of Flow State and Videogame Self-efficacy*. Paper presented at the Human Factors and Ergonomics Society 54th Annual Meeting, San Francisco, CA.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A Motivational Model of Video Game Engagement. *Review of General Psychology*, 14(2), 154-166. doi: 10.1037/a0019440

Rabin, S. (2005). *Introduction to Game Development*. Boston: Charles River Media.

Reeve, J., & Deci, E. L. (1996). Elements of the Competitive Situation that Affect Intrinsic Motivation. *Personality and Social Psychology Bulletin*, 22(1), 24-33. doi: 10.1177/0146167296221003

Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. In P. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 185-218). Amsterdam: Elsevier.

Richardson, J. C., & Newby, T. (2006). The Role of Students' Cognitive Engagement in Online Learning. *American Journal of Distance Education*, 20(1), 23-37.

Sansone, C., Fraughton, T., Zachary, J., Butner, J., & Heiner, C. (2011). Self-regulation of Motivation when Learning Online: The Importance of Who, Why and How. *Educational Technology Research and Development*, 59(2), 199-212. doi: 10.1007/s11423-011-9193-6

Schell, J. (2008). *The Art of Game Design: A Book of Lenses*. Burlington, MA: Morgan Kaufmann.

## REAL-TIME PREDICTORS OF ENGAGEMENT

- Schüler, J., Sheldon, K. M., & Fröhlich, S. M. (2010). Implicit Need for Achievement Moderates the Relationship Between Competence Need Satisfaction and Subsequent Motivation. *Journal of Research in Personality*, 44(1), 1-12. doi: DOI: 10.1016/j.jrp.2009.09.002
- SGL. (2011). The Emergence of a Serious Games Industry. Retrieved November 2011, from <http://www.seriousgames.org/about2.html>
- Sharek, D. (2009). Block Walk (Version 1.0) [Computer Game]. Raleigh, NC.
- Sharek, D. (2010). *The Influence of Flow in the Measure of Engagement*. M.S. Master's Thesis, NCSU, Raleigh, NC. Retrieved from <http://www.lib.ncsu.edu/theses/available/etd-02182010-132225/>
- Sharek, D. (2011). GridBlocker (Version 1.0) [Computer Game]. Raleigh, NC.
- Sharek, D., & Wiebe, E. (2011). *Using Flow Theory to Design Video Games as Experimental Stimuli*. Paper presented at the Human Factors and Ergonomics Society 55th Annual Meeting, Las Vegas, NV.
- Skadberg, Y. X., & Kimmel, J. R. (2004). Visitors' Flow Experience while Browsing a Web site: Its Measurement, Contributing Factors and Consequences. *Computers in Human Behavior*, 20(3), 403-422.
- Tsang, P. S., & Velazquez, V. L. (1996). Diagnosticity and Multidimensional Subjective Workload Ratings. *Ergonomics*, 39(3), 358 - 381.
- Webster, J., & Ho, H. (1997). Audience Engagement in Multimedia Presentations. *SIGMIS Database*, 28(2), 63-77. doi: <http://doi.acm.org/10.1145/264701.264706>

## REAL-TIME PREDICTORS OF ENGAGEMENT

- Webster, J., Trevino, L. K., & Ryan, L. (1993). The Dimensionality and Correlates of Flow in Human-computer Interactions. *Computers in Human Behavior*, 9(4), 411-426.
- Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing Online Games Against Computer- vs. Human-controlled Opponents: Effects on Presence, Flow, and Enjoyment. *Computers in Human Behavior*, 24(5), 2274-2291. doi: DOI: 10.1016/j.chb.2007.11.002
- Welsh, E. T., Wanberg, C. R., Brown, K. G., & Simmering, M. J. (2003). E-learning: Emerging Uses, Empirical Results and Future Directions. *International Journal of Training and Development*, 7(4), 245-258. doi: 10.1046/j.1360-3736.2003.00184.x
- Whitton, N. (2011). Game Engagement Theory and Adult Learning. *Simulation & Gaming*, 42(5), 596-609. doi: 10.1177/1046878110378587
- Wiebe, E. N., Branoff, T. J., & Shreve, M. A. (2011). Online Resource Utilization in a Hybrid Course in Engineering Graphics. *Advances in Engineering Education*, 2(3).
- Wiebe, E. N., Roberts, E., & Behrend, T. S. (2010). An Examination of Two Mental Workload Measurement Approaches to Understanding Multimedia Learning. *Computers in Human Behavior*, 26, 474-481.

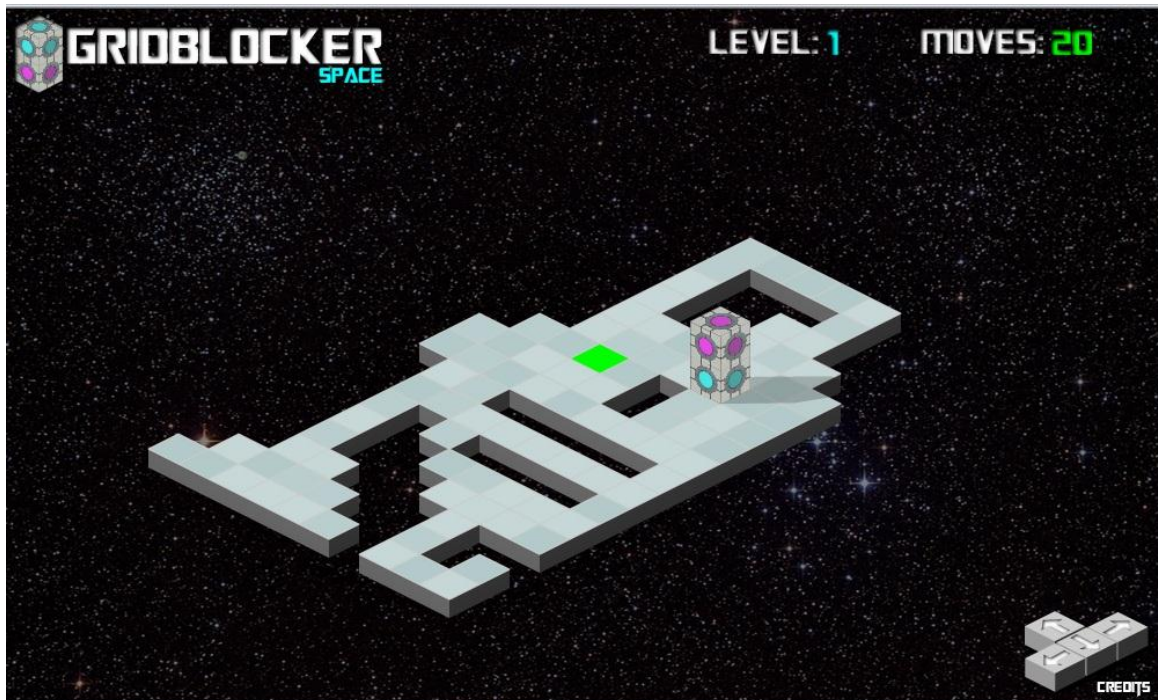
# REAL-TIME PREDICTORS OF ENGAGEMENT

## APPENDICIES

# REAL-TIME PREDICTORS OF ENGAGEMENT

## Appendix A: Study One Screenshots

### GridBlocker (Sharek, 2011) Game Board



### Instructions

**GAME ELEMENTS**  
Mouse over each element to learn more about them. Also, during the game, you can mouse over the block and goal elements to be reminded about what they do.



**CONTROLS**  
Use your keyboard's arrow keys to move the block up, down, left, and right. The image at right shows the direction that each key will move the block.



**LEVELS**  
There are several levels of varying difficulty in this game. Good luck!

**START GAME**



# REAL-TIME PREDICTORS OF ENGAGEMENT

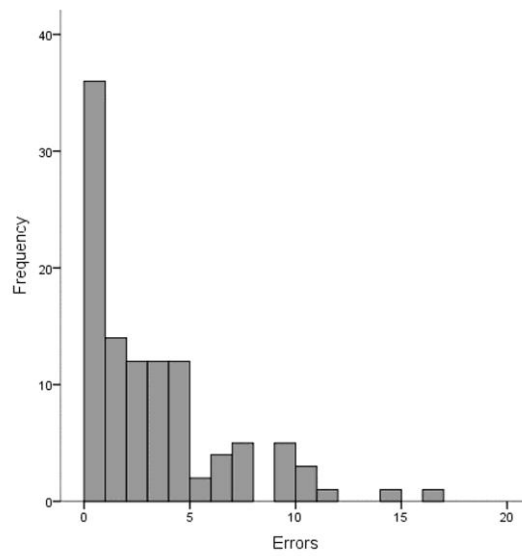
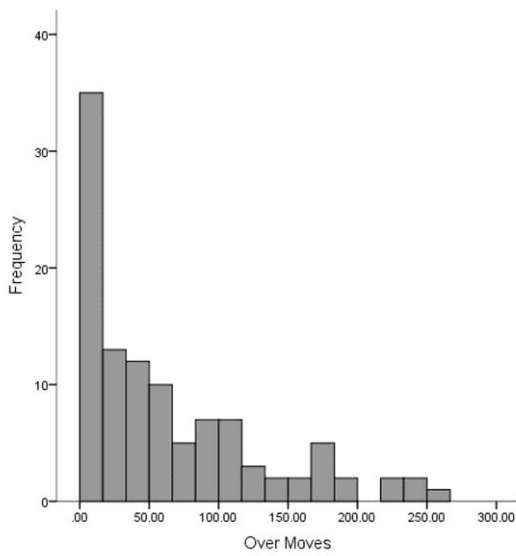
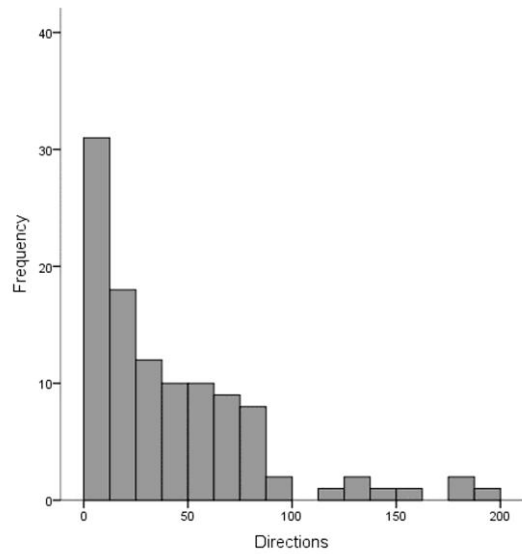
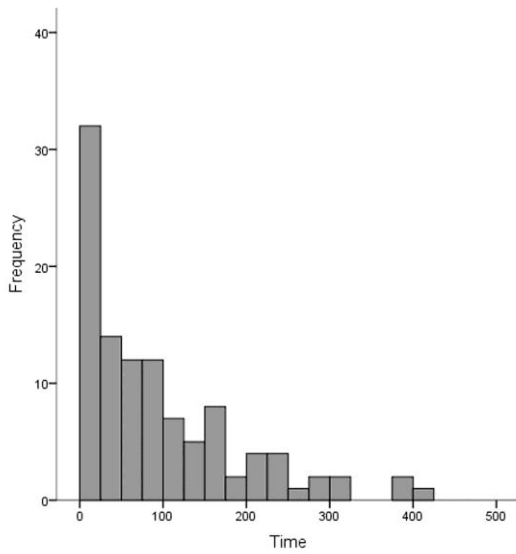
## Appendix B: Study One Informed Consent Form

<b>INFORMED CONSENT FORM for RESEARCH</b> <b>North Carolina State University</b> VIDEO GAME STUDY	
<p><b>What are some general things you should know about research studies?</b> You are being asked to participate in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate or to stop participating at any time. The purpose of research studies is to gain a better understanding of a certain topic or issue. You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those that participate. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask the researcher for clarification or more information. You may print this consent form using your Web browser's print option. If at any time you have questions about your participation, do not hesitate to contact the researcher named below.</p> <p><b>What is the purpose of this study?</b> You are invited to participate in a research study. The goal of this study is to gather ratings of individual game levels after they have been played. This information will be used to further develop game levels for the game, gridBlocker.</p> <p><b>What will happen if your child takes part in the study?</b> If you agree to participate in this study, you will be asked to play a web-based puzzle game. The goal of the game is to move a 3D block along a game board and position it over the goal. This game can help improve strategy skills and the ability to determine how a 3D shape would appear when rotated. After each level you will be asked to rate your experience with the level. When you have completed the task, you will be asked to answer a few questions about yourself (such as your gender, age) and how the game made you feel including how much effort you felt like you put into playing the game. This study should not take longer than 20 minutes of your time.</p> <p><b>Risks:</b> There are no risks or discomforts associated with this study, although some of the levels may be more difficult to play and may cause minor frustration.</p> <p><b>Benefits:</b> This research will be used to help us understand how to successfully design a game that people find interesting and fun to play. We will be very interested in hearing what you think.</p> <p><b>Confidentiality:</b> The information in the study records will be kept confidential to the full extent allowed by law. Data will be stored securely in a secured database. No reference will be made in oral or written reports which could link you to the study. You will NOT be asked to enter your name in any study materials so that no one can match your identity to the answers that you provide. Your computer IP address is not captured or stored.</p> <p><b>Compensation:</b> If you are participating in this experiment through Amazon's Mechanical Turk program, you will receive the compensation described in the HIT page once you have entered the experimental completion code back into the HIT.</p> <p><b>What if you have questions about this study?</b> If you have questions at any time about the study or the procedures, you may contact the researcher, David Sharek, at <a href="mailto:mturk@playgraph.com">mturk@playgraph.com</a> or Eric Wiebe at <a href="mailto:eric_wiebe@ncsu.edu">eric_wiebe@ncsu.edu</a>.</p> <p><b>What if you have questions about your rights as a research participant?</b> If you feel you have not been treated according to the descriptions in this form, or that your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919.515.4514).</p> <p><b>Consent To Participate.</b> By clicking on the I CONSENT button below, you are agreeing with the following statement: <b>"I understand that I have been asked to participate in a study. I have read and understand the above information. I know that I can print a copy of this form. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled."</b></p>	
<b>I CONSENT</b>	<b>I DO NOT CONSENT</b>

# REAL-TIME PREDICTORS OF ENGAGEMENT

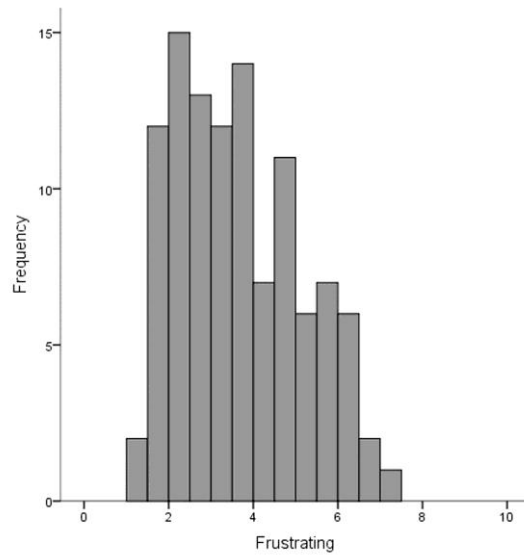
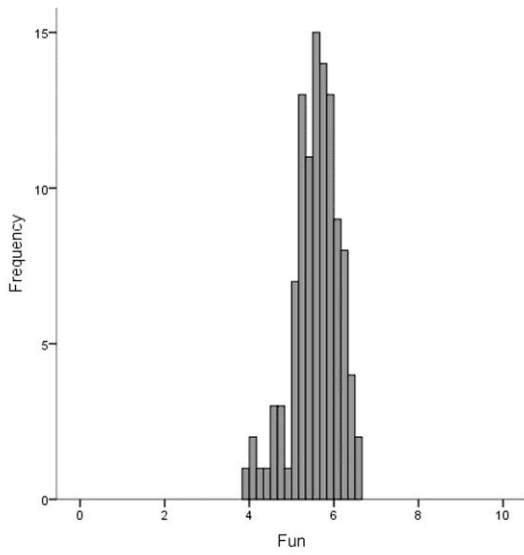
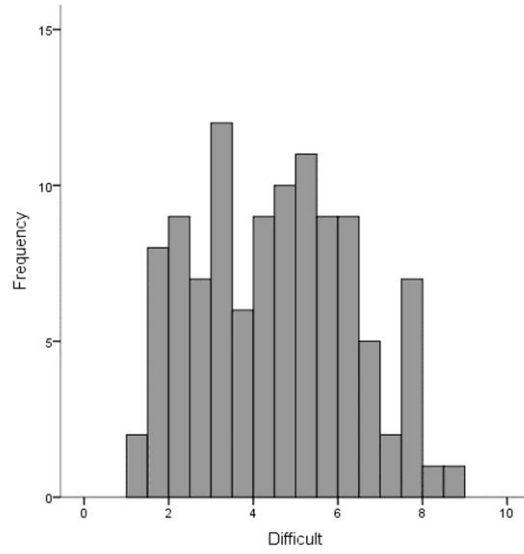
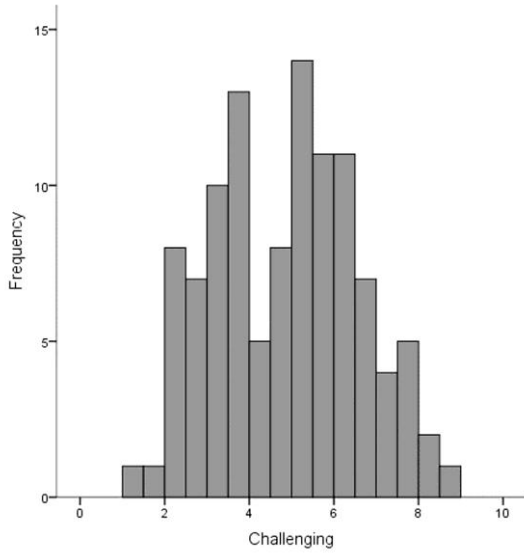
## Appendix C: Study One Performance and Self-report Distributions

### Performance Distributions



# REAL-TIME PREDICTORS OF ENGAGEMENT

## Self-report Distributions



## REAL-TIME PREDICTORS OF ENGAGEMENT

### Appendix D: Game Engagement Questionnaire (GEQ) Items

Item	Dimension
1. I lose track of time	Presence
2. Things seem to happen automatically	Presence
3. I feel different	Absorption
4. I feel scared	Absorption
5. The game feels real	Flow
6. If someone talks to me, I don't hear them	Flow
7. I get wound up	Flow
8. Time seems to kind of standstill or stop	Absorption
9. I feel spaced out	Absorption
10. I don't answer when someone talks to me	Flow
11. I can't tell that I'm getting tired	Flow
12. Playing seems automatic	Flow
13. My thoughts go fast	Presence
14. I lose track of where I am	Absorption
15. I play without thinking about how to play	Flow
16. Playing makes me feel calm	Flow
17. I play longer than I meant to	Presence
18. I really get into the game	Immersion
19. I feel like I just can't stop playing	Flow

(Brockmyer, et al., 2009)

## REAL-TIME PREDICTORS OF ENGAGEMENT

### Appendix E: Intrinsic Motivation Inventory

For each of the following statements, please indicate how true it is for you, using the following scale:

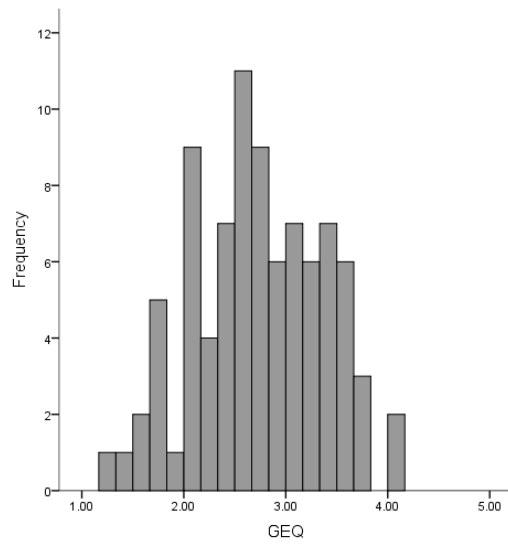
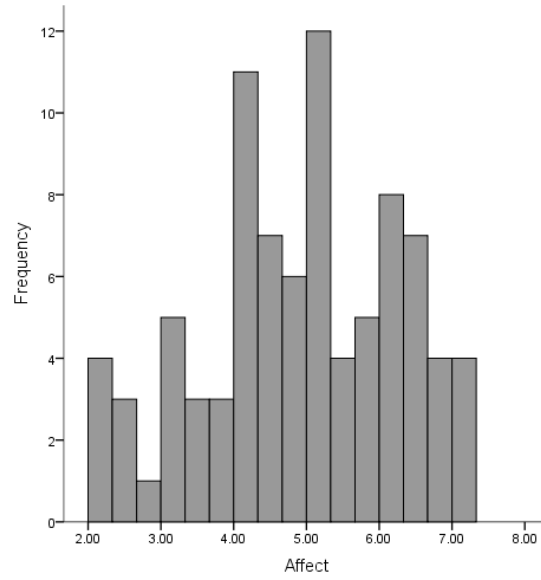
1	2	3	4	5	6	7
Not at all true		Somewhat true			Very true	

#### **Interest/Enjoyment**

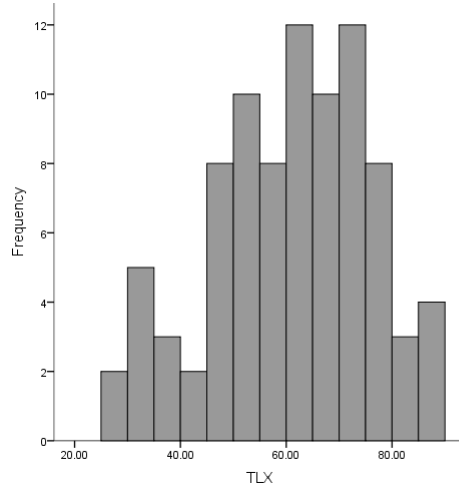
1. I enjoyed doing this activity very much
  2. This activity was fun to do.
  3. I thought this was a boring activity. (R)
  4. This activity did not hold my attention at all. (R)
  5. I would describe this activity as very interesting.
  6. I thought this activity was quite enjoyable.
  7. While I was doing this activity, I was thinking about how much I enjoyed it.
- (Deci, 2009)

# REAL-TIME PREDICTORS OF ENGAGEMENT

## Appendix F: Study Two Self-report Distributions



# REAL-TIME PREDICTORS OF ENGAGEMENT



## REAL-TIME PREDICTORS OF ENGAGEMENT

### Appendix G: User Engagement Scale

The scale was administered using a five-point scale with “strongly disagree” and “strongly agree” at the respective endpoints. Items identified with an asterisk (\*) indicate items that were reverse-coded.

#### **Focused Attention**

1. I lost myself in this gaming experience. 1
2. I was so involved in the game that I lost track of time. 17
3. I blocked out things around me when I was playing the game. 2
4. When I was playing the game, I lost track of the world around me. 18
5. The time I spent playing the game just slipped away. 3
6. I was absorbed in the game. 19
7. During the gaming experience I let myself go. 4

#### **Felt Involvement**

8. I was really drawn into the game. 20
9. I felt involved in the game. 5
10. The gaming experience was fun. 21

#### **Novelty**

11. I continued to play the game out of curiosity. 6
12. The content of the game incited my curiosity. 22
13. I felt interested in the game. 7

#### **Endurability**

14. Playing the game was worthwhile. 23
15. I consider my gaming experience a success. 8
16. The gaming experience did not work out the way I had planned.\*24
17. The gaming experience was rewarding. 9
18. I would recommend this game to my friends and family. 25

#### **Aesthetics**

19. The game was attractive. 10
20. The game was aesthetically appealing. 26
21. I liked the graphics and images used in the game. 11
22. The game appealed to my visual senses. 27
23. The screen layout of the game was visually pleasing. 12

#### **Perceived Usability**

24. I felt frustrated while playing the game.\*28
25. I found the game confusing to use.\* 13
26. I felt annoyed while playing the game.\*29
27. I felt discouraged while playing the game.\*14
28. Playing the game was mentally taxing.\*30
29. The gaming experience was demanding.\*15
30. I felt in control of my gaming experience. 31
31. I could not do some of the things I needed to do in the game.\*16