修　士　論　文

Finding the Fun:
Gameful Design of Classroom Goal Structures for Motivating Student Performance

指導教員　　宮田一乘　教授

北陸先端科学技術大学院大学
知識科学研究科知識科学専攻

1250209　Robert Wesley Songer

審査委員：　宮田　一乘　　教授（主査）
内平　直志　　教授
神田　陽治　　教授
永井　由佳里　教授

2015 年 8 月

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This research would not have been possible without the support and encouragement from my faculty advisor, Kazunori Miyata, as well as the faculty and staff at Kanazawa Technical College.

I would like to extend my sincerest gratitude to Professor Miyata for his attentive counseling and unwavering assistance since I joined his lab as a bright-eyed and hopeful amateur researcher looking for a place to make a mark. His means of guidance and active engagement with me have provided invaluable opportunities to see view my research and its academic context from alternate perspectives.

The administration at Kanazawa Technical College provided the chance to complete this research, and for that I am eternally indebted to their generosity. Maintaining my status as an instructor at Kanazawa Technical College has provided me with an advantageous viewpoint for my research which otherwise might not have been possible as a full-time student.

Finally, I would like to express my whole-hearted appreciation for the cooperation of the teachers at Kanazawa Technical College, Omihito Matsushita, Matthew Bailey, Sarah Lubold, Isaac Roelfsema, Nicholas Duff, and Jenny Brown, who welcomed my requests to administer experiments and surveys in their courses. Their collaboration was crucial to the success of my study, and their feedback was vital to the development of the gamification add-on.

Robert W. Songer
August 2015

1 INTRODUCTION

In recent years the term "gamification" has become a buzzword among companies and educators describing a new revolution in motivation and engagement. Described simply as "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled, & Nacke, 2011), gamification techniques have been criticized by researchers and game designers alike for their superficial application of game symbolism and operant conditioning style of motivating users. Many of its applications build game-like reward structures around an existing core activity, which is argued to be ultimately harmful to the user's well-being in the long-term. A cornerstone review of studies into human motivation shows that activities which satisfy psychological needs are more beneficial than those that only offer external rewards (Deci, Koestner, & Ryan, 1999).

Nevertheless, the success of the video game industry is proof that game design is effective at motivating and engaging users in seemingly insurmountable tasks. Educators in particular see great value in the potential to engage students in learning the same way they engage with video games. The traditional approach of using extrinsic rewards such as grade points and certificates of achievement is limited in its efficacy. One possible solution may come from the subcategory of gamification known as "gameful design" where the word "gameful" represents the experiential and behavioral qualities of gameplay (McGonigal, 2011). McGonigal describes, gamers as highly optimistic, empowered individuals with a strong sense of purpose and urgency for achieving a major victory. Games provide these experiences of striving towards some "epic win" not by implementing a reward structure, but instead through the design application of things like narrative, challenge, discovery and thrill. This is the basis of gameful design, and a major premise of this paper.

Much research still needs to be done before any conclusive arguments can be made for likening real-world activities to games. The existing research shows that the efficacy of gamification is highly conditional and depends on individual traits of the user as well as the context (Hamari, Koivisto, & Sarsa, 2014). An example of individual traits influencing a user's experience is the mood a person adopts, often

involuntarily, when approaching some activity. This relates to positive and negative affect, which research has shown to influence individual work behavior and enjoyment (Isen & Reeve, 2005). Contextual factors include those pertaining to the physical space of the activity, the associated social environment, and the psychological mindset of the participants. Two studies are presented herein with the aim of examining both individual and contextual factors that contribute to the efficacy of game-like experiences for engagement and learning.

We begin below with a discussion of existing research and background research that supports our premise of gameful design. After covering the similarities and differences of game-based learning, gamification, and gameful design, we then present our own findings from research that attempts to align the motivating aspects of games to the classroom. The contributions of this research include the results of two studies in which we examine contextual factors and individual traits related to the outcomes of gameful classroom activities. The first study examines aspects of the playful mindset of participants in a game-based learning activity and their relationships to enjoyment of the activity itself. For evaluation purposes, we developed the Playful Affordances Model as a synthesis of previous frameworks and design literature surrounding play and pleasure. The second study examines the individual traits of motivation, confidence, and affect in students enrolled in a course with a gamification element. These traits are identified and measured according to their correspondence with the elements of optimal experience, as described by flow theory (Csikszentmihalyi, 1990). Our findings have implications for the analysis and design of instructional, game-like activities in education and training.

2 LITERATURE REVIEW

This chapter presents foundational research that connects games, play, and learning as a background for our two studies of gameful design. We begin with a review of game-based learning research illustrating specific aspects of gaming and gameplay that are attractive to educators. This is followed by a discussion of gamification and its background, as well as the limitations it must overcome as described by its critics. Gameful design is then proposed as a user-centered design philosophy with similar goals to gamification. In order to promote game-like experience, though, gameful design requires an understanding of games and play which we give in a review of game design models and the philosophy of play. Case studies present the opportunity to deepen this understanding by analyzing what researches have attempted to accomplish in the past. Finally, we draw on this background to propose what we call "gameful learning" as the application of gameful design to promote optimal learning activities that simultaneously engage users and attend to individual psychological needs.

2.1 GAME-BASED LEARNING

Researchers of game-based learning advance the claim that games promote learning for a variety of reasons (Steinkuehler & Squire, 2014). Practical applications of such research attempt to exploit the learning that happens within a video game for educational purposes. One such game-based learning model proposes a cycle of user judgments, behavior, and feedback that may be employed to achieve desirable learning outcomes (Garris, Ahlers, & Driskell, 2002).

James Paul Gee writes extensively about the ways in which video games can provide great learning experiences (Gee, 2007). Good video games embody domain-specific meaning in objects and symbols in a way similar to specialty domains of various professions. Players can take on identities and role-play situations they might not otherwise get to experience in a classroom. Information is provided through various forms of media that form an interconnected network of situational meaning. The problems presented in games are well-defined subsets of larger problems with necessary information being provided both on-demand and just-in-time. The freedom of the fantasy element in video games allows for relevant cultural models to be

thought about critically within the context of the game. Finally, the act of gaming exercises the use of distributed knowledge networks consisting of tools, symbols, technologies, and even other players in the social group. Gee's principles are all backed by research in cognitive science and education.

Morris, Croker, & Zimmerman (2013) look at game design from a science education perspective and emphasize the ways that games can act as tools to support learning. Their premise is that games can contribute to scientific literacy by designing activities that enhance the acquisition of content knowledge, process skills, or an understanding of the nature of science. Furthermore, the study proposes three levels of scaffolding for game mechanisms: motivational scaffolds that engage students, cognitive scaffolds that compensate for the limits of human cognition, and metacognitive scaffolds that provide opportunities for identity association.

Another perspective proposes the idea of conceptual play spaces as a means of creating meaningful contexts for learning (Barab, Ingram-Goble, & Warren, 2009). Students project themselves into a fictional problem context where their ability to achieve the learning outcomes shows immediate effects in the environment. Example systems illustrate how scaffolds can be used to provide a perceptual environment, contextual details, rules and metrics to regulate activity, and engagement through interaction with other students.

The model presented by Garris, Ahlers, & Dreskell (2002) is based on a specific type of gameplay which breaks down under the scrutiny of an ecological perspective. Linderoth challenges the basic assumption that these cycles enabling a player's progression within a game is evidence of learning (Linderoth, 2012). If implemented correctly, the signals and affordances within a game environment enable players to navigate increasingly complex situations once their individual meanings can be recognized. Linderoth states that observation of such behavior has led to the assumption that games inherently direct continual learning. Studies in situated cognition suggest that this is actually a fundamental aspect of acquiring literacy in specialized contexts such as technology (Gee, 2010). Nevertheless, Linderoth's point is that the cues in video games may be designed to be understood quickly and used liberally within the game environment so that the player experiences little frustration, and consequently little learning as well.

The issue of contextual significance in a learning activity is another concern for game-based learning. It has been shown that different physical and social contexts surrounding two given tasks influence the ability of students to transfer learning (Klahr & Chen, 2011). Games and gameplay rely on setting contexts in the physical, social, and psychological dimensions (Stenros, 2012). The psychological dimension provides a frame of reference for the interpretation of game actions, effectively bounding the space in which game-based learning happens. This can be problematic when the fictional and imaginative nature of the game's narrative differs significantly from the authentic context in which the learning is meant to be applied. Games used to teach real-world skills require pre- and post-activities that scaffold and debrief the students such as those advocated in by Garris, Ahlers, & Driskell (2002), or risk losing their effectiveness to the issue of knowledge transfer.

2.2 GAMIFICATION

Traditional instances of gamification have focused on the use of visual elements that represent accomplishments in electronic games (e.g., points, badges, and leaderboards) and promoted their use in other contexts. Numerous definitions of gamification exist in various domains, such as web applications (Zichermann & Cunningham, Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps, 2011), organizational management (Kumar & Herger, 2013), and enterprise (Werbach & Hunter, 2012; Zichermann & Linder, 2013). The broadest and most inclusive definition of gamification is "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled, & Nacke, 2011). This non-prescriptive definition further describes "game design elements" that can be anything from superficial interface elements to complicated design activities. The levels are laid out in rising levels of abstraction as shown in Table 2-1.

At the most concrete level are "game interface design patterns" which include components from interaction design, such as points and badges. The next level is "game design patterns and mechanics" including parts of game design that emphasize gameplay. Beyond that are "game design principles and heuristics" that assist in the approach to design problems. "Game models" provide descriptions of abstract game systems and how they afford gaming experiences. Finally, the highest level of game design elements is "game design methods" which describe the practices

involved in the process of designing a game. According to Deterding et al., gamification is the strategy of using such elements without any explicit intent or purpose.

Table 2-1 Levels of Game Design Elements adapted from (Deterding, Dixon, Khaled, & Nacke, 2011)

Level	Description	Example
Game interface design patterns	Common, successful interaction design components and design solutions for a known problem in a context, including prototypical implementations	Badge, leaderboard, level
Game design patterns and mechanics	Commonly reoccurring parts of the design of a game that concern gameplay	Time constraint, limited resources, turns
Game design principles and heuristics	Evaluative guidelines to approach a design problem or analyze a given design solution	Enduring play, clear goals, variety of game styles
Game models	Conceptual models of the components of games or game experience	MDA; challenge, fantasy, curiosity; game design atoms; CEGE
Game design methods	Game design-specific practices and processes	Playtesting, playcentric design, value conscious game design

Findings from a meta-analysis of empirical studies aimed at the efficacy of gamification (Hamari, Koivisto, & Sarsa, 2014) suggest various external influences such as points, goals and feedback have generally positive effects on motivation. This analysis includes many diverse studies from multiple fields, but ultimately shows what little conclusive evidence there is surrounding gamification techniques. Although many of the studies included in the meta-analysis claim positive effects, their results appear to be heavily dependent upon the given context and user characteristics.

One possible explanation for characteristics that influence a user's reception of gamified systems may be player preferences. Several studies in the past have

attempted to understand why players engage in games and what makes them enjoyable. Bartle's Player Types (Bartle, 2004) is a classification of play styles for players of massively multiplayer online role-playing games (MMORPGs). Bartle created this two-dimensional characterization based on players acting on or interacting with other players or the world (see Figure 2-1). The names of each style are chosen according to the primary activity engaged in by the player. Killers prefer to exercise power over other players, engaging most frequently in player-versus-player content. Achievers enjoy completing quests and accomplishing goals set up by the designers. Socializers play MMORPGs for the chance to make friends and cooperate with other players. Finally, explorers are attracted by the lure of discovery in the content-rich setting of the game world.

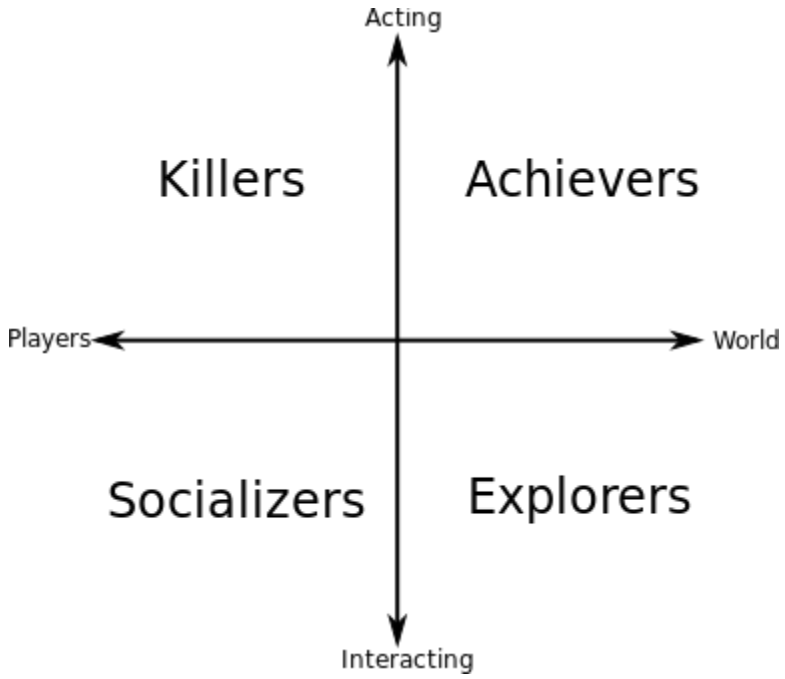

Figure 2-1 Bartle's Player Types

A more robust model of player motivations is Yee's facets (Yee, 2006) that describe various attributes of play styles present in every player to a certain degree. Yee's analysis identifies subcomponents within online gameplay ranging from the built-in systems of achievement to friendly or competitive forms of social activity, and immersive qualities such as escapism. The specific feedback from players reveals a highly diverse range of interests. For example, while some players are driven by the progression system of leveling up within these games, others find it to be tedious and

prefer to reach the max level as quickly as possible so they can enjoy content only available to high-level players. Both Bartle's player types and Yee's facets highlight specific motivational qualities designed into the gameplay of MMORPGs and reflect the diversity of player preferences. If these analyses can be transferred to users of gamified systems, they indicate that not all users will enjoy interacting with typical progression mechanics if the underlying activity itself is not intrinsically appealing.

The theoretical justification for gamification techniques often refers to a theory of motivation from behavioral psychology known as Self-Determination Theory (SDT). SDT is a macro-theory that describes the core incentives of human behavior to reside with needs for autonomy, competency and relatedness (Ryan & Deci, 2000). Expanding on this principle, the theory describes the contents of goals to be either intrinsic (a desirable pursuit for its own sake) or extrinsic (a means to achieve or avoid some consequential end state). A meta-analytic review of these distinctive traits in empirical research has shown not only that intrinsic motivation is more beneficial to a person's well-being, but also that extrinsic motivations tend to inhibit or distract from intrinsic motivations (Deci, Koestner, & Ryan, 1999). The review found that extrinsic regulations (e.g. rewards) have the potential to motivate a subject when perceived as informational and thus promoting a sense of competency. Detrimental effects occur when the subject perceives the extrinsic factor instead as controlling and thus inhibiting a sense of autonomy.

SDT has been related to a wide variety of fields including the issue of engagement with electronic games. One such study looks at habitual game play through the lens of SDT and reveals engagement with electronic gaming to be linked to the satisfaction of psychological needs necessary for well-being (Przybylski, Rigby, & Ryan, 2010). In the study, compulsive game play is portrayed as a substitution for feelings of satisfaction surrounding needs that are not met in daily life. This suggests that obsessive play is a symptom of a deprived well-being rather than a cause. Further empirical investigations in the same study emphasize the importance of these universal human needs by examining the effects of gaming on aggression. The results showed that gameplay which thwarts feelings of competency is a better predictor of aggressive behavior than the existence of violent content alone.

According to SDT, digital rewards similar to those that originate from game design are considered to be extrinsic regulators and may include anything from practical goods to intangible feelings of praise or social status. In contrast, an example of intrinsic motivation would be the positive feelings associated with learning, such as the delight of seeing a newly understood concept used effectively in a real context. Given that extrinsic rewards such as points, badges, and leaderboards are key elements of existing gamification practices, the field appears to be settled on a superficial foundation. Nevertheless, electronic games continue to be wildly popular on various platforms despite their uninhibited use of similar tactics. Tulloch (2014) argues that the game industry enjoys its success due to a long history of refining reward programs that signify player achievements and quantify progress. Such support is necessary for the onboarding of players engaging with video games that may be highly complicated systems with steep learning curves. Game designers recognize that humans are habitual creatures and that players often require incentives to complete learning tasks for progressing through the system and discovering advanced or emergent gameplay (Koster, 2004). Despite being composed of incremental levels of systemic complexity created by rules, games are inherently objects of entertainment and operate under the premise that the player must be engaged. From this observation, Tulloch offers an interpretation of gamification as "a form of training built upon the techniques used in, and heritage of, games rather than traditional pedagogy" (Tulloch, 2014, p. 326). While traditional pedagogy motivates students by awarding grades and credits, good games use rewards with endogenous value within the context of play. When implemented appropriately, such rewards can contribute to the intrinsic motivations of players through enhanced social interactions, self-reflection, fun, and experiences of flow (Wang & Sun, 2011).

Good video games provide a holistic experience in which an intrinsically rewarding activity is directed and enhanced by the added value of progression metrics and goal-oriented rewards. Many existing gamified systems have failed to recognize this intricate relationship and focus primarily on extrinsic rewards. The resulting applications then are criticized for having characteristics that betray the principles of good game design (Dichev, Dicheva, Angelova, & Agre, 2014). Such applications are superficial, adding a layer of game mechanics rather than taking a systematic approach to the design of the experience. They are reward-oriented, ignoring

opportunities for supporting intrinsically motivating experiences. They are not user-centric, promoting organizational goals of the producer rather than considering possible user goals. Finally, they are pattern-bound, uncritically adopting formulas for interaction design from popular gamification applications that came before.

2.3 GAMEFUL DESIGN

Given the broad definition of gamification as simply as the use of game design elements in non-game contexts, the term "gameful design" has been proposed for designing with the specific goal of game-like experience as an outcome (Deterding, Dixon, Khaled, & Nacke, 2011). The emphasis of gamification lies with the strategy of using game design elements, whereas gameful design explicitly sets the goal of achieving experiential and behavioral qualities similar to gameplay. Achieving such a goal likely involves the use of game design elements, so gameful design can be considered a subdomain of gamification.

The experience of the player is a major component of game design, as illustrated by the MDA framework (Hunicke, LeBlanc, & Zubek, 2004). This framework is named after the three distinct game components in an abstract model of games: mechanics, dynamics, and aesthetics. While representing different layers of abstraction in the composition of a game, the framework of MDA proposes that 1) games function more like artifacts than media, providing the players with opportunities for interaction rather than simply a conduit for experiences; and 2) developers approach a game differently than players do. As shown in Figure 2-2, developers build games from their most basic components (i.e., mechanics) while the players get into the game from the emotional responses it elicits (i.e., aesthetics). The dynamics serve as the middle-ground where players interpret possible action and consequence through examining the inputs and outputs of game mechanics. While game mechanics represent information about the internal state of the game and the algorithms pertaining to the game's flow, dynamics are the integration of these rules into a comprehensive gaming experience. The game of golf, for example, has the mechanics of hitting a ball with a club, a hole for the ball to ultimately arrive in, and a metric for measuring the number of swings. The dynamics are then represented by concepts residing at a higher level of complexity, such as gauging difficulty with par and striving for a hole-in-one. A player's experience of the game—the enjoyment of being

outdoors, the challenge of hitting the ball, and the sense of accomplishment when getting it in the hole—these signify the aesthetics, or experiential qualities, of golf.

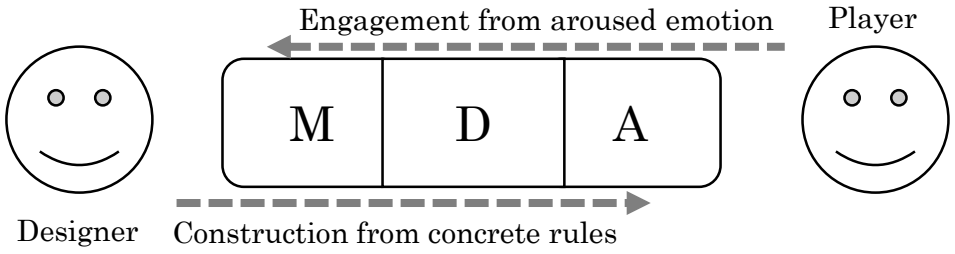

Figure 2-2 Different perspectives of a game according to the MDA framework, adapted from (Hunicke, LeBlanc, & Zubek, 2004)

The notion of the player's experience being essential to the composition of a game has been overlooked in traditional definitions of games and play. Long before the birth of electronic games, Johan Huizinga wrote about games as the frivolous activities that happen within the bounded ritual of play (Huizinga, 1955). His idea of the "consecrated spot" in which play happens is attributed as the foundational concept of the "magic circle" in game studies (Salen & Zimmerman, 2004). Roger Caillois echoes similar notions when describing the playing of games as "free (voluntary), separate (in time and space), uncertain, unproductive, governed by rules, make-believe" (Caillois, 1961, pp. 10-11). Furthermore, Caillois introduces a means of distinguishing games from play by placing the two concepts at opposite ends of a continuum. Games are represented by the Latin term "ludus" which emphasizes structured activities with explicit rules and goals. Play is represented in contrast by the Latin term "paidia" which describes the activity as spontaneous and unstructured. In the years since Huizinga and Caillois, numerous scholars have contributed to the definition of games by identifying characteristics that pertain to the structure of the activity itself. Few address a player's interaction with the system presented by the game, and those that do speak only in terms of goals or conflict (Juul, 2003). In contrast to these definitions, game designer Jesse Schell defines a game as "a problem-solving experience approached with a playful attitude" (Schell, 2008, p. 37) where "problem-solving" implies formal elements such as clear goals and boundaries, methods of solving, and the ultimate consequence of overcoming the problem or surrendering to its challenge. This definition is significant in its

recognition that the inextricable relationship between the game system and the disposition of the player is essential to the experience of a game.

In contrast to setting the experience as the desired outcome of a gamified system, popular approaches to gamification have extracted game elements from the systematic whole of the game with the expectation that they will recreate the same psychological influence of gameplay (Robertson, 2010). Operationalizing game mechanics outside their natural context in such a way raises the question of what may be stripped away in the resulting experience. Guided by Self-Determination Theory, one study examined the psychological effects of typical game interface elements and found an observable difference in user performance (Mekler, Brühlmann, Opwis, & Tuch, 2013). In the study, an image tagging experiment was set up to test the effects of points vs. no points as well as a meaningful frame vs. no frame of meaning. Subjects in the meaningful frame were told that their tagging results would contribute to a scientific cause. The results show that although there was no difference in user motivation, the groups in the meaningful frame produced higher-quality tags while the best results came from those who additionally received points for each tag.

A case study at the University Politechnica of Bucharest describes an instance of using game rules and constraints to redesign a mandatory work activity for university faculty members (Rughiniş, 2013). Researchers designated an explicit time and place for completing the required annual reviews of educational materials in a gamified way. Participants were brought in to a computer lab and placed in teams for the duration of the event. Each team had a single work station (computer and chair) at the beginning. They spent time reviewing the available materials and submitting revision tickets to other participants acting as arbiters. Each revision ticket would earn the team points for purchasing items that enabled further work, such as chairs, computers, food and drink. Through observing the activity and its participants, the authors illustrate contention between aspects of work and play. They observed a kind of "half-engagement" with the game rules and an emphasis on work objectives. For example, some participants would use chairs that had not been purchased or collaborated with members of other teams to improve revision accuracy. Rughiniş describes such behavior as the consequence of a "transparent magic circle" in which the physical, social, and psychological contexts of play overlap with those

of work, and an "ambivalent invitation" is extended for players to engage in a partial attitude of playfulness. The consequence of intersecting work and play thus may have the unintended consequence of goals not being taken seriously. Additionally, the work goals may be perceived as not being taken seriously as was the case here when a number of participants opted out of the gamified activity and did the work on their own.

As the above case study shows, the structural goals of a work activity may be convoluted by (possibly conflicting) structural goals of the game. Games create a holistic experience by providing clear rules and boundaries that give meaning to the actions of players. Consecutive transgressions into the state of play (i.e., free-form, exploratory activity) and then into the state of the game (i.e., engaging with a system) can be distinguished from the mundane state of the non-game activity (Walther, 2003). After performing these transgressions, the player must then balance between the mode of play, where the goal is to maintain a pleasurable state, and the mode of the game, where finer goals are defined by rules and structure. Without this balance, the autotelic qualities of the experience may be undermined by the rigidity of the game, or the importance of the goals may be diminished by the frivolity of play.

The psychological context of play is a reoccurring theme in the literature. An anthropological review of play illustrates an lack of consistency when distinguishing the act of playing a game from the mode of playful experience (Malaby, 2009). As an activity, the common perception of play is almost exclusive to the form of the activity irrespective of the attitude of the player. On the other hand, the mode of play is characterized by the player's readiness to improvise and seek creative order within an indeterminate scale of possibility—a psychological state of curious enthusiasm. Without assuming this state, engagement with a game system becomes mere operation even though the activity may be described as play. In such a case, the psychological context for play is not satisfied and the player's experience may not necessarily be one of enjoyment.

Michael J. Apter describes the playful mindset as "paratelic" and compares it to the serious, or "telic" mindset (Apter, 1991). A mindset that is telic will experience anxiety in states of high arousal whereas a mindset that is paratelic will experience

excitement in similar situations. This dual nature of human psychology is also reflected in other emotional situations, such as conforming or deviating from rules, which Apter describes in great detail with Reversal Theory (Apter, 2001). Here, the Greek word "telic" expresses the existence of an intended goal or purpose while the prefix "para-" indicates running alongside or contrary to said purpose. Additionally, the term "autotelic" has been used to characterize playful participation in an activity as both a means and an end within itself. Play creates meaning through internal and implicit communication among participants, encouraging further play and engagement through its autotelic nature (Bateson, 1972). Play begets play by maintaining a separation from the seriousness of normalized social environments and continuously providing a range of possibilities to be explored.

Considering the literature reviewed up to this point, we can say that the challenge for gameful design is to elicit a playful mindset without undermining the non-game purpose of the activity. Empirical evidence suggests that this may be achieved by simply creating a psychological frame of play without the adoption of game goals. This can be seen in the image tagging exercise mentioned above (in which the awarded points had no further meaning besides tracking progress), as well as another study where no significant difference was observed between a gameful framing condition and a full game condition (Lieberoth, 2015). In the study, student participants were asked to facilitate discussions surrounding responses to a school satisfaction survey. Two conditions were tested: a framing condition in which participants interacted with game artifacts such as cards, pawns, and a game board; and a full-game condition in which participants interacted with the same game artifacts and competed for a win condition. The control group had their discussion materials on simple paper and followed basic instructions for the activity. The study then collected self-reported data for intrinsic motivation and behavioral data of time spent on the optional continuation of the task. Compared to the control group, both game conditions had significantly higher ratings for the interest/enjoyment dimension of the intrinsic motivation survey, although the control group spent more of the optional time to complete tasks. All other intrinsic motivation scores (such as importance/effort, relatedness, and value/usefulness) were consistent across the three groups. Furthermore, the only observed difference between the two game

conditions was that the full-game condition received higher scores for being more "like a game".

2.4 GAMEFUL LEARNING

When trying to understand the motivation for engagement with video games, researchers surveyed self-identified gamers and found that dimensions of dispositional flow and intrinsic regulation are the most influential factors (Wang, Khoo, Liu, & Divaharan, 2008). Other factors investigated include harmonious or obsessive passion, extrinsic regulation, and positive affect. A cluster analysis was used to discover profiles of gamers with high, low, and average levels of both harmonious and obsessive passion for play. The resulting profiles all shared high scores in dispositional flow and autonomous/intrinsic regulation while controlled regulation was among the lowest. The measurements of flow used in this study include the nine dimensions originally proposed by Csikszentmihalyi (Flow, 1990):

1. A balance of challenge and skill
2. Merging of actions and awareness
3. Clear goals
4. Unambiguous feedback
5. Concentration on the task at-hand
6. A sense of control
7. Loss of self-consciousness
8. Transformation of time
9. Autotelic experience

It is suggested that the nine elements of flow can be divided into those that are conditions for achieving flow state and those that are outcomes of reaching flow. In order to achieve flow, the important elements are clear goals, unambiguous feedback, challenge-skill balance, and a sense of control. Reaching the flow state then produces the outcomes of action-awareness merging, total concentration, a loss of self-consciousness, and transformation of time. Studies suggest this categorization to be consistent with gamified experiences as well with the exception of autotelic experience. In pure gaming activities, autotelic experience appears to be an associated outcome; whereas, gamified exercise has shown that autotelic experience relates more strongly to the conditions than the outcomes (Hamari & Koivisto, 2014).

The theory of flow and its elements have been the center of much discussion around the design of game play experience in both professional game design (Salen & Zimmerman, 2004; Fullerton, 2008) and the application of game design elements to non-game contexts (Pavlas, 2010; Hamari & Koivisto, 2014). Given that educators recognize the value of several flow elements for learning (such as clear goals and unambiguous feedback), an experiential gaming model based on flow elements and playfulness has been proposed as a bridge between educational theory and game design (Kiili, 2005).

Although most elements of flow are explicitly defined, the element of autotelic experience seems vague and difficult to implement. In common terms, autotelic experience may be described simply as "fun" within a context of enjoyment. As mentioned in the discussion of the psychological context of play above, a playful mindset is integral to autotelic experience. Bateson asserts that the manner of framing an activity as play is more central to the experience than the activity itself (Bateson, 1972). Once the frame of mind has been established, then play becomes an exploration of possibilities within a determined space. A similar perspective is expressed by Csikszentmihalyi & Bennett when they define play as "a state of experience in which the actor's ability to act matches the requirements for action in his environment" (Csikszentmihalyi & Bennett, 1971, p. 45).

Following the previous discussions of game-based learning, gamification, and gameful design, we can see several issues related to the use of game design elements in learning contexts: the problem of knowledge transfer between differing contexts; the drawbacks of unoriginal, superficial applications of game-like rewards; and the complicated matter of the psychological mindset of play. We propose an approach that combines the psychological context of gaming with the physical and social contexts of a learning environment (Songer & Miyata, in press). The domain of an educational topic can be dissected into overlapping physical and social contexts similar to those of the "magic circle" of play. The physical dimension includes spatial and temporal boundaries surrounding the activity as well as the artifacts involved. The social dimension establishes social borders agreed upon by all participants which, in terms of a game, allows for the playful mindset to be shed during times of serious play. Whereas a typical educational game would produce its own artifacts symbolic of the target domain, an application of game design might instead

incorporate real artifacts from the target domain into a gameful activity. Similarly, the social boundaries of the game can be expanded to include actions and protocols typically used within the real-world target domain. This idea draws on concepts from alternate reality games and pervasive games, which actively explore the fusion of games with the real world (McGonigal, 2008; Montola, Stenros, & Waern, 2009).

3 PLAYFUL AFFORDANCES MODEL

In order to test the applicability of the gameful learning principle, we developed a model for the evaluation of gameful design, called the Playful Affordances Model (Songer & Miyata, 2014). The model embodies several concepts from the above discussion such as transgressing into the psychological context of play and promoting the autotelic qualities of game-like experiences. It aims to identify concepts of "fun" and "pleasure" by drawing from the philosophy of play, the design of interactive artworks, and the analysis of video games. An existing "pleasure framework" details 13 categories of pleasure determined by a literature review and proved through application (Costello & Edmonds, 2007). The result is a one-dimensional list of terms used to describe the various forms of pleasurable, playful experiences. The PLEX framework extends this list to cover the range of experiences specifically afforded by video games (Korhonen, Montola, & Arrasvuori, 2009). It was developed by analyzing engagement with video game systems and so follows the assumption of play as a form of activity, including experiences such as "suffering" or "completion" as types of playful experience. Such experiences may be enjoyable when a playful mindset is adopted; however, the frustrating activities included in the PLEX framework will likely be less pleasurable in non-game contexts.

In contrast to the previous frameworks, the Playful Affordances Model forms a multi-dimensional categorization of autotelic experiences identified from the literature. At the most abstract level, the four play categories of "agon", "alea", "mimicry", and "ilinx", which loosely translate into contest, chance, imagination, and vertigo, are adopted from philosopher Roger Caillois (Caillois, 1961). These categories provide themes for exploring terms from existing frameworks and grouping them based on similarity. A look at the resulting groups revealed two types of terms in each category: those expressing action and those expressing state. Both types were then generalized and matched as action-state pairs representing each of the four categories. Table 3-1 shows this categorization of terms as unordered lists under their respective themes. The bottom row shows the pairings of play behaviors and experiential states proposed as the representative concepts in each category.

Table 3-1 Play terms mapped to pairs of behavioral and experiential states

Agon	Alea	Mimicry	Ilinx
Challenge	Discovery	Fantasy	Sensation
Competition	Curiosity	Narrative	Simulation
Difficulty	Exploration	Fiction	Danger
Control	Risk	Creation	Sensory
Achievement & Completion	Beauty & Immersion	Cognitive Synergy	Physical Activity
Contest & Challenge	**Exploration & Discovery**	**Imagination & Creativity**	**Sensation & Arousal**

Now it is worth mentioning a number of terms that did not appear to fit into our selected categories of play. The first is the pleasure of socializing and fellowship, and the second is expression. Undoubtedly play is a major instrument in the construction of our social frameworks; however, we assert that the pleasure of fellowship belongs more to the social context of play than to the playful activity itself. Meanwhile, expression is an activity of play in itself since a multitude of playful actions (such as self-discovery and creation) exist as possible ways that expression is achieved, and so one enters into play by seeking out these ways to express oneself. Furthermore, the concept of expressing oneself suggests the existence of some partner or audience to whom the communication of expression is taking place. This again reflects a connection of play with its social context and further indicates that such qualities are emergent properties of the whole experience rather than of any single delineable instance. Just as play theorist Brian Sutton-Smith describes a variety of rhetoric that is constructed through play (Sutton-Smith, 2001), the complicated matters of self-expression and socialization can be seen as independent though not unrelated activities to which play is complementary. We therefore conclude that feelings of fellowship and expression are more appropriately described as significant outcomes of play within its social context. This is further supported by literature in play philosophy and radical game design (Flanagan, 2009; De Koven, 2013).

Two other terms that seem incompatible with our chosen categories of play are those of submission and negativism. Submission, as described by game designer Marc

LeBlanc[1], is the pleasure of surrendering oneself to play as a pastime. The opposite term, negativism, is described as provocative rule-breaking by Apter (1991). Our omission of these two terms is due to their dichotomous nature. Given that the activities of play can been seen as a subset of possibilities derived from a chaotic world (Malaby, 2009) or an unbounded region for the fabrication of rules (Walther, 2003), we see submission and negativism as relating directly to the player's interpretation of the play space. Where one player may take pleasure in abiding by the rules as the established bounds of activity, a deviant player will enjoy actively seek out ways to challenge and expand upon these limitations. For these reasons, we find submission and negativism to be better expressed as modalities of play existing to some degree irrespective of the activity type.

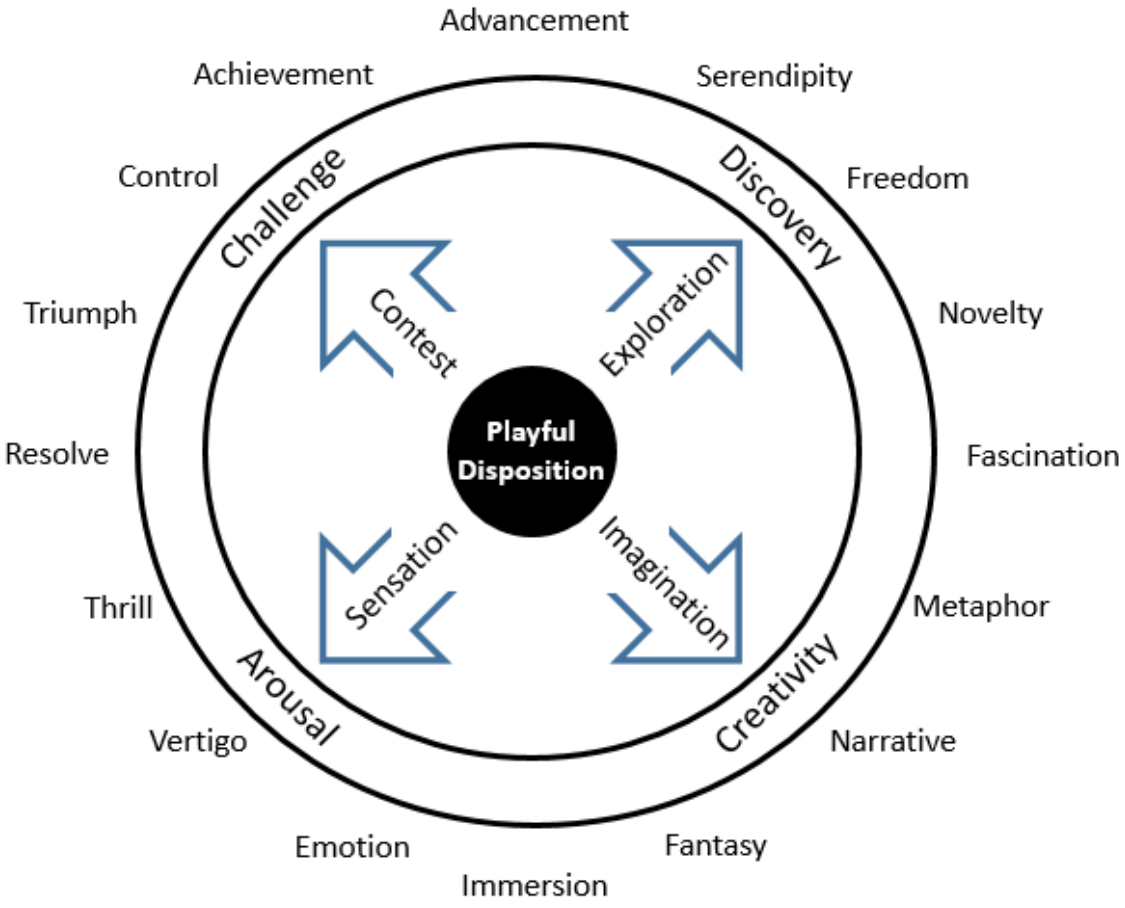

Figure 3-1 Playful Affordances Model

[1] http://algorithmancy.8kindsoffun.com/

When likened to the concept of design affordances, pairs of employable actions and resulting emotional states illuminate a way to achieve desired experiences by creating opportunities for certain types of playful actions. The representative pairs from the categories in Table 3-1 are proposed as "playful affordances" and form the basic components of the model. In addition, the prerequisite of a playful disposition is included to fully capture the act of transgressing into a state of play. The playful disposition is at the center of the experience with forms of activity branching out in a radial pattern to support the aesthetic states along the rim, as shown in Figure 3-1. Activities relating to contest, exploration, imagination, and sensation afford the expansion of playful experience when approached with a playful attitude. The resulting experiences are represented by the general terms of challenge, discovery, creativity, and arousal.

Around the rim of our model are more concrete, descriptive terms fitted according to their correspondence to the four experiential states. Since it is unrealistic to expect user experiences to be described with only the four terms of challenge, discovery, creativity and arousal, the expanded terms are provided as examples of where more specific terms might be situated with respect to others in the model. Furthermore, the terms at the cardinal points—advancement, fascination, immersion, and resolve—are given as examples that combine aspects of adjoining dimensions within the model. Advancement describes the experience of realizing one's skill and ability through putting it to the test, thus combining challenge and discovery. Fascination describes the lure of an imaginary context and desire to discover more about its intricacies. Immersion is the emotional or physical sensation in a non-real/imaginary context, and resolve is a strong urge to rise to a challenge.

Finally, as a radial pattern, the model is inherently multi-directional. We are in no way trying to imply that any given experience will expand out in only one direction from the center. Experiences of play are often best described as conglomerations of these affordances. A scavenger hunt illustrates this point as it provides players with opportunities for contest and exploration simultaneously. An Easter egg hunt is a particular example of a scavenger hunt that allows for imagination as the participating children believe they are searching for treasures hidden by an anthropomorphic rabbit. When the children then discover eggs planted in a nest up in a tree, they experience the sensation of exhilaration as they climb through the

obstacles set by the tree branches to reach their prize. Following this exercise of identifying experiences that include multiple styles of playful affordances, we would then hypothesize that the most fulfilling experiences are those that incorporate a broad range of the affordances present in the model.

4 STUDY 1: PLAYFUL AFFORDANCES

Our first study aims to identify the autotelic experiences of play afforded by the design of a game-based learning activity. The opportunity for this study came during an exchange program between Singapore Polytechnic and Kanazawa Technical College (KTC) when the Singaporean students participated in a technical English class at KTC. The teachers planned for the students to play a team-based business negotiations game, called *The Shosha*[2], in which players must trade cash, resources, and project cards to gather the requirements for establishing businesses. This instance of the game presented the opportunity to examine playful affordances for players with the personal trains of: no experience with the game; repeated experience with the game; native language skill; and low-to-intermediate language skill.

4.1 METHOD

The game consisted of three rounds. Each round included of a planning phase, in which players could only talk amongst their teams, and an action phase for making deals with other teams and completing sets. At the end of each round, the teams reported their score as the sum of cash on hand and the fixed assets of their established businesses. The winning team was determined according to the highest accumulated score after all three rounds. In this instance, the game was played in English with mixed teams of Singaporean and Japanese students.

The playful experiences were captured by a survey of the players conducted after the game had completed. The survey had a single pre-post evaluation construct—the player's anticipated enjoyment versus actual enjoyment—which was asked retrospectively due to time constraints. Survey items consisted of a total of 18 Likert-type questions: one for anticipated enjoyment before play; one for overall enjoyment after play; eight for behaviors engaged in during play; and eight for experiences of fun. The behavioral items reflected the action terms in the Playful Affordances Model—contest, exploration, imagination, and sensation. Each of the four behavioral items had a positively worded item and a negatively worded item as shown in Table

[2] http://www.projectdesign.co.jp/the-shosha (Japanese)

4-1, and each solicited a response on a scale from 0 (completely disagree) to 5 (completely agree). The experiential items were simple one-word descriptors chosen from the outer edge of the model to which respondents rated intensity experienced during play on the scale of 0 (not at all) to 5 (a great amount). Two terms were chosen from each of the four dimensions for a complete list of achievement, arousal, challenge, creativity, curiosity, discovery, fantasy, and thrill. The two items for expected and actual enjoyment used scales similar to the previous items, ranging from 0 (no enjoyment at all) to 5 (greatly enjoyable). Appendix A: Playful Affordances Survey contains the full text of the English and Japanese versions.

Table 4-1 Survey items and their corresponding behavioral terms

Survey Item	Behavioral Term
I put a lot of effort into performing as best I could in the game.	Contest
The game was too easy.	Contest
I was excited to make deals / establish businesses / work towards a high score.	Sensation
The game was too slow or boring for me.	Sensation
I tried various different ways to make deals / operate in my team.	Exploration
I did not change my tactics during the game.	Exploration
I do not care for the business theme in the game.	Imagination
I could imagine what it must be like to form a business / be a businessman.	Imagination

In this instance, the Singaporean students had no previous experience with the game while the Japanese students were playing it for their fourth time. The previous three times were done in the same course over a two-month span prior to the arrival of the Singaporean students. The first instance was performed in Japanese so the students could get accustomed to the rules. The second and third instances were in English with translation sheets of common negotiation phrases. In between the second and third instances, the students participated in a focused scenario activity for practicing the specific phrases used when making deals.

After the play finished, the students participated in a reflection activity in which they discussed the qualities of other players they had recognized as good business

partners. The survey was distributed after the reflection activity completed, and it was collected again within the same day. All students responded to items written in the language for which they are most proficient (English for the Singaporean students and Japanese for the Japanese students).

The responses were analyzed in subsets defined by nationality as well as self-reported anticipation of enjoyment. The primary statistic used in the data analysis was Pearson correlation values. These values were calculated from survey item scores of actual enjoyment versus playful experiences; behaviors and their respective experiential terms from the Playful Affordances Model; and pairs of experiential terms taken from same dimensions of the model. Subsets to be analyzed were determined to be those from Singaporean students, Japanese students, and students who anticipated low, medium, or high amounts of enjoyment. This distinction was made to examine the dispositions of students according to their experience with the game as well as apparent interest at the start of play.

4.2 RESULTS

Our findings show that the business negotiations game was a pleasurable experience for new players and experienced players alike, while enjoyment directly related to playful experiences had during play. The elements of arousal, sensation, and thrill correlated the strongest among the four dimensions of the Playful Affordances Model. Furthermore, the correlation of arousal with enjoyment shows that the students had achieved a playful state of mind instead of a serious one. Overall, the activity was rated high for contest, discovery, and arousal, while low ratings were given for imagination and exploration.

Out of the 43 respondents, 7 were discarded due to acquiescence bias on positively and negatively worded behavioral items. The remaining responses were divided into the following subsets: Singaporean students (n = 12); Japanese students (n = 24); cohort 1 as students reporting low anticipated enjoyment (n = 9); cohort 2 as those reporting medium anticipated enjoyment (n = 17); and cohort 3 as those reporting high anticipated enjoyment (n = 10). Students experiencing the game for the first time made up 44% of cohort 1, 35% of cohort 2, and 20% of cohort 3. On average, participants enjoyed the game more than they had anticipated, and the difference was significant (M_1 = 3.08, SD_1 = 0.94; M_2 = 4.11, SD_2 = 1.06; $p < .001$).

Singaporean students reported lower anticipated enjoyment than the Japanese students (M_{SG} = 2.75, SD_{SG} = 0.87; M_{JP} = 3.25, SD_{JP} = 0.94), but higher actual enjoyment (M_{SG} = 4.42, SD_{SG} = 1.00; M_{JP} = 3.96, SD_{JP} = 1.08). Anticipated enjoyment values did not correlate to either overall enjoyment of the game (p = .78) or playful experiences (all eight of which satisfied p > .05).

Table 4-2 Correlations of actual enjoyment to playful experiences. Subsets are Singaporean students (SG), Japanese students (JP), cohort 1 (C1), cohort 2 (C2), cohort 3 (C3), and whole population (P).

	SG $r(10)$	JP $r(22)$	C1 $r(7)$	C2 $r(15)$	C3 $r(8)$	P $r(34)$
Challenge	0.64	0.47*	0.30	0.10	0.17	0.39*
Discovery	0.66*	0.62*	0.72*	0.54*	0.86*	0.64*
Creativity	0.02	0.59*	0.10	0.59*	0.27	0.50*
Arousal	0.69*	0.72*	0.85*	0.24	0.88*	0.70*
Achievement	0.66*	0.59*	0.21	0.06	0.81*	0.60*
Curiosity	0.32	0.57*	0.67*	0.50*	0.67*	0.53*
Fantasy	0.52	0.40	0.64	0.31	0.53	0.44*
Thrill	0.82*	0.66*	0.78*	0.58*	0.79*	0.72*

*p < .05

Reported scores for playful experiences correlated with overall enjoyment of the activity. Correlation values for each of the subsets are shown in Table 4-2 with cohort 3 having the overall highest significant values. Out of the eight experiential items measured, arousal and thrill had the strongest correlation with enjoyment for all students. Similarly, arousal was the strongest predictor of enjoyment for the Japanese students, cohort 1, and cohort 3 while thrill was the strongest for Singaporean students. Cohorts 1 and 3 both showed strong correlations for arousal, thrill, and discovery while cohort 3 also had a strong correlation for achievement. The four significant correlations for cohort 2 were moderate across all items, although their actual enjoyment was the highest among the cohorts (M_{C1} = 4.00, SD_{C1} = 1.32; M_{C2} = 4.24, SD_{C2} = 0.75; M_{C3} = 4.00, SD_{C3} = 1.33). No conclusive data was given for fantasy among any of the subsets, and challenge was a significant indicator only for the Japanese students and population as a whole.

The correlations between behavior-experience and experience-experience pairs within each dimension of the Playful Affordances Model are shown in Table 4-3.

Arousal had the strongest correlations to thrill and sensation across all subsets as well as the whole. Arousal and thrill were very strongly related for students with high expectations of enjoyment (cohort 3). No correlations could be found between exploration and discovery, or imagination and creativity, except for the latter pair which strongly correlated for cohort 3. As for discovery, its relationship to curiosity proved to be strongest among students with high expectations. Challenge and achievement had a high positive correlation for Japanese students and students with low expectations (cohort 1) but a significant negative correlation was found with the Singaporean students.

Table 4-3 Correlations between items within the same dimensions of the Playful Affordances Model. Subsets are Singaporean students (SG), Japanese students (JP), cohort 1 (C1), cohort 2 (C2), cohort 3 (C3), and the whole population (P).

	SG	JP	C1	C2	C3	P
	$r(10)$	$r(22)$	$r(7)$	$r(15)$	$r(8)$	$r(34)$
Challenge/Achievement	-0.58*	0.71*	0.81*	0.47	0.60	0.54*
Challenge/Contest	0.00	0.76*	-0.12	0.51*	0.64*	0.44*
Discovery/Curiosity	0.45	0.57*	0.27	0.61*	0.81*	0.56*
Discovery/Exploration	-0.15	0.10	-0.08	0.42	-0.27	0.12
Creativity/Fantasy	0.40	0.60*	0.15	0.66*	0.45	0.49*
Creativity/Imagination	-0.01	0.22	-0.10	0.11	0.85*	0.27
Arousal/Thrill	0.89*	0.81*	0.73*	0.66*	0.97*	0.81*
Arousal/Sensation	0.67*	0.86*	0.84*	0.75*	0.78*	0.78*

*$p < .05$

Table 4-4 Behavioral items, their correlations, and final ratings

Behavioral Term	Positive Item		Negative Item			
	M	*SD*	*M*	*SD*	$r(34)$	Final Rating
Contest	4.08	0.84	3.77	1.31	0.16	3.77
Sensation	3.57	1.44	3.56	1.21	-0.27	3.56
Exploration	3.47	0.97	3.11	1.30	-0.17	3.11
Imagination	2.94	1.43	3.04	1.34	-0.00	3.04

At the time of development, we suggested that the radial nature of the Playful Affordances Model allows descriptors to be used as dimensions of a radar chart in

the holistic evaluation of playful activities (Songer & Miyata, 2014). In this instance of the business negotiations game, correlation values between behavioral and experiential items did not justify aggregating their ratings into the same dimensions. For this reason, we opted instead to represent the eight primary terms separately. We assumed equal weights on positively worded item scores and their corresponding negatively worded item scores (reversed) to aggregate the means for behavioral items; although, there was no significant correlation between these scores for any of the four behavioral terms (Table 4-4). For the experiential terms, the mean scores of items chosen from the same dimensions of the model were aggregated with equal weights (Table 4-5) and labeled with the primary term. As reported previously in Table 4-3, the scores for each of these descriptor pairs showed significant correlation. The final results reveal that contest, discovery, and arousal were the strongest elements of play while exploration and imagination scored the lowest (Figure 4-1).

Table 4-5 Mean ratings for individual experiential terms and with equal weighting

Experiential Term	*M*	*SD*	Paired Term	*M*	*SD*	Final Rating
Challenge	3.56	1.03	Achievement	3.53	1.21	3.54
Discovery	3.78	1.05	Curiosity	3.75	1.03	3.76
Creativity	3.75	1.08	Fantasy	3.03	1.21	3.39
Arousal	3.81	1.33	Thrill	3.58	1.63	3.69

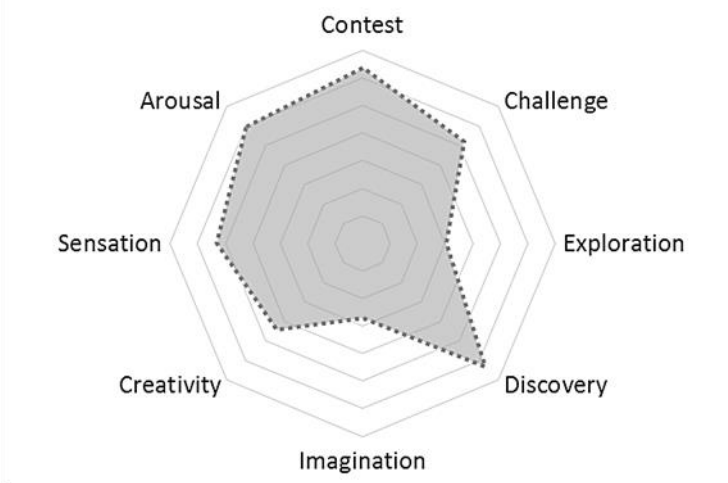

Figure 4-1 Final ratings for playful affordances and experience

4.3 DISCUSSION

The differences between the Singaporean students and Japanese students have implications for the effects of novelty, academic intensity, and different cultures of the participants of this game-based learning activity. The Singaporean students were playing the game for the first time, which may have influenced their higher overall enjoyment despite having relatively lower expectations in the beginning. By comparison, the Japanese students, who were playing for their fourth time, made up the majority of students in cohort 3 with high anticipated enjoyment, indicating that the game activity is a genuinely pleasurable one.

Since the game was carried out in English and relied heavily on communication between participants, we assumed the Japanese students would experience more difficulty compared to the Singaporean students who could speak English fluently. Although the Japanese students reported moderate challenge on average, the data analysis shows that challenge was associated with their overall enjoyment, and strongly related to both the contest of their abilities as well as their sense of achievement. This may reflect levels of self-confidence in their ability to communicate with other players, which was a crucial element in the game's design. Here we should note the meaning of the Japanese word chosen for the translation of "challenge" is associated less with difficulty and more with actively pushing one's limits. In this sense it has a near similar meaning to contest, for which there was a strong correlation with challenge among the Japanese students.

The correlation between arousal, sensation, and thrill indicate that the sensation/arousal dimension of the Playful Affordances Model was the strongest in the design of this activity. This may be attributed to limited resources and time constraints that created a sense of urgency among players. If sensation is interpreted as sheer stimulation in the gaming environment, then its strong correlations to arousal and thrill as well as the positive correlation to enjoyment of all three suggest the students had achieved a playful state of mind, as described in Apter's Reversal Theory (Apter, 2001).

Overall, the data has some implications for the design of game-like activities. First, it is worth pointing out the significant correlations of playful experiences to levels of enjoyment reported by the students. The game was designed with minimal reward

mechanics—sets of cards and a team score. Other mechanics such as time limits, limited resources, and competitive/cooperative dynamics are likely tied to experiences of arousal, thrill, contest, and discovery, which were most prevalent during gameplay. Students who anticipated a moderate amount of enjoyment appeared less sensitive to the effects of playful experiences than those who anticipated either a great amount or a little amount. Regardless, this study demonstrates how students participating in a stimulating game activity can achieve a playful mindset no matter of their degree of anticipated enjoyment.

4.4 LIMITATIONS

In this study, the design of the survey was a limitation with respect to its accordance with the design of the game. The original game designers were not available to comment explicitly their design intentions and how they may relate to elements from the Playful Affordances Model. As such, the questions items used for capturing behavior may not have been accurately paired to playful actions realized by the design of the game. For example, exploration was measured by asking about trial-and-error practices in the development of player tactics, although it might have been more suitable to instead ask about seeking out new people with whom to make deals.

The results relied heavily on player self-report and interpretation of the survey language. The English and Japanese survey items were reviewed by two separate bilingual experts; however, no rigorous or proven translation method was used in the development of the survey. Responses were measured on a Likert-like scale which further relies on individual interpretations of interval levels.

Finally, the sample selection was limited by availability of students in the exchange program as well as those admitted to KTC. The Singaporean students on the exchange program were those that passed a set of criteria for the study abroad program, whereas the Japanese students included everyone within their 4th year of study in the KTC Global Information Technology department. Valid responses were received from 100% of Singaporean participants but only 69% of Japanese participants.

In the future, research that expands on the study presented here should examine both the generalizability of these findings as well as the validity of the Playful

Affordances Model as tool for evaluation. Further experiments with larger sample sizes and different types of game-based learning activities should examine the relationships embedded in the model. Whereas this study found strong correlation for the sensation/arousal dimension, the same may not hold for activities with different designs. Further work should determine if our results are attributed specifically to the design of the chosen game, or if behavioral items in the model do not generally correspond with the given experiences.

5 STUDY 2: SKILL GROWTH

Our second study aims to examine how personal traits relate to learning English at a Japanese school, with possible effects from a game-like skill bars element. We sought to examine motivation and confidence as they relate to positive and negative affect as well as dispositional flow. The individual traits of interest, effort, and instrumental motivation were used to identify motivation, while confidence was represented by self-efficacy and self-concept. Additionally introducing a game-like element allowed us to test the hypothesis that we could use gamification techniques to enhance the learning experience by encouraging flow states.

The timeframe for this study was the spring semester of the 2015 academic calendar at Kanazawa Technical College (KTC). The samples were selected from students enrolled in the Advanced English I: Conversation course and the International Communication II course. The experiments were performed with the expressed consent and cooperation of the English teachers assigned to the courses in the General Studies Department.

5.1 METHOD

Advanced English I: Conversation is an elective course for 4th-year students in the three departments of KTC: Electrical & Electronics Engineering, Mechanical Engineering, and Global Information Technology. The course was designed to be taught by three different teachers in three separate sections. Students were assigned to each section according to their English ability level as observed by their teachers in previous classes. International Communication II is a mandatory course for all students in their 5th year of study in the Global Information Technology department. This course was taught by two different teachers in two separate sections. Students were again assigned according to their observed ability with English. Specifically, the 4th-year students were placed with other students of the same ability level; whereas, the 5th-year students were assigned equally to each section in order to achieve a balance of high and low ability in both sections.

5.1.1 Gamification Add-On

A gamification element was added to both courses by implementing an add-on for an online learning management system (LMS). From the beginning, both courses had planned to use the LMS administered by KTC and run on open-source software called Moodle[3]. The add-on was conceived to be a feedback device that simulates the growth of student skills throughout the course. Borrowing the idea of progress bars representing character skills in video games, the add-on was designed and developed to similarly reflect student progress with a set of skills from the course learning objectives. Key features in addition to the display of progress bar graphics were defined with input from the English teachers as the primary users.

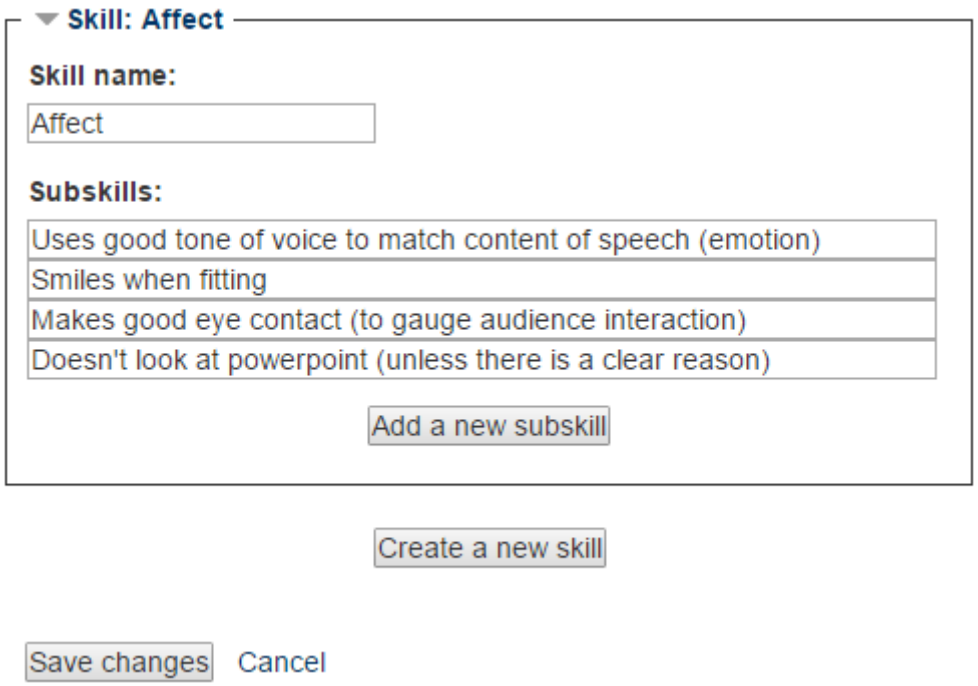

Figure 5-1 Skill Edit screen for teacher users to define skills and subskills

Users of the LMS with teacher privileges in the course could define skills to appear alongside a skill bar in the course add-on block. Once the add-on was added to a course, the teacher would edit the skills for that course using the interface shown in Figure 5-1. Each primary skill was represented by a simple name with associated subskills to provide greater detail. These subskills did not prescribe any specific

[3] https://moodle.org/

mechanics for raising or lowering the skill bar values; however, the teachers in both courses elected to use the subskills as criteria when deciding how to change the primary values for each skill.

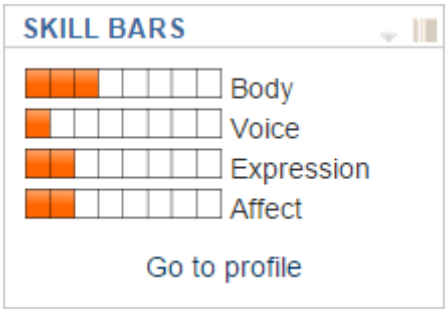

Figure 5-2 Skill bars block shown on the course page for International Communication II

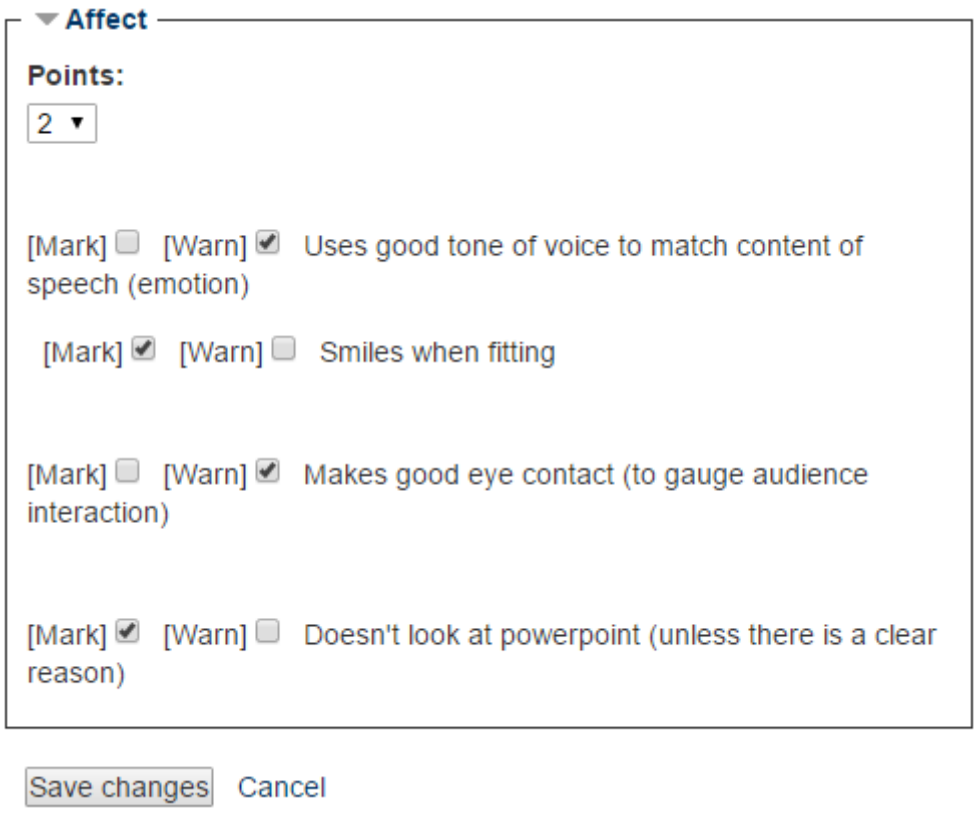

Figure 5-3 Skill Update screen for teacher users to adjust skill points and subskill marks

After logging in for the first time to a course with the add-on, student users initiated a profile. This one-time operation required the students to select one strong skill and one weak skill. The selection determined initial scores for each of the student skills.

For example, the student whose bars are shown in Figure 5-2 chose Body as a strong skill and Voice as a weak skill.

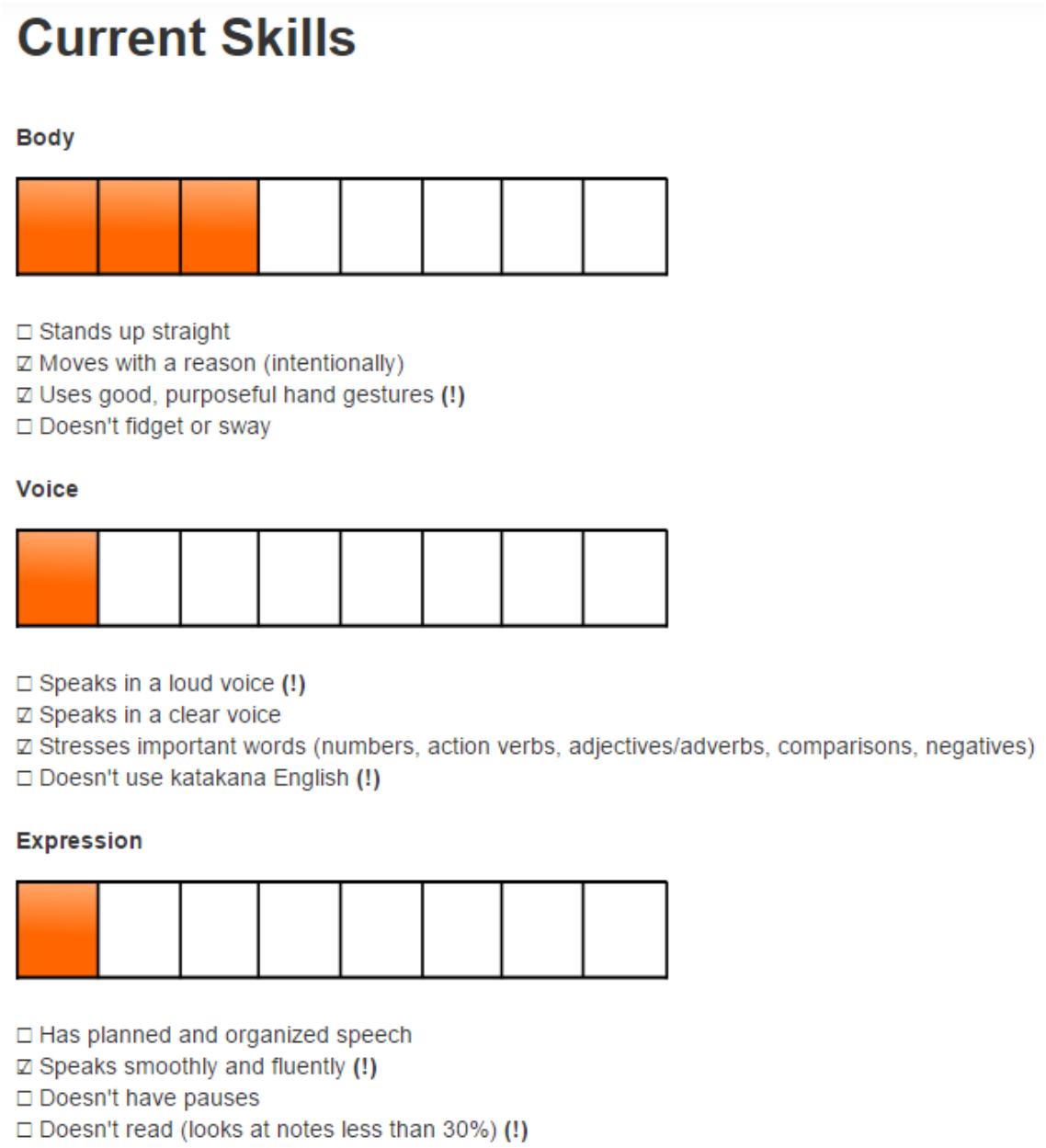

Figure 5-4 Student Profile screen shows skill bars and subskill details

After an activity had been performed and the teacher was ready to give feedback to a student, the skill bars would then be updated through the interface shown in Figure 5-3. Teachers had direct control over the skill bar value as well as the completion/warning marks of each subskill. Students could then view their skill bars

and subskills in greater detail on the profile screen, as shown in Figure 5-4. Subskills that had been achieved were represented by a checked box while those that needed special attention (i.e. a warning) were denoted with an exclamation mark (!).

5.1.2 Activity Cycles

With the introduction of the skill bars add-on, it was necessary for the teachers of each section to administer activity cycles that incorporate feedback through the tool. Using the selected skill set as a target, the teachers would direct learning activities with the students, assess the results, update the skill bars, and then provide feedback to the students. In Advanced English I: Conversation, the skill set related to English speaking ability and each cycle involved a voice recording assignment. The first assignment was used as a baseline for student ability to be assessed against in future assignments. The low ability group then proceeded to complete the feedback loop three times while the high ability group completed it four times. Students were directed to check their updated skill bars on the LMS after the second activity, but each proceeding cycle used printouts of student skill bars. The teachers did not have planned activities on the computers for every class, so the printouts were provided to ensure that the students received their feedback. However, these printouts showed results for all students with skill bars whereas the online interface only revealed skill bars for the individual.

In the International Communication II course, activity cycles revolved around English presentation assignments. The teacher would assess each student following a rubric of criteria that formed the set of subskills for each skill bar. In activities where students were paired, the teacher chose each partner randomly with no consideration for student ability. Each time a student received full marks for a set of subskills, the corresponding skill bar would increase a single point. The first feedback was given as a graded rubric of all the subskills. Each of the following activity cycles were then completed with feedback given as a printout of all student skill bars similar to the one for Advanced English I: Conversation above. A total of four activity cycles were completed in this way.

5.1.3 Evaluation

The division of students into sections within each course enabled us to test for effects on individual traits related to the gamification condition and controlled by English

ability. In the 4th-year class, students in the low ability section and the high ability section used the skill bars mechanic while those in the medium ability section did not. The 5th-year class was divided evenly by ability into two sections, so one became the control group while the other used the skill bars. Although each section had different teachers, the separate sections all followed the same schedule and course materials.

Individual traits of the students were measured with an electronic survey presented at the beginning of the spring semester, and once more at the end, about ten weeks later. The survey was administered on the LMS as a series of multiple choice Likert-like items. The pretest survey included 25 items: three items each for interest in English, effort and persistence, self-efficacy, instrumental motivation, and self-concept; five positive affect terms; and five negative aspect terms. The posttest survey included 18 items: three items each for interest in English, effort and persistence, and self-efficacy; and nine items for the aspects of dispositional flow. Each survey was voluntary, and the students were presented at the start of each one with a statement of consent for allowing their results to be used anonymously in educational research.

The survey items regarding motivation and confidence were adapted from the Organization for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA) Student Approaches to Learning inventory (Artelt, Baumert, Julius-Mc-Elvany, & Peschar, 2003). Whereas the original items were written with regards to the main topics of the PISA—mathematics, science, and reading (of a native language)—our items were written to pertain specifically to English. The items for instrumental motivation and self-concept were excluded from the posttest due to their long-term nature as being not likely to change dramatically over a 10-week period.

Positive and negative affect was tested using the short-form of the International Positive and Negative Affect Schedule (PANAS) (Thompson, 2007). This list of ten terms was presented in English and empirically tested for consistency in responses from international university students. In the pretest survey, these terms were included alongside Japanese translations in order to assist interpretation by students with lower English ability. Positive affect terms included alert, inspired,

determined, attentive, and active; while negative affect terms included upset, hostile, ashamed, nervous, and afraid.

Dispositional flow was tested with items from the Short Dispositional Flow Scale-2 (DFS-2) (Jackson, Martin, & Eklund, 2008) and its Japanese translation which has been validated with Japanese adults (Kawabata, Mallett, & Jackson, 2008). The original short DFS-2 includes a single item for each of the nine aspects of flow mentioned in Section 2.4 above. The original items were written to measure a respondent's tendency to experience flow during physical activities such as exercising. These items were then modified to relate instead to participation in the English classes.

5.1.4 Data Analysis

Survey results were analyzed for correlation and effect size. Pearson's correlation was used to compare the dependent variables of interest, effort, self-efficacy, and dispositional flow to the independent variables of instrumental motivation, self-concept, positive and negative affect, and the skill bars condition. For effect size, Cohen's *d* was chosen as a common metric for comparing difference scores of samples with inconsistent sizes and variance (Borenstein, Hedges, Higgins, & Rothstein, 2009). Described simply as standardized mean difference, Cohen's *d* is the difference of means between two sets divided by the pooled standard deviation of both sets (Equation 5-1). For paired sets (such as interest, effort, and self-efficacy on the pretest and posttest surveys) the pooled standard deviation was calculated by taking the correlation score of the two sets into account (Equation 5-2). While the survey response scales remained consistent as 5-point scales ranging from -2 to 2, we also present the raw mean difference *D* where applicable.

$$d = \frac{M_1 - M_2}{SD_{pool}} \qquad \text{(Equation 5-1)}$$

$$SD_{within} = \frac{SD_{diff}}{\sqrt{2(1-r)}} \qquad \text{(Equation 5-2)}$$

5.2 RESULTS

Survey responses were analyzed as separate samples for the International Communication II course ($n = 32$) and the Advanced English I: Conversation course

(n = 41). Each course sample was further divided into the samples of the control group (n = 17 for International Communication II; n = 12 for Advanced English I: Conversation) and the group that used the skill bars add-on. In the Advanced English I: Conversation course, the skill bars sample was further divided according to the sections of high ability students (n = 19) and low ability students (n = 10).

In the International Communication II course, the control group showed relationships for individual self-concept and instrumental motivation with initial interest, effort, and self-efficacy. Gains in motivation and confidence may be attributed to dispositional flow which correlated strongly with self-efficacy for the whole of the students. The skill bars add-on did not appear to have an effect, although the students using it did report lower positive affect, higher negative affect, and lower frequency of flow experiences.

In the Advanced English I: Conversation course, instrumental motivation scores correlated with initial interest for the students with low and intermediate ability while self-concept correlated with interest for the low and high ability students. The control group showed a gain in interest which is unexplainable in terms of the data we collected, but a drop in self-efficacy appears to be connected to low positive affect, a minor drop in effort, and infrequent flow experiences during the course. Students of high and low ability using the skill bars reported significantly higher scores for clear goals and unambiguous feedback, showing a potential effect from the use of the gamification add-on to give clear, direct reports on student performance.

5.2.1 International Communication II

Student scores in International Communication II showed mostly positive changes for interest, effort, and self-efficacy. As shown in Figure 5-5, the larger differences were in the interest scores of the skill bars group (d = 0.48) and the effort scores of the control group (d = 0.51). Pretest and posttest scores along with their correlations in the control group are presented in Table 5-1 and Table 5-2 respectively.

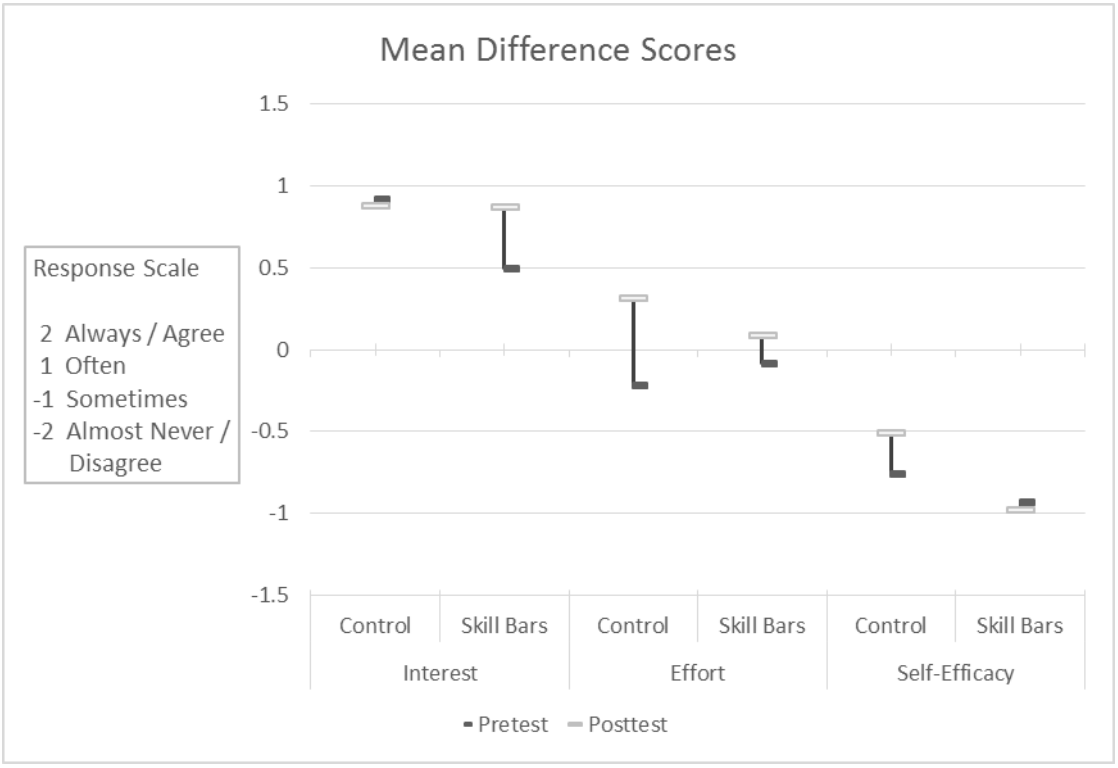

Figure 5-5 Mean difference scores for control group and skill bars group

5.2.1.1 Group Scores

In the pretest survey, the scores of the control group showed self-concept scores to have strong relationships with the scores for self-efficacy and instrumental motivation. Self-efficacy and effort & persistence also correlated, while the two of them had significant relationships to instrumental motivation scores. Additionally, student effort & persistence scores appeared to be connected to their interest in English.

Table 5-1 Control group pretest mean scores and correlation values

Pretest (Control)			$r(15)$			
	M	*SD*	**IN**	**EP**	**SE**	**IM**
Interest (**IN**)	0.92	0.87				
Effort & Persistence (**EP**)	-0.22	1.03	0.50*			
Self-Efficacy (**SE**)	-0.76	0.91	0.41	0.61*		
Instrumental Motivation (**IM**)	-0.47	1.32	0.38	0.51*	0.49*	
Self-Concept (**SC**)	0.02	1.48	0.47	0.40	0.70*	0.72*

*$p < .05$

On the posttest survey, the correlation of effort & persistence with interest remained significant. Likewise, the correlation of effort & persistence with self-efficacy remained significant but with smaller values. The relationship between interest and self-efficacy was insignificant just as it was in the pretest although the correlation value dropped.

Table 5-2 Control group posttest mean scores and correlation values

Posttest (Control)			r(15)		
	M	*SD*	**IN**	**EP**	**SE**
Interest (**IN**)	0.88	0.82			
Effort & Persistence (**EP**)	0.31	1.05	0.70*		
Self-Efficacy (**SE**)	-0.51	1.00	0.27	0.56*	
Dispositional Flow (**DF**)	-0.01	0.80	0.38	0.50*	0.80*

*$p < .05$

While the instrumental motivation scores of the control group were more broadly related to the items for effort and self-efficacy, it was self-concept for the skill bars group that appeared to have a stronger relationship. On the pretest survey, scores in the skill bars group showed significant correlations to self-concept for interest, effort & persistence, and self-efficacy. Instrumental motivation had a moderate correlation value with self-efficacy but the values for interest and effort & persistence were low. Furthermore, interest showed a strong relationship with effort & persistence which in turn showed a strong relationship with self-efficacy. However, these correlations were weaker in the posttest as only the relationship of interest with effort & persistence remained significant.

Table 5-3 Skill bars group pretest mean scores and correlation values

Pretest (Skill Bars)			r(13)			
	M	*SD*	**IN**	**EP**	**SE**	**IM**
Interest (**IN**)	0.49	0.70				
Effort & Persistence (**EP**)	-0.09	1.14	0.61*			
Self-Efficacy (**SE**)	-0.93	0.79	0.51	0.60*		
Instrumental Motivation (**IM**)	-1.24	0.99	0.29	0.23	0.53*	
Self-Concept (**SC**)	0.42	1.19	0.62*	0.60*	0.58*	0.31

*$p < .05$

Table 5-4 Skill bars group posttest mean scores and correlation values

Posttest (Skill Bars)			$r(13)$		
	M	*SD*	**IN**	**EP**	**SE**
Interest (**IN**)	0.87	0.85			
Effort & Persistence (**EP**)	0.09	0.77	0.56*		
Self-Efficacy (**SE**)	-0.98	0.74	0.18	0.43	
Dispositional Flow (**DF**)	-0.61	0.61	0.55*	0.39	0.54*

*$p < .05$

Comparing these relationships with the pre-post difference scores in both groups, we see possible explanations for the different results between samples. The control group reported more frequent flow experiences on the posttest, which showed a moderate correlation to effort & persistence and a strong relationship to self-efficacy. In the pretest, effort & persistence and self-efficacy related to instrumental motivation which was higher for the control group than the skill bars group ($d = 0.67$). The changes both effort & persistence and self-efficacy for the control group were both positive. Meanwhile, the posttest scores of flow experience in the skill bars group showed moderate correlation to self-efficacy and interest, both of which showed correlation to self-concept in the pretest. Self-efficacy scores also correlated moderately with instrumental motivation. The skill bars group had higher self-concept scores ($d = 0.45$) than the control group, but considerably lower scores for instrumental motivation. The results for the skill bars group show a significant growth in interest but lower frequency of flow. Overall, the noticeably larger scores for dispositional flow in the control group ($d = 0.77$) may be connected to the combination of a larger score for the related individual trait in the pretest (i.e., instrumental motivation vs. self-concept) and a larger difference score in the related outcomes over the course of the experiment. The scores of both groups for these individual traits and flow are presented in Figure 5-6.

The scores for positive and negative affect were also analyzed in the pretest. The skill bars group had slightly lower positive affect ($d = -0.30$) and considerably higher negative affect ($d = 0.50$). The affect scores were largely unrelated to all other items in the pretest, except for positive affect which correlated with effort & persistence in the skill bars group ($r(13) = 0.61$, $p = .02$). Furthermore, positive and negative affect

scores in the pretest appeared to be unrelated to scores for dispositional flow in the posttest.

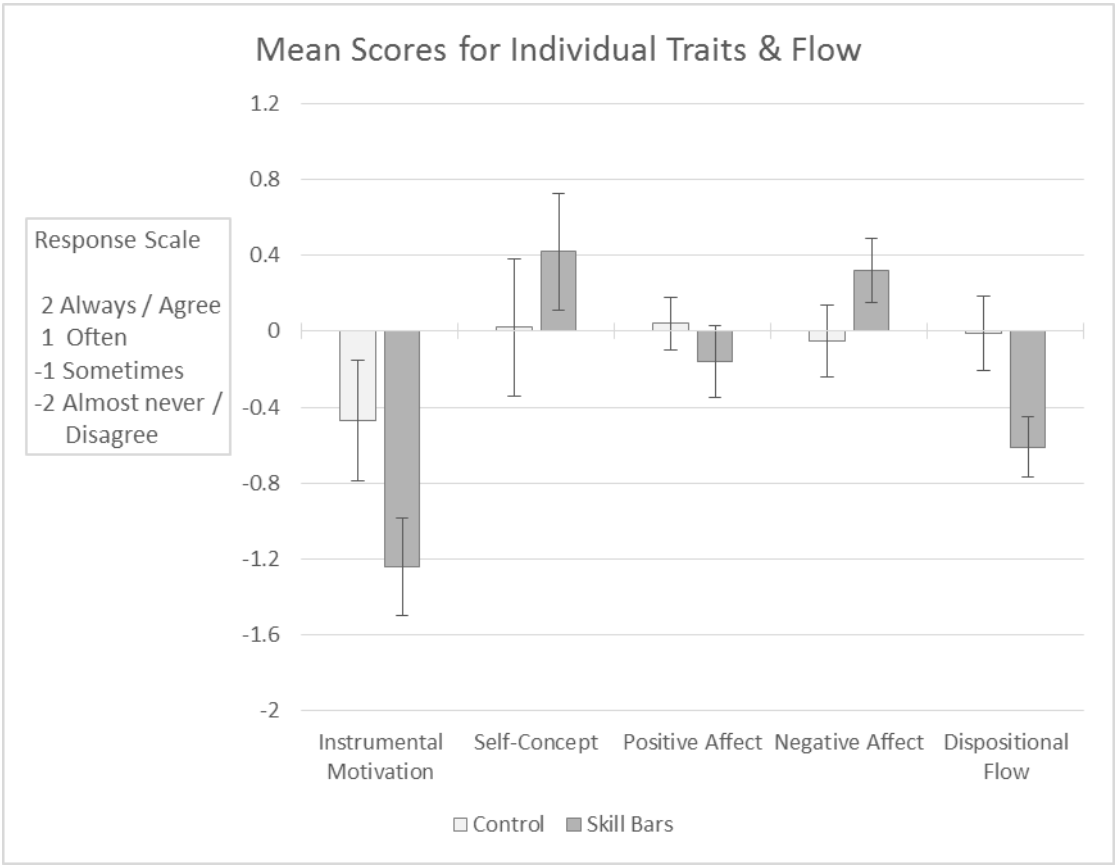

Figure 5-6 Sample scores for individual traits in the pretest and dispositional flow in the posttest

A look at the individual items for dispositional flow shows additional characteristics of the students' experiences in the separate sections of the course (Figure 5-7). The students as a whole had similar experiences concerning action-awareness merging, total concentration, and a sense of control. The skill bars group reported substantially lower scores for clear goals (d = -0.80), transformation of time (d = -0.81), and autotelic experience (d = -0.74). Large differences can also be seen for the items of challenge-skill balance (d = -0.60) and loss of self-consciousness (d = -0.68). This shows that the skill bars group had more difficulty approaching the challenges of the public speaking assignments, tended to be more self-conscious in class, and felt that the course was less enjoyable than the control group. Nevertheless, students in the skill bars group that experienced greater degrees of flow also reported higher scores for interest in the posttest ($r(13) = 0.53$, $p = .04$).

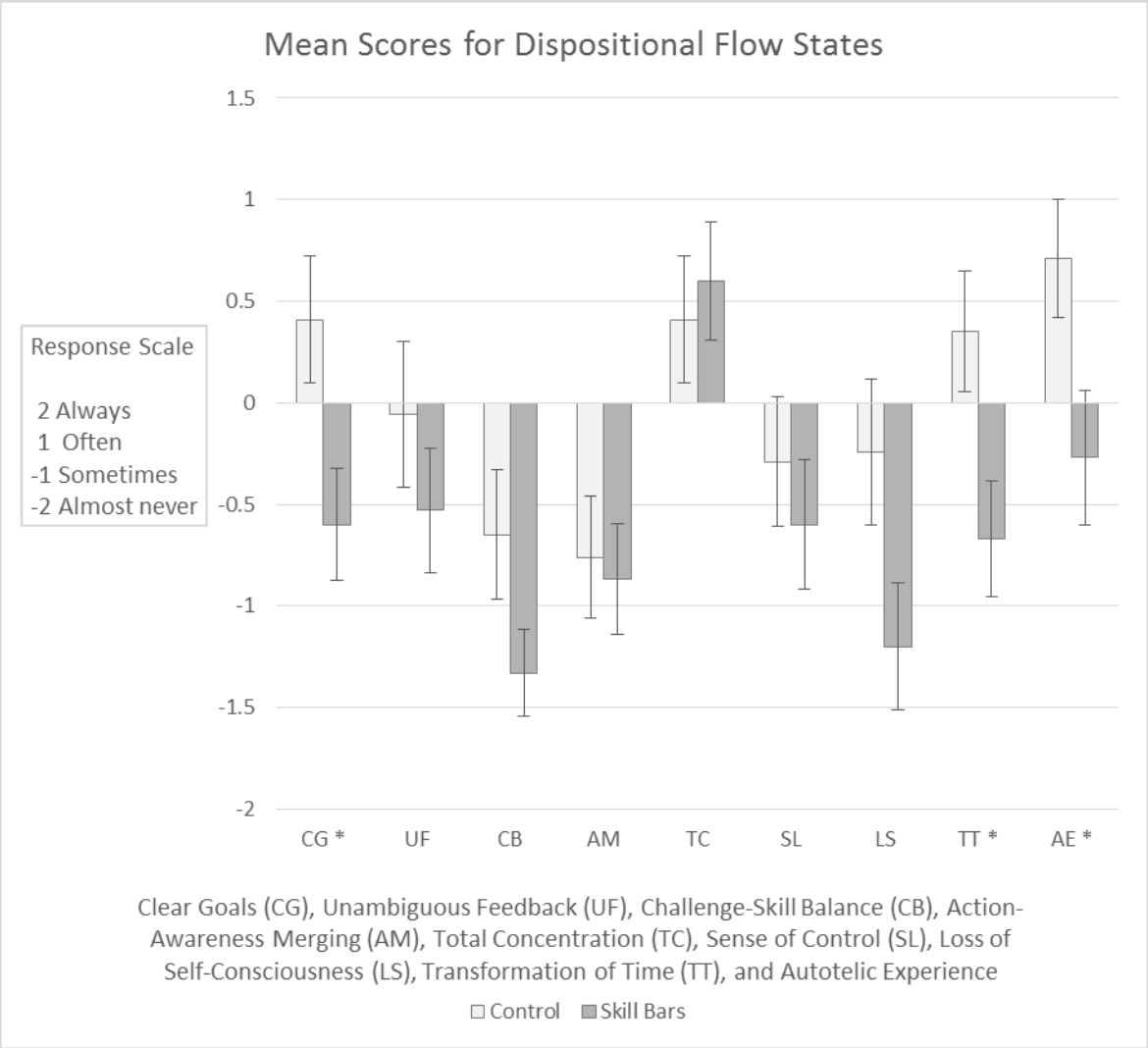

Figure 5-7 Scores for separate flow states; (*) indicates statistically significant difference between control and skill bars groups as determined by a two-tailed Student's t-Test ($p < .05$)

Overall, both samples showed gains in student interest, effort & persistence, and self-efficacy. The control group began the course with higher instrumental motivation scores than the skill bars group, experienced higher gains in effort & persistence and self-efficacy, and reported more frequent flow experiences. Higher scores for clear goals, challenge-skill balance, and autotelic experience for the control group coincide with higher ratings in loss of self-consciousness and transformation of time. Meanwhile, the flow item that most closely relates to the design of the skill bars add-on, unambiguous feedback, showed a generally lower effect in the group that used the skill bars during class ($d = -0.35$).

5.2.1.2 Whole Population Scores

In terms of the whole population, Pearson correlation values for difference scores in interest, effort & persistence, and self-efficacy showed no significant correlations with any of the independent variables of instrumental motivation, self-concept, positive affect and negative affect. Nor did any of the difference scores appear to relate directly to frequency of flow experiences. However, when considering the possibility of non-linear correlation, we discovered a large difference between the dispositional flow scores of students who reported a loss in self-efficacy and those who reported a gain. This was found by dividing the students into cohorts according to individual self-efficacy difference scores. Cohorts were formed based on distance from the mean difference score of the whole. Students with a large drop in self-efficacy during the course (beyond one standard deviation below the mean, $D < -0.69$) made up cohort 1 while students with large gains in self-efficacy (beyond one standard deviation above the mean, $D > 0.92$) made up cohort 4. Cohorts 2 and 3 consisted of the students within one standard deviation below and above the mean ($M = 0.11$), respectively. Then, comparing the mean scores of dispositional flow for each cohort showed that greater frequencies of flow experiences were associated with gains in self-efficacy ($M_{C1} = -0.69$, $SD_{C1} = 0.60$; $M_{C2} = -0.33$, $SD_{C2} = 0.89$; $M_{C3} = -0.39$, $SD_{C3} = 0.69$; $M_{C4} = 0.36$, $SD_{C4} = 0.49$). The difference in flow scores between cohort 1 and cohort 4 is $D = 1.05$ on our 5-point response scale, which translates into a standardized mean difference of $d = 2.29$. Furthermore, a two-tailed Student's t-Test with unequal variance showed that the flow scores of cohort 1 and cohort 4 were statistically significant ($p = .03$).

Similarly, negative affect scores were lower in the cohort with higher gains in self-efficacy. Cohort 1 had reported the highest mean negative affect score in the pretest while the mean of each successive cohort was smaller than the one before ($M_{C1} = 0.40$, $SD_{C1} = 0.43$; $M_{C2} = 0.23$, $SD_{C2} = 0.75$; $M_{C3} = 0.15$, $SD_{C3} = 0.78$; $M_{C4} = -0.40$, $SD_{C4} = 0.73$). The difference in negative affect scores between the low and high cohorts then is $D = -0.80$, or $d = -1.28$; although, these samples were not statistically significant ($p = .08$).

Following this same method of analysis, our comparison of the other difference scores to the independent variables scores and dispositional flow did not produce any other noteworthy results.

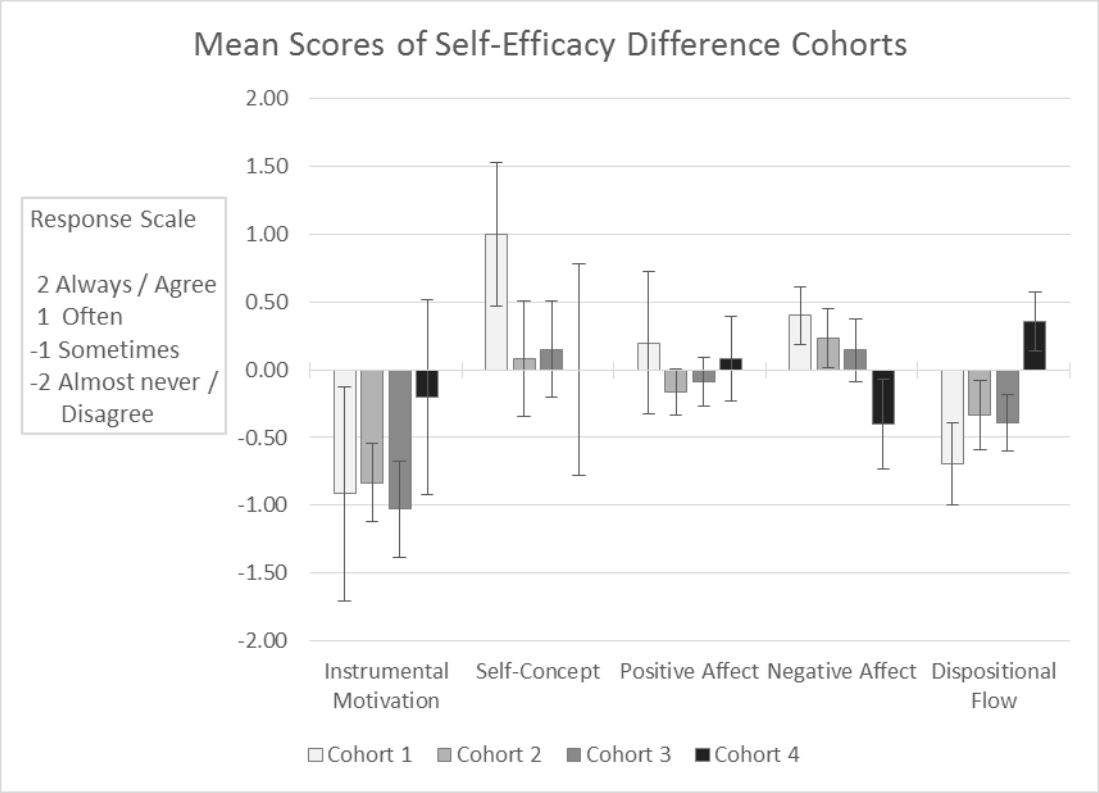

Figure 5-8 Mean scores of individual traits for cohorts determined by self-efficacy difference scores

5.2.2 Advanced English I: Conversation

In the Advanced English I: Conversation course, the differences in pretest and posttest scores show both gains and losses for the control group, the low ability group and the high ability group (Figure 5-9). The control group (consisting of students with intermediate language ability) showed a growth in interest ($d = 0.42$), but also a significant drop in self-efficacy ($d = -0.57$). The skill bars group was divided into a high ability group and a low ability group. While the high ability group saw an increase in self-efficacy ($d = 0.49$), the changes in the low ability group were relatively small. The mean scores and correlations values of items on the pretest and posttest are presented in tables below.

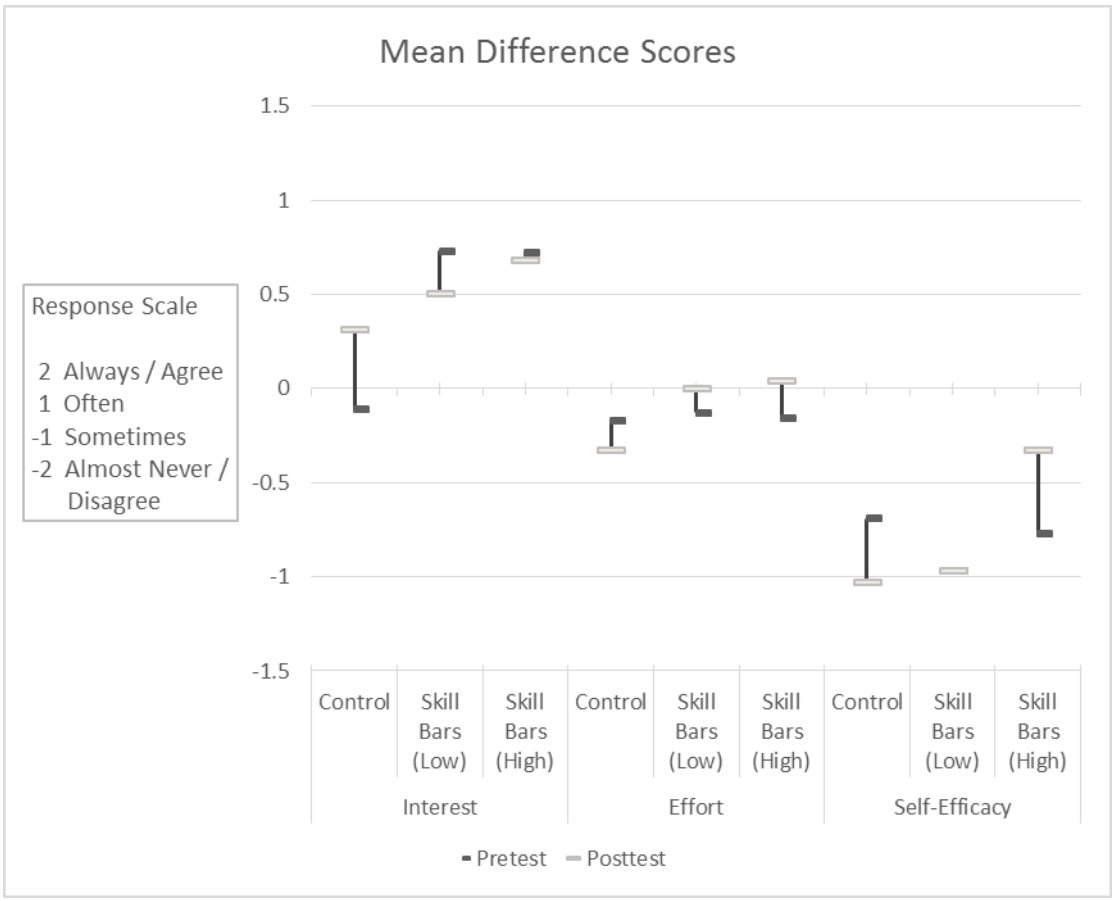

Figure 5-9 Mean difference scores for control group and skill bars groups of low and high ability

5.2.2.1 Group Scores

In the pretest survey (Table 5-5), the control group showed strong relationships for instrumental motivation with both interest and effort. Interest had a moderate correlation with effort & persistence while self-efficacy correlated to self-concept. Table 5-6 shows that these correlation values did not persist. Posttest scores for effort & persistence showed a stronger relationship to self-efficacy than interest. The correlation values for self-efficacy and interest were insignificant in the posttest, which remained consistent from the pretest despite a slightly higher value.

The low ability group began the course with the lowest self-efficacy scores out of all three groups. However, self-efficacy showed no significant correlation to any other item on the pretest (Table 5-7). Interest and instrumental motivation showed a strong correlation while both items additionally showed strong correlation with self-concept. Meanwhile, effort & persistence correlated with interest and self-concept to nearly equal degrees.

Table 5-5 Control group pretest mean scores and correlation values

Pretest (Control)			r(10)			
	M	*SD*	**IN**	**EP**	**SE**	**IM**
Interest (**IN**)	-0.11	1.00				
Effort & Persistence (**EP**)	-0.17	0.99	0.59*			
Self-Efficacy (**SE**)	-0.69	0.69	0.04	0.42		
Instrumental Motivation (**IM**)	-0.53	1.25	0.73*	0.76*	0.36	
Self-Concept (**SC**)	-0.33	0.91	0.03	0.20	0.64*	0.05

*$p < .05$

Table 5-6 Control group posttest mean scores and correlation values

Posttest (Control)			r(10)		
	M	*SD*	**IN**	**EP**	**SE**
Interest (**IN**)	0.31	1.01			
Effort & Persistence (**EP**)	-0.33	0.70	0.46		
Self-Efficacy (**SE**)	-1.03	0.41	0.22	0.71*	
Dispositional Flow (**DF**)	-0.60	0.49	0.41	0.56	0.69*

*$p < .05$

Table 5-7 Skill bars (low ability) group pretest mean scores and correlation values

Pretest (Low Ability)			r(8)			
	M	*SD*	**IN**	**EP**	**SE**	**IM**
Interest (**IN**)	0.73	0.97				
Effort & Persistence (**EP**)	-0.13	1.15	0.66*			
Self-Efficacy (**SE**)	-0.97	0.82	0.03	0.12		
Instrumental Motivation (**IM**)	-0.43	1.05	0.71*	0.38	0.20	
Self-Concept (**SC**)	-0.70	1.21	0.84*	0.65*	0.46	0.78*

*$p < .05$

In the posttest, the scores of the low ability group showed that the correlation between effort & persistence and interest remained strong (Table 5-8). The correlation value between self-efficacy and effort & persistence became high, although not significant due to the small sample size.

Table 5-8 Skill bars (low ability) group posttest mean scores and correlation values

Posttest (Low Ability)			*r*(8)		
	M	*SD*	**IN**	**EP**	**SE**
Interest (**IN**)	0.50	1.17			
Effort & Persistence (**EP**)	0.00	0.97	0.86*		
Self-Efficacy (**SE**)	-0.97	0.51	0.45	0.60	
Dispositional Flow (**DF**)	-0.49	0.55	0.71*	0.66*	0.37

*$p < .05$

Table 5-9 Skill bars (high ability) group pretest mean scores and correlation values

Pretest (High Ability)			*r*(17)			
	M	*SD*	**IN**	**EP**	**SE**	**IM**
Interest (**IN**)	0.72	0.72				
Effort & Persistence (**EP**)	-0.16	1.05	0.42			
Self-Efficacy (**SE**)	-0.77	0.80	0.21	0.51*		
Instrumental Motivation (**IM**)	-0.44	0.98	0.23	0.59*	0.46*	
Self-Concept (**SC**)	0.30	0.97	0.66*	0.29	0.42	0.22

*$p < .05$

Table 5-10 Skill bars (high ability) group posttest mean scores and correlation values

Posttest (High Ability)			*r*(17)		
	M	*SD*	**IN**	**EP**	**SE**
Interest (**IN**)	0.68	0.67			
Effort & Persistence (**EP**)	0.04	0.92	0.69*		
Self-Efficacy (**SE**)	-0.33	0.97	0.53*	0.71*	
Dispositional Flow (**DF**)	-0.33	0.53	0.59*	0.62*	0.68*

*$p < .05$

Compared to the other two groups, the high ability group also reported similarly low self-efficacy but higher self-concept on the pretest. Self-concept scores correlated with interest while self-efficacy scores showed a correlation with effort & persistence (Table 5-9). Instrumental motivation also had a moderate correlation with both self-efficacy and effort & persistence. In the posttest survey, all correlation values became stronger than in the posttest (Table 5-10). The correlation of self-efficacy with effort & persistence remained the strongest while the relationship between self-

efficacy and interest became significant. Interest scores also showed a strong correlation with effort & persistence.

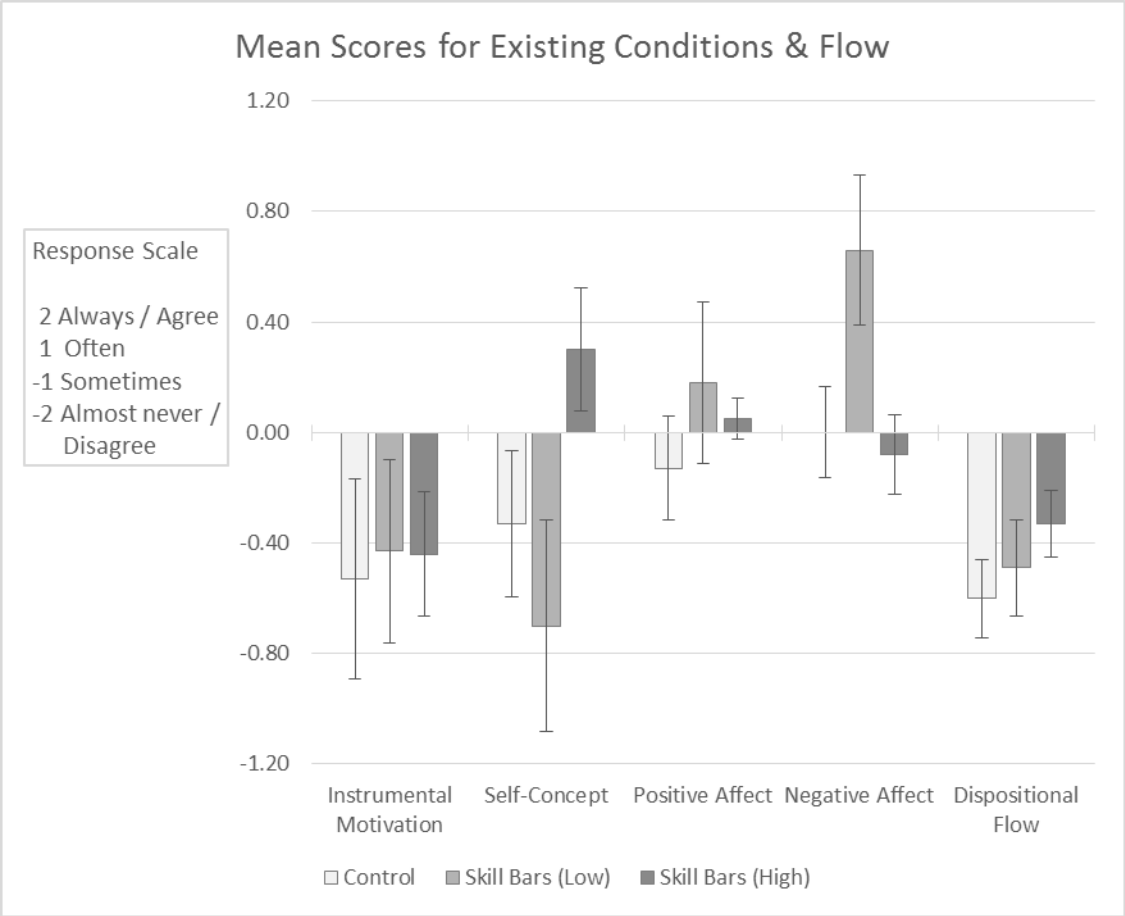

Figure 5-10 Sample scores for individual traits in the pretest and dispositional flow in the posttest

Examining the relationships between items for each group may give some insight into the reasons for the observed pre-post difference scores. Overall, the control group change in interest was significant compared to the combined skill bars group ($d = 0.56$); however, interest scores did not appear to relate to any of the measured variables. Although there was a correlation with instrumental motivation on the pretest, the scores for instrumental motivation were equally low among all three groups, as can be seen in Figure 5-10. So changes in interest scores for the control group are likely influenced by some external factor. Compared to the combined skill bars group, the control group also saw a significant change in self-efficacy ($d = -0.65$). Self-efficacy correlated with self-concept in the pretest, whereas in the posttest it correlated strongly with effort & persistence as well as dispositional flow. So, the

change in self-efficacy for the control group appeared to be connected to a low self-concept at the start in addition to a slight drop in effort & persistence and infrequent flow experiences through the course.

Within the skill bars groups, there were only slight changes to student motivation and confidence except in the high ability group which saw a significant gain in self-efficacy. Their self-efficacy scores showed a moderate correlation with instrumental motivation in the pretest and a strong correlation to dispositional flow in the posttest. As shown in Figure 5-10 above, flow experiences were generally higher for the skill bars groups compared to the control group ($d = 0.42$), with the high ability group reporting higher scores than the low ability group ($d = 0.30$). Furthermore, posttest scores of self-efficacy in the high ability group correlated strongly with dispositional flow. This suggests that flow experiences may have been a factor for change in the self-efficacy scores of the high ability group.

Positive and negative affect scores in the pretest revealed further characteristics of each student group. The low ability group had the highest negative affect scores and showed significant correlation for negative affect with instrumental motivation ($r(10) = 0.68$, $p = .03$). The control group and the high ability group both showed correlation for positive affect with self-efficacy ($r(12) = 0.60$, $p = .04$; and $r(19) = 0.52$, $p = .02$). On the posttest, scores of dispositional flow correlated strongly with pretest scores of positive affect for the high ability students ($r(19) = 0.65$, $p = .003$), which is attributed mainly to the scores of the flow items for loss of self-consciousness ($r(19) = 0.72$, $p = .001$) and autotelic experience ($r(19) = 0.62$, $p = .005$). These same two items also showed a moderate inverse relationship with negative affect for the high ability group, where loss of self-consciousness was $r(19) = -0.49$ ($p = .03$) and autotelic experience was $r(19) = -0.49$ ($p = .03$). This suggests that the initial motivation of the low ability students is tied to pressure from a perceived obligation to learning English. Meanwhile, the confidence of the intermediate and high ability students is connected to pleasurable experiences had when learning and practicing their skills.

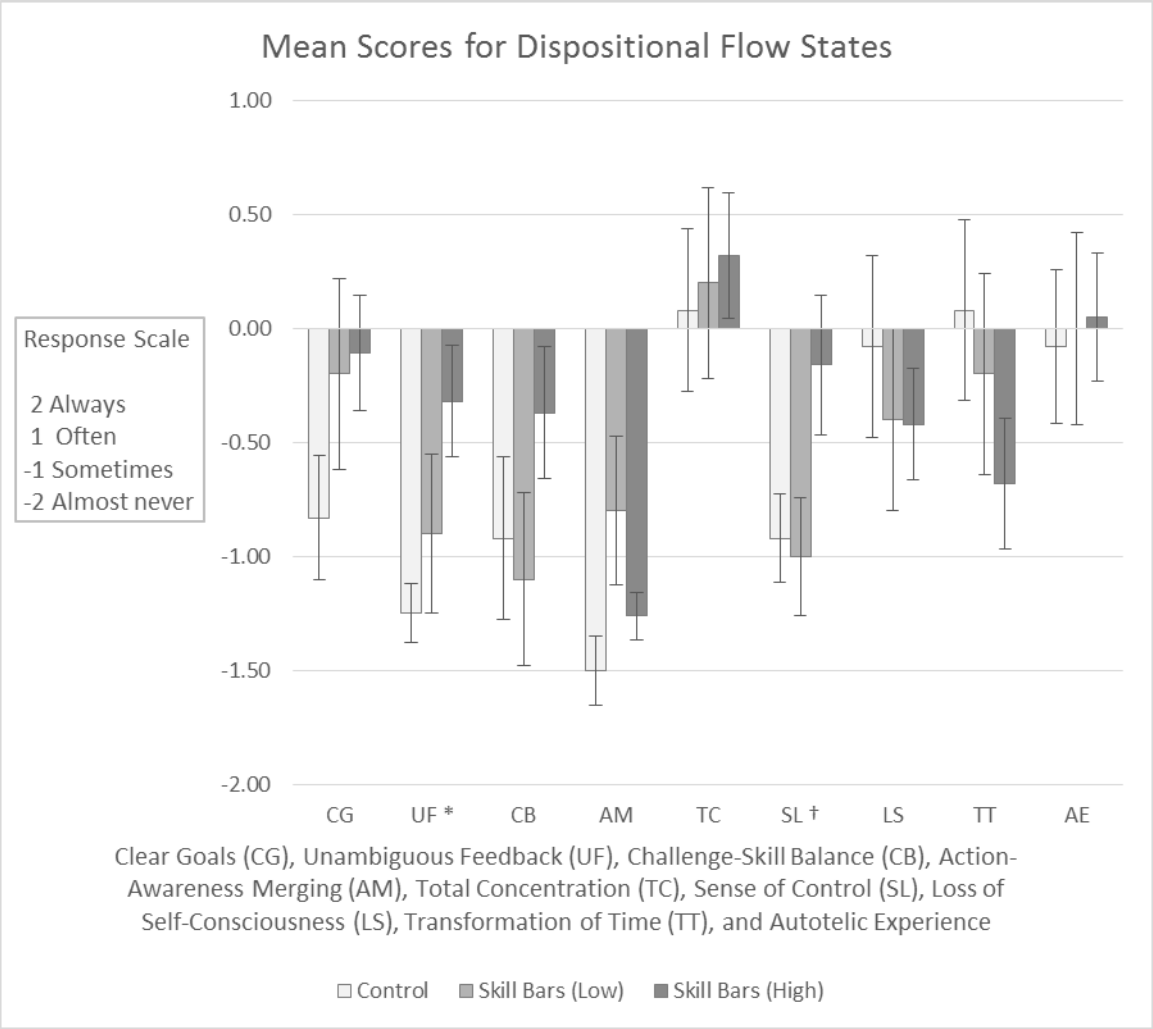

Figure 5-11 Scores for separate flow states; statistically significant differences are indicated by (*) for control vs. skill bars and (†) for low ability vs. high ability, as determined by a two-tailed Student's t-Test ($p < .05$)

Sub-items for dispositional flow revealed further differences in the students' experiences (Figure 5-11). All three groups reported similar scores for total concentration and autotelic experience. The skill bars groups combined showed significantly higher scores for clear goals ($d = 0.61$), unambiguous feedback ($d = 0.73$), action-awareness merging ($d = 0.58$), and sense of control ($d = 0.42$), but lower scores for transformation of time ($d = -0.45$). Compared to the low ability group, scores for the high ability group showed significantly higher scores for unambiguous feedback ($d = 0.54$), challenge-skill balance ($d = 0.58$), and sense of control ($d = 0.68$), but significantly lower scores for action-awareness merging ($d = -0.64$). Very little

difference was observed between high and low ability groups for the items of clear goals, total concentration, loss of self-consciousness, and autotelic experience. Given that both the low ability group and the high ability group had similar scores for clear goals while the scores of the control group were significantly lower, this is strong evidence that the skill bars add-on was effective in providing students with distinct goals for improving their English ability. Furthermore, a two-tailed, unequal variance Student's t-Test comparing the unambiguous feedback scores for the control group and the combined skill bars group shows a statistically significant difference ($p = .004$), suggesting the add-on also had a positive effect on student ability to interpret feedback. Since the scores of the low ability group were significantly lower than those of the high ability group, this effect may have been weaker for students with low ability. Challenge-skill balance and sense of control had large differences between the ability groups and so were likely influenced more by English ability than other factors. Interestingly though, a large negative difference for action-awareness merging in the high ability students suggests that the more accomplished students were less spontaneous with their English use and had to think more about it in the process.

5.2.2.2 Whole Population Scores

The difference scores for the shared pre-post survey items show how individual traits and flow experiences were related to changes in motivation and confidence. Pearson correlation values showed relationships for the difference scores of effort & persistence with both instrumental motivation ($r(41) = -0.42$, $p = .006$) and dispositional flow ($r(10) = 0.41$, $p = .007$). Using the same cohort analysis method explained previously, we divided the whole based on individual difference scores of effort & persistence. Cohort 1 consisted of students beyond one standard deviation below the mean ($D < -1.03$), cohort 2 and cohort 3 consisted of students within one standard deviation below and above the mean ($M = 0.07$), and cohort 4 consisted of everyone beyond one standard deviation above the mean ($D > 1.18$). Their mean scores of the cohorts for instrumental motivation, self-concept, affect, and flow are shown in Figure 5-12. In addition to the Pearson correlation being quite apparent, the figure reveals a large difference in scores between cohorts for both instrumental motivation and dispositional flow. For instrumental motivation, the mean difference between the scores of cohort 1 and cohort 4 was $D = -1.40$ ($d = -1.76$). For

dispositional flow, this was $D = 0.76$ ($d = 1.67$). Additionally, a two-tailed Student's t-Test with unequal variance showed that the differences between cohort 1 and cohort 4 were statistically significant ($p = .01$ and $p = .01$).

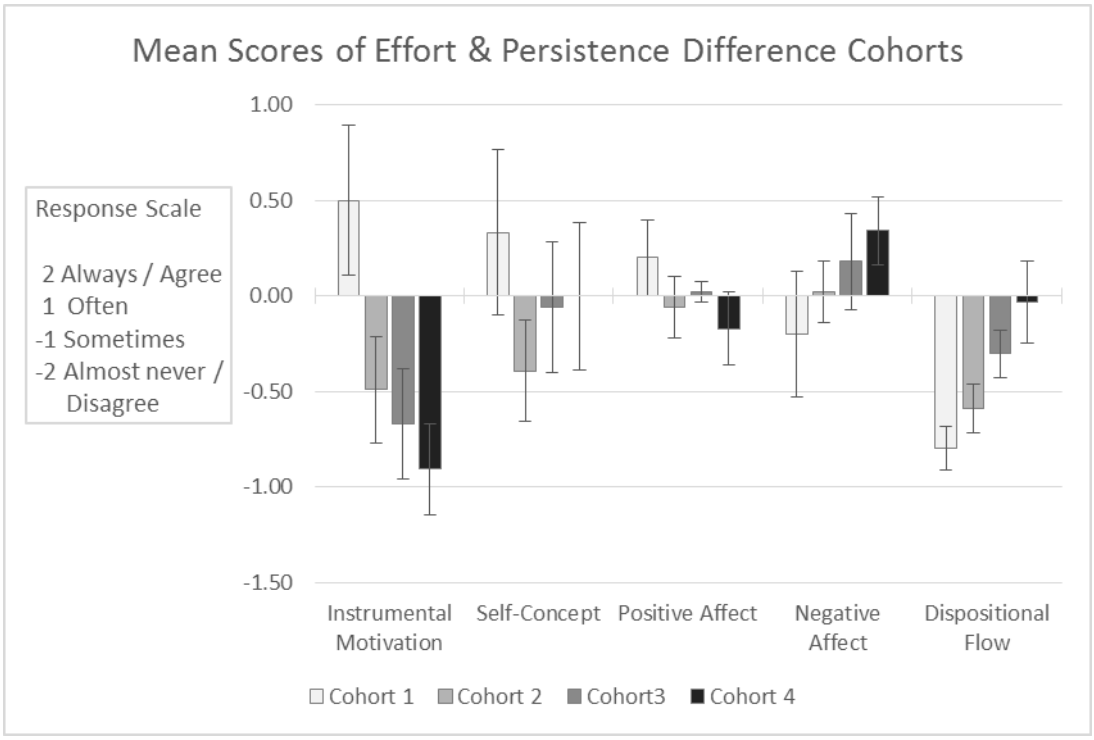

Figure 5-12 Mean scores of individual traits for cohorts determined by effort & persistence difference scores

Using the same method for mean difference scores of self-efficacy shows positive affect had little apparent influence on changes in confidence. Figure 5-13 shows the results of dividing students into cohorts based on self-efficacy individual difference scores. Here, cohort 1 consisted of students with difference scores below $D = -0.84$, cohort 2 and cohort 3 consisted of students within one standard deviation of the mean ($M = 0.11$), and the scores of cohort 4 were all greater than $D = 1.05$. Dispositional flow scores showed correlation to individual self-efficacy difference scores ($r(41) = 0.34$, $p = .03$). There is a noticeable difference between the mean scores of cohort 1 and those of cohort 4 ($D = 0.84$, $d = 1.56$), although the difference is not statistically significant ($p = .06$). On the other hand, the relationship between positive affect scores and difference scores for self-efficacy was rather inconsequential ($M_{C1} = 0.04$, $SD_{C1} = 0.09$; $M_{C2} = 0.11$, $SD_{C2} = 0.62$; $M_{C3} = -0.05$,

SD_{C3} = 0.76; M_{C4} = -0.07, SD_{C4} = 0.12), suggesting that the correlations we observed on the pretest were not indicative of changes observed over time.

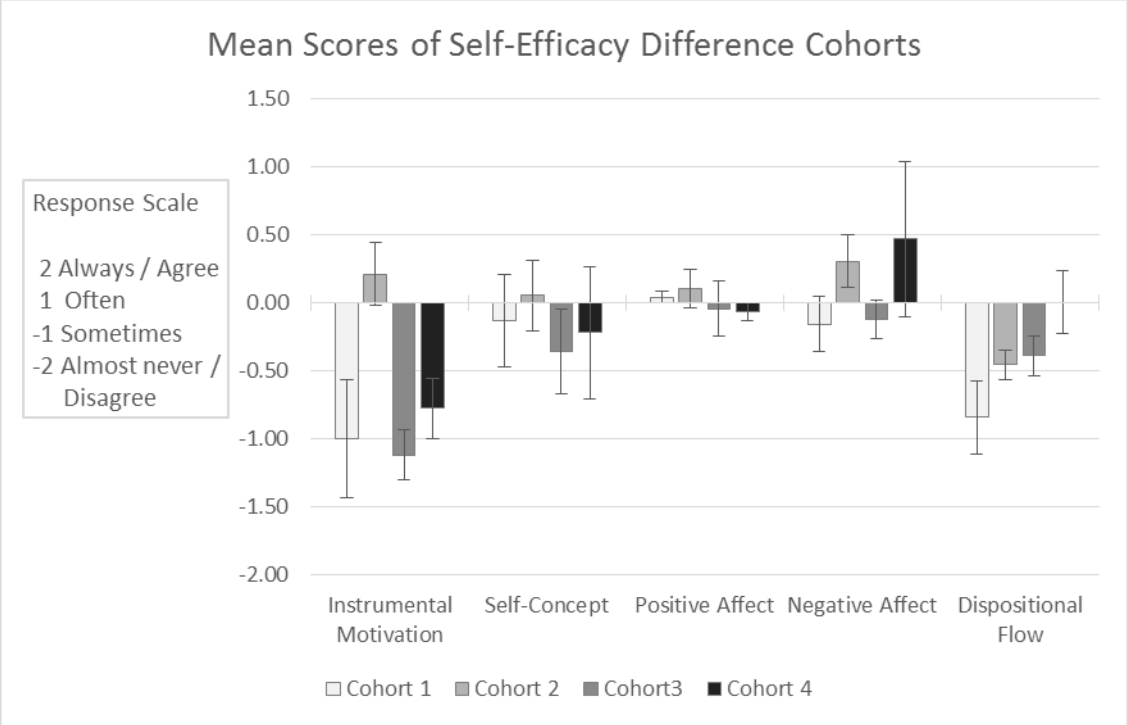

Figure 5-13 Mean scores of individual traits for cohorts determined by self-efficacy difference scores

5.3 DISCUSSION

Our findings illustrate a complex picture for the motivation and confidence of individual students. Given that one of the major criticisms of gamification is its superficial use of rewards for motivation, we focused specifically on aspects of intrinsic motivation and psychological needs in our measurements. Preexisting traits such as instrumental motivation and self-concept correlated with qualities of interest, effort, and self-efficacy for students at the beginning of semester. Experiences of the various flow elements throughout the semester were then connected to some psychological outcomes but not others. Also, the use of the gamification element seemed effective when paired with the Advanced English I: Conversation materials, but not the International Communication II materials.

Changes in motivation and confidence in the International Communication II course appear to be connected to initial levels of instrumental motivation and self-concept. Instrumental motivation and dispositional flow were the key factors for changes in

the self-efficacy of both groups and the effort of the control group. Meanwhile the initial interest of the students in the skill bars group was related to self-concept, and their initial effort correlated with positive affect. In the end, the skill bars add-on did not appear to have an effect on the flow experiences of the students, although it might have influenced gains in interest. Their interest scores were initially lower than those of the control group but came to match them in the end. As for the class as a whole, the relationship between dispositional flow and self-efficacy appeared to remain significant especially for the students with high gains or losses in self-efficacy. Similarly, students with high negative affect reported the largest drops in self-efficacy while students with low negative affect reported the largest gains.

The Advanced English I: Conversation sample shows variation in the relationships of individual traits and dispositional flow with psychological outcomes. The control group showed a change in interest that was unrelated to any of the measured variables of motivation, confidence, affect or flow. The change may instead be attributed to other factors that influence student performance in a classroom setting, such as developmental qualities of the students or teacher efficacy. Changes in self-efficacy were related to dispositional flow for the whole class. While the high ability group reported the most frequent flow experiences and saw the greatest improvement in self-efficacy, the control group reported the lowest flow scores and saw a drop in self-efficacy. This is significant especially because it shows that even the low ability students could perform better in this regard than the control group. Similarly, effort & persistence scores improved more for students who reported a greater frequency flow experiences. The mean difference scores for each group were small, but positive for the skill bars groups and negative for the control group. Moreover, the flow items directly pertaining to the design of the skill bars add-on showed significant differences between the students that used the add-on and the ones that did not.

The common theme for the two course samples appears to be that flow experiences directly contribute to feelings self-efficacy. Relating these findings to the three basic psychological needs proposed by Self-Determination Theory, we can see that feelings of self-efficacy are synonymous with feelings of competence. The other needs of autonomy and relatedness are not directly measured by our constructs for

motivation and confidence; although, their existence to various degrees in the students likely contributed to the resulting psychological outcomes.

The skill bars add-on showed positive effects for the Advanced English I: Conversation course, but not the International Communication II course. Based on our observations, we postulate this is because of a number of factors that hampered the potential effectiveness of the skill bars. First, the course assignments centered on public speaking activities that required students to practice English speaking skills in front of their peers. Second, the students using the skill bars in particular were more inclined to negative affect and less inclined to positive affect, which was likely exacerbated by the public speaking exercises. Another potential contributing factor that was not deliberately measured was the degree of clarity in the way the skills and subskills were written.

For comparison, the low ability group of students in the Advanced English I: Conversation course managed higher flow scores than the skill bars group in International Communication II despite high negative affect scores and final interest, effort and self-efficacy scores that were no higher. Advanced English I: Conversation was an elective course, which implies a higher degree of autonomy for enrolling in the course, and assignments were carried out without an audience. The teachers in Advanced English I: Conversation also provided considerably finer detail in their feedback to the students compared to the general rubric for presentations in International Communication II. Overall, we believe this shows that the efficacy of gamification tools and practices, such as the skill bars, relies heavily on how they are used, as demonstrated by the English conversation assignments.

5.4 LIMITATIONS

The sample selection is a major limitation for this study as it includes only students studying at Kanazawa Technical College. Students were assigned based on English ability according to the personal judgement of individual English teachers which may have been limited in its objectivity. Furthermore, the International Communication II course only included students in the Global Information Technology department, which are likely not representative of students studying in other areas. The students taking Advanced English I: Conversation elected to enroll in the course, which indicates a certain degree of volition in their participation;

whereas, the students in International Communication II were all required to enroll. This aspect alone may have inhibited a sense of autonomy for some students but not others. Moreover, given that games are commonly regarded as voluntary activities (Juul, 2003), the students' ability to experience the activities as game-like may also have been limited. In future experiments it would be worth exploring the implications of gamifying a required course as compared to an elective one.

This study aimed to identify the effects of four independent variables on four dependent variables. Such a high level of complexity in addition to the combination of multiple groups with pre-post style surveys resulted in an experimental design that was overly complicated in the interpretation of its results. The decision to run this experiment in a real classroom instead of a testing environment likely introduced numerous external factors that could have skewed the results. Implementation of the skill bars condition varied between courses as well as between sections within the Advanced English I: Conversation course. The decision to use the add-on mechanic was made after the course had already been designed, which limited the degree to which the gamification element could be integrated into activities.

Additional factors such as teacher efficacy and developmental qualities of the students were not controlled for and may have influenced the results. An extensive, international meta-analytic review of empirical studies on student achievement determined that teachers on average are able to achieve an effect size between $d = 0.20$ and $d = 0.40$ over a given year (Hattie, 2009). For this reason we looked specifically for effect sizes over $d = 0.40$, but the individual effects of the teachers in this case could not be measured.

Future research might be to define simpler, more precise experiments based on the implications of these findings. Larger sample sizes and very deliberate planning within test environments would be the most effective for supporting or contradicting our findings here.

6 CONCLUSIONS

In this paper, we presented a discussion about gamification and gameful design as applied to educational contexts. Compared to game-based learning, which has a potential issue with knowledge transfer between contexts, gameful learning is proposed as an approach to learning in authentic contexts that are structured in a way that satisfies basic psychological needs through playful experiences. We introduced the Playful Affordances Model as a means of evaluating autotelic qualities of game-like experiences, and hypothesized that various types of play can give rise to intrinsic motivation. Finally, we reported the results of two studies: the evaluation of autotelic experiences presented by a game-based learning activity, and the identification of individual qualities that may influence the prospect of optimal engagement in a gamified classroom setting.

Our findings of the first study showed specific predictors of enjoyment that are likely influenced by the design of the game. Correlations between items within the same dimensions of the model proved to be strong for sensation/arousal and contest/challenge but weak for exploration/discovery and imagination/creativity. Further work should examine other types of games to determine if these results are attributed specifically to the design of the chosen game, or if behavioral items in the model do not generally correspond with the given experiences. Nevertheless, we offer this study as an example of measuring game-like experiences for autotelic qualities that pertain to learner motivations.

The second study revealed details of student motivation and confidence and their relationships to particular psychological needs in the classroom. Overall, student self-efficacy of individual English ability appeared to be the most sensitive learning outcome to experiences of flow. The use of a skill bars gamification element in the public speaking class did not enhance the dispositional flow of students with higher negative affect, but may have been connected to a rise in interest. Meanwhile, student groups in the conversation class showed that successful promotion of flow experiences may be effective no matter the level of English ability. Students using the skill bars to complete speaking exercises showed gains in self-efficacy and effort, which correlated with reported frequencies of flow.

Self-Determination Theory provides a broad yet firm basis for the study of human motivation. Given the high-level of abstraction from which the theory describes the psychological desires of human beings, it is difficult to devise specific solutions for any given situation. Gamification was inspired by the success of video games, but gameful design was born from a close inspection of the experiential affordances therein. Games and play are powerful aspects of human culture which deserve a greater understanding if they are to be applied more broadly to daily life. The challenge of gameful design is to provide a playful context where one might otherwise not exist, and enhance goal-seeking behaviors towards some genuine purpose. As researchers, we present this attempt to navigate a minefield of distractions and find some deeper potential behind the shallow claims of gamification. We anticipate that our fellow researchers may find analytical value in our conceptualization of playful affordances, and practitioners might gain inspiration from our skill bars add-on. On the whole, educators in general should gain insights into their students and potential factors that make them lose interest, become lazy, or come to doubt themselves in class. Human motivation is complicated, but it must be understood for the future of human learning.

7 REFERENCES

Apter, M. J. (1991). A structural phenomenology of play. In *Adult play: a reversal theory approach* (pp. 13-30).

Apter, M. J. (2001). *Motivational styles in everyday life: A guide to reversal theory.* Washington D.C.: American Psychological Association.

Artelt, C., Baumert, J., Julius-Mc-Elvany, N., & Peschar, J. (2003, Sep 29). Learners for Life: Student Approaches to Learning. OECD Publishing. Retrieved Apr 05, 2015, from http://www.oecdbookshop.org/en/browse/title-detail/?isb=9789264103917

Barab, S. A., Ingram-Goble, A., & Warren, S. (2009). Conceptual Play Spaces. In R. E. Ferdig, *Handbook of Research on Effective Electronic Gaming in Education* (Vol. 3, pp. 989-1009). Hershey, PA: IGI Global.

Bartle, R. A. (2004). *Designing Virtual Worlds.* New Riders Publishing.

Bateson, G. A. (1972). Theory of Play and Fantasy. In G. A. Bateson (Ed.), *Steps to an Ecology of Mind* (pp. 177-193). Chicago: The University of Chicago Press.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis.* John Wiley & Sons, Ltd.

Caillois, R. (1961). *Man, play, and games.* Chicago: University of Illinois Press.

Costello, B., & Edmonds, E. (2007). A study in play, pleasure and interaction design. *Proceedings of the 2007 conference on Designing pleasurable products and interfaces* (pp. 76-91). ACM.

Csikszentmihalyi, M. (1990). *Flow.* New York: Harper Perennial.

Csikszentmihalyi, M., & Bennett, S. (1971). An Exploratory Model of Play. *American Anthropologist, 73*(1), 45-58.

De Koven, B. (2013). *The Well-Played Game.* The MIT Press.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin, 125*(6), 627 - 668.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification". *MindTrek '11* (pp. 9-15). New York: ACM.

Dichev, C., Dicheva, D., Angelova, G., & Agre, G. (2014). From Gamification to Gameful Design and Gameful Experience in Learning. *Cybernetics and Information Technologies, 14*(4), 80-100.

Flanagan, M. (2009). *Critical Play: Radical Game Design.* The MIT Press.

Fullerton, T. (2008). *Game Design Workshop.* Boca Raton: CRC Press.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation and learning: A research and practice model. *Simulation & gaming, 33*(4), 441-467.

Gee, J. P. (2007). *What Video Games Have to Teach Us About Learning and Literacy.* Palgrave Macmillan.

Gee, J. P. (2010). A situated-sociocultural approach to literacy and technology. In E. A. Baker (Ed.), *The new literacies: Multiple perspectives on research and practice* (pp. 165-193).

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale-2. *Computers in Human Behavior, 40*, 133-143.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? — A Literature Review of Empirical Studies on Gamification. *47th Hawaii International Conference on System Sciences.* Hawaii, USA.

Hattie, J. A. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge.

Huizinga, J. (1955). *Homo ludens: a study of the play element in culture.* Boston: Beacon Press.

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. *AAAI workshop on Challenges in Game AI.* AAAI Press.

Isen, A. M., & Reeve, J. (2005, December). The Influence of Positive Affect on Intrinsic and Extrinsic Motiation: Facilitating Enjoyment of Play, Responsible Work Behavior, and Self-Control. *Motivation and Emotion, 29*(4), 297-325.

Jackson, S. A., Martin, A. J., & Eklund, R. C. (2008). Long and Short Measures of Flow: The Construct Validity of the FSS-2, DFS-2, and New Brief Counterparts. *Journal of Sport and Exercise Psychology, 30*, 561-587.

Juul, J. (2003). The Game, the Player, the World: Looking for a Heart of Gameness. *Level Up: Digital Games Research Conference Proceedings*, (pp. 30-45). Utrecht.

Kawabata, M., Mallett, C. J., & Jackson, S. A. (2008). The Flow State Scale-2 and Dispositional Flow Scale-2: Examination of factorial validity and reliability for Japanese adults. *Psychology of Sport and Exercise, 9*, 465-485.

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *Internet and Higher Education, 8*, 13-24. doi:10.1016/j.iheduc.2004.12.001

Klahr, D., & Chen, Z. (2011). Finding one's place in transfer space. *Child Development Perspectives, 5*(3), 196-204.

Korhonen, H., Montola, M., & Arrasvuori, J. (2009). Understanding playful user experience through digital games. *International Conference on Designing Pleasurable Products and Interfaces*, (pp. 274-285).

Koster, R. (2004). *A Theory of Fun for Game Design.* Paraglyph Press.

Kumar, J. M., & Herger, M. (2013). *Gamification at Work: Designing Engaging Business Software.* The Interaction Design Foundation.

Lieberoth, A. (2015). Shallow Gamification: Testing Psychological Effects of Framing an Activity as a Game. *Games and Culture, 10*(3), 229-248.

Linderoth, J. (2012). Why gamers don't learn more: An ecological approach to games as learning environments. *Journal of Gaming and Virtual Worlds, 4*(1), 45-62. doi:10.1386/jgvw.4.1.45_1

Malaby, T. M. (2009). Anthropology and play: the contours of playful experience. *New Literary History, 40*(1), 205-218.

McGonigal, J. (2008). Why I love bees: A case study in collective intelligence gaming. In *The ecology of games: connecting youth, games, and learning* (pp. 199-228).

McGonigal, J. (2011). *Reality Is Broken: Why Games Make Us Better and How They Can Change the World.* Penguin Group.

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013). Disassembling Gamification: The Effects of Points and Meaning on User Motivation and Performance. *CHI 2013 Extended Abstracts* (pp. 1137-1142). Paris, France: ACM.

Montola, M., Stenros, J., & Waern, A. (2009). *Pervasive Games.* Burlington: Morgan Kaufmann.

Morris, B. J., Croker, S., & Zimmerman, C. (2013). Gaming science: the "Gamification" of scientific thinking. *Frontiers in psychology, 4*.

Pavlas, D. (2010). A Model of Flow and Play in Game-Based Learning: The Impact of Game Characteristics, Player Traits, and Player States. Orlando, Florida: University of Central Florida.

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

Robertson, M. (2010, October 6). *Can't play, won't play*. (Hide&Seek) Retrieved June 2014, from http://hideandseek.net/2010/10/06/cant-play-wont-play/

Rughiniş, R. (2013). Work and Gameplay in the Transparent 'Magic Circle'of Gamification. In *Design, User Experience, and Usability. Health, Learning, Playing, Cultural, and Cross-Cultural User Experience* (pp. 577-586). Berlin Heidelberg: Springer.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78.

Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals.* The MIT Press.

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

Songer, R. W., & Miyata, K. (2014). A Playful Affordances Model for Gameful Learning. *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality.* ACM.

Songer, R. W., & Miyata, K. (in press). A Playful Affordances Approach to the Design of Gameful Learning. *International Journal of Engineering Education: Special Issue on Gamification in Engineering Education*.

Steinkuehler, C., & Squire, K. (2014). Videogames and Learning. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences, Second Edition.* New York: Cambridge University Press.

Stenros, J. (2012). In Defence of a Magic Circle: The Social and Mental Boundaries of Play. *DiGRA Nordic 2012 Conference*, (pp. 1-18).

Sutton-Smith, B. (2001). *The Ambiguity of Play.* Harvard University Press.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the postitive and negative affect schedule (PANAS). *Journal of cross-cultural psychology, 38*(2), 227-242.

Tulloch, R. (2014). Reconceptualising Gamification: Play and Pedagogy. *Digital Culture & Education, 6*(4).

Walther, B. K. (2003). Playing and Gaming: Reflections and Classifications. *The International Journal of Computer Game Research, 3*(1).

Wang, C. J., Khoo, A., Liu, W. C., & Divaharan, S. (2008). Passion and Intrinsic Motivation in Digital Gaming. *CyberPsychology & Behavior, 11*(1), 39-45. doi:10.1089/cpb.2007.0004

Wang, H., & Sun, C.-T. (2011). Game Reward Systems: Gaming Experiences and Social Meanings. *Proceedings of the 5th Digital Games Research Association*, (pp. 14-17).

Werbach, K., & Hunter, D. (2012). *For the Win: How Game Thinking Can Revolutionize Your Business.* Wharton Digital Press.

Yee, N. (2006, December). Motivations for Play in Online Games. *CyberPsychology & Behavior, 9*(6), 772-775.

Zichermann, G., & Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps.* O'Reilly Media.

Zichermann, G., & Linder, J. (2013). *The Gamification Revolution: How Leaders Leverage Game Mechanics to Crush the Competition.* McGraw-Hill.

APPENDIX A: PLAYFUL AFFORDANCES SURVEY

[English Version]

Please answer the following questions by selecting the most appropriate number on the given scale.

1. How would you rate your anticipated enjoyment of the game BEFORE you started playing?

(No enjoyment at all) 0 1 2 3 4 5 (Greatly enjoyable)

2. How would you rate your enjoyment of the game now AFTER you have finished playing?

(No enjoyment at all) 0 1 2 3 4 5 (Greatly enjoyable)

3. I put a lot of effort into performing as best I could in the game.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

4. The game was too slow or boring for me.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

5. I tried various different ways to make deals / operate in my team.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

6. I do not care for the business theme in the game.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

7. I could imagine what it must be like to form a business / be a businessman.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

8. I did not change my tactics during the game.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

9. I was excited to make deals / establish businesses / work towards a high score.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

10. The game was too easy.

(Completely disagree) 0 1 2 3 4 5 (Completely agree)

11. Please answer to what degree you experienced the following things during the game:

a. Achievement (Not at all) 0 1 2 3 4 5 (Great amount)
b. Arousal (Not at all) 0 1 2 3 4 5 (Great amount)
c. Challenge (Not at all) 0 1 2 3 4 5 (Great amount)
d. Creativity (Not at all) 0 1 2 3 4 5 (Great amount)

e. Curiosity	(Not at all)	0	1	2	3	4	5	(Great amount)
f. Discovery	(Not at all)	0	1	2	3	4	5	(Great amount)
g. Fantasy	(Not at all)	0	1	2	3	4	5	(Great amount)
h. Thrill	(Not at all)	0	1	2	3	4	5	(Great amount)

12. Other than the above, what else did you think was fun about this game, if anything?

[Japanese Version]

Please answer the following questions by selecting the most appropriate number on the given scale.

1. 今回のゲームを始める前に、どれくらい楽しみを期待しましたか？
（全然期待しなかった） 0 1 2 3 4 5 （とても期待した）
2. 今ゲームが終了して、どれくらい楽しみましたか？
（全然楽しまなかった） 0 1 2 3 4 5 （とても楽しんだ）
3. ゲーム中、出来る限り努力をして活動に取り組んだ。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
4. 私にとって今回のゲームはスピードが遅すぎて退屈だった。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
5. 交渉したり、チーム運営をするために様々な方法を試した。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
6. ビジネスというテーマのゲームにあまり興味がない。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
7. ビジネスをすることやビジネスマンとはどのようなものか想像することができた。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
8. ゲーム中、自分の戦術を変えなかった。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
9. 交渉、ビジネス設立、ハイスコアを得ることに興奮した。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）
10. このゲームは簡単過ぎだった。
（そうは思わない） 0 1 2 3 4 5 （とてもそう思う）

11.ゲーム中、次の感情をどの程度経験できたか答えてください。

a. 達成感	（全然なかった）	0	1	2	3	4	5 （とても感じた）
b. 盛り上がり	（全然なかった）	0	1	2	3	4	5 （とても感じた）
c. 挑戦	（全然なかった）	0	1	2	3	4	5 （とても感じた）
d. 創造	（全然なかった）	0	1	2	3	4	5 （とても感じた）
e. 好奇心	（全然なかった）	0	1	2	3	4	5 （とても感じた）
f. 発見	（全然なかった）	0	1	2	3	4	5 （とても感じた）
g. ファンタジー（空想）	（全然なかった）	0	1	2	3	4	5 （とても感じた）
h. スリル感（興奮度）	（全然なかった）	0	1	2	3	4	5 （とても感じた）

12. 上記以外、ゲームは「楽しい」と思ったところがあったら書いてください。

APPENDIX B: ENGLISH CLASS SURVEYS

[Pretest Survey]

This is an optional survey about your feelings for learning English. Please answer only if you agree to let this information be used anonymously for educational research. Do not answer if you cannot answer honestly.
このアンケートは英語学習に対する研究調査のためのものであり、回答は任意のものです。すべての情報は教育研究の目的で使用され匿名で扱われます。このアンケートの利用目的に同意いただける場合には回答をお願いします。同意できない、あるいは正直に答えることができない場合は、回答の必要はありません。

1.) Because English is fun, I wouldn't want to give it up.
英語は楽しいから諦めたくない。
Select one:
・ disagree / そう思わない
・ disagree somewhat / あまりそう思わない
・ agree somewhat / まあそう思う
・ agree / そう思う

2.) English is important to me personally.
英語は個人的に重要なものである。
Select one:
・ disagree / そう思わない
・ disagree somewhat / あまりそう思わない
・ agree somewhat / まあそう思う
・ agree / そう思う

3.) When practicing English, I sometimes get totally absorbed.
英語を練習すればとても夢中になることがある。
Select one:
・ disagree / そう思わない
・ disagree somewhat / あまりそう思わない
・ agree somewhat / まあそう思う
・ agree / そう思う

4.) When practicing English, I try as hard as possible.
英語を練習するとき、一所懸命やる。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

5.) When practicing English, I keep at it even if it is difficult.
英語を練習するとき、難しくてもよく頑張る。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

6.) When practicing English, I try my best to acquire the skills taught.
英語を練習するとき、教わった技能を習得できるようにする。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

7.) I'm confident I can understand the most difficult material presented in English class.
英語科目のもっとも難しい内容を理解できる自信がある。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

8.) I'm confident I can do an excellent job on English assignments.
英語の課題を上手く仕上げる自信がある。
Select one:

・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

9.) I'm certain I can master the skills being taught.
教わった技能を習得できる自信がある。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

10.) I study English to increase my job opportunities.
就職の機会を増やすために英語を勉強している。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

11.) I practice English because I expect to use it often in the future.
将来、頻繁に使うと考えているから英語を勉強している。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

12.) I practice English to get a good job.
良い仕事に就くために英語を勉強している。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

13.) I'm hopeless in English class.

私は英語となると全くお手上げである。

Select one:

・ disagree / そう思わない

・ disagree somewhat / あまりそう思わない

・ agree somewhat / まあそう思う

・ agree / そう思う

14.) English is one of my best subjects.

英語は得意な科目の一つである。

Select one:

・ disagree / そう思わない

・ disagree somewhat / あまりそう思わない

・ agree somewhat / まあそう思う

・ agree / そう思う

15.) I have always done well in English classes.

英語の科目はいつも優秀である。

Select one:

・ disagree / そう思わない

・ disagree somewhat / あまりそう思わない

・ agree somewhat / まあそう思う

・ agree / そう思う

16.) Thinking about yourself and how you normally feel, how often do you generally feel each of the following? Please write in a number 1-5 on the following scale: (never) 1 2 3 4 5 (always)

下記の事項について、自分自身のあり方について普段どのように感じていますか？当てはまる数字を書いてください。

（全く感じない）1 2 3 4 5 （いつも感じる）

_____ Upset / いらいらする

_____ Hostile / 愛想が無い

_____ Alert / 抜け目の無い

_____ Ashamed / 恥ずかしがり

_____ Inspired / 意欲がある

_____ Nervous　/　緊張する

_____ Determined　/　意志の固い

_____ Attentive　/　用心深い

_____ Afraid　/　心配する

_____ Active　/　積極的

[Posttest Survey]

This is an optional survey about your feelings for learning English in this class. Please answer only if you agree to let this information be used anonymously for educational research. Do not answer if you cannot answer honestly.
このアンケートは授業中の英語学習に対する研究調査のためのものであり、回答は任意のものです。すべての情報は教育研究の目的で使用され匿名で扱われます。このアンケートの利用目的に同意いただける場合には回答をお願いします。同意できない、あるいは正直に答えることができない場合は、回答の必要はありません。

Please answer the following questions according to your experience studying in this class.
この授業中の学習体験に基づき、以下の質問に答えてください。

1.) Because English is fun, I wouldn't want to give it up.
英語は楽しいから諦めたくない。
Select one:
・ disagree / そう思わない
・ disagree somewhat / あまりそう思わない
・ agree somewhat / まあそう思う
・ agree / そう思う

2.) English is important to me personally.
英語は個人的に重要なものである。
Select one:
・ disagree / そう思わない
・ disagree somewhat / あまりそう思わない
・ agree somewhat / まあそう思う
・ agree / そう思う

3.) When practicing English, I sometimes get totally absorbed.
英語を練習すればとても夢中になることがある。
Select one:
・ disagree / そう思わない
・ disagree somewhat / あまりそう思わない

・ agree somewhat / まあそう思う
・ agree / そう思う

4.) When practicing English, I try as hard as possible.
英語を練習するとき、一所懸命やる。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

5.) When practicing English, I keep at it even if it is difficult.
英語を練習するとき、難しくてもよく頑張る。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

6.) When practicing English, I try my best to acquire the skills taught.
英語を練習するとき、教わった技能を習得できるようにする。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

7.) I'm confident I can understand the most difficult material presented in English class.
英語科目のもっとも難しい内容を理解できる自信がある。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

8.) I'm confident I can do an excellent job on English assignments.
英語の課題を上手く仕上げる自信がある。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

9.) I'm certain I can master the skills being taught.
教わった技能を習得できる自信がある。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

10.) I have a strong sense of what I should do to improve my English.
私の英語力を向上するために何をすればいいかを強く自覚している。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

11.) I have a good idea while practicing about how well I am doing.
練習している間、自分がどれだけ上手くできているかが感じられる。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

12.) I feel I am competent enough to meet the high demands of the class.
授業中の高い要求に対応するだけの能力が、自分にはある気がする。
Select one:
・ almost never / ほとんどそうではない

・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

13.) I can use English spontaneously and automatically without having to think.
考えようとしなくても、自然にパットと英語を使うことができる。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

14.) I can completely focus on the task at hand.
目の前の課題に完全に集中している。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

15.) I have a feeling of total control over what I am doing.
自分がしていることを完全にコントロールしているような気がする。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

16.) I do not care about what other people might think of me.
他の人が自分のことをどのように思っていようと、気にならない。
Select one:
・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

17.) The way time passes during class seems to be different from normal.

授業中、時間の経ち方が普段とは違うような気がする。

Select one:

・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである

18.) Class is extremely rewarding.

授業にはすごくやりがいがある。

Select one:

・ almost never / ほとんどそうではない
・ sometimes / ときどきそうである
・ often / よくそうである
・ always / いつもそうである