POSITIVE PSYCHOLOGY AND SECOND LANGUAGE MOTIVATION: EMPIRICALLY VALIDATING A MODEL OF POSITIVE L2 SELF

A Dissertation
Submitted
to the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education

By
J. Lake
December, 2015

Examining Committee Members

Mark Sawyer, Advisory Chair, Kwansei Gakuin University
David Beglar, Teaching and Learning
Tamara Swenson, External Member, Osaka Jo Gakuin University
Yoko Kozaki, External Member, Mukogawa Women's University
James Elwood, External Member, Meiji University

ABSTRACT

Positive psychology is rapidly developing as a field in psychology. Many constructs associated with positive psychology have been developed but relationships have not been demonstrated to second language (L2) learning or L2 learning motivation. The main purpose of this study was to explore empirically some core constructs of positive psychology and L2 learning motivation by testing a structural model of the causal relationships among levels of self-concept, and L2 proficiency. In order to do that, it was first necessary to validate measureable components of each of the levels. The self-concept constructs were: a global *positive self-concept*, a domain-specific *positive L2 self*, and L2 skill specific *self-efficacy*. The various *self*-constructs were organized into finer levels of specificity, from the global to L2 domain to L2 domain skills.

A structural model was created from three latent variables that were in turn created from measured variables at each level of specificity. For the latent positive self-concept the measured variables consisted of *flourishing*, *hope*, and *curiosity*. For the latent variable of positive L2 self the measured variables consisted of an *interested-in-L2 self*, *passion-for-L2-learning self*, and *L2 mastery goal orientation*. For the latent motivational variable of L2 self-efficacy the measured variables were *L2 speaking self-efficacy*, *L2 listening self-efficacy*, and *L2 reading self-efficacy*. The measured variables were based on adapted or newly created self-reports.

To demonstrate that the model holds beyond self-reports, objective L2 proficiency measures were also modeled with the latent variables of positive self-concept and positive L2 self. To demonstrate the generalizability of the self-model with L2 proficiency, a cross-validation study was done with two different objective measures of L2 proficiency, TOEIC and TOEIC Bridge.

The results for the study were all positive for the creation of composite variables and fit to causal models. Latent variables were created for a composite positive self-concept, a composite positive L2 self, and a composite L2 motivation variable. The positive self-concept and positive L2 self also fit a model that included an objective measure of L2 proficiency. Finally, structural equation modeling confirmed causal relationships among positive self-concept, positive L2 self with both L2 motivation and with L2 proficiency.

This study showed how constructs from the rapidly expanding field of positive psychology can be integrated with second language motivation. This study showed one way positive psychology can be applied to second language learning and suggests that positive psychology might invigorate future L2 motivation studies.

ACKNOWLEDGMENTS

The research and writing of this dissertation took many years to complete. This dissertation also reflects the knowledge that I gained in the courses for the Temple University doctoral program. I am grateful for all help, support, and encouragement that I received over the years.

I would like to thank my dissertation committee, Mark Sawyer, David Beglar, Tamara Swenson, Yoko Kozaki, and James Elwood, for their many suggestions and detailed feedback. I wish to especially thank Mark Sawyer who has spend countless hours and effort over several years in providing me feedback and steering me through the dissertation process to completion.

I appreciate what I have learned from fellow students I have studied with at Temple in Tokyo, Osaka, and Fukuoka. I would especially like to thank Brad Visgatis, Matt Apple, Peter Neff, Lance Burrows, Harumi Kimura, Tsuyuki Miura, Keita Kikuchi, Noriko Lake, Paul Westrick, John Eidswick, and Andrew Obermeier. Brad Visgatis was especially encouraging during the final drafts of the dissertation and helped me through a lot of confusion.

I learned much from the classes I took at Temple and they have done much in shaping what I know that went into the dissertation. Especially, as related to my dissertation, I would like to thank: Mark Sawyer for his classes on educational psychology, David Beglar for his classes on quantitative research, and J. D. Brown for his classes on quantitative research. In relation to my dissertation, I also benefited from classes or seminars with: Ken Schaefer, Steve Ross, Peter

MacIntyre, George Engelhard, Tim McNamera, Trevor Bond, Nick Ellis, Lyle Bachman, Zoltan Dörnyei, Jim Purpura, Marshall Childs, and Rod Ellis.

Friends and colleagues have also contributed towards my thinking in regards to my dissertation. I am thankful for discussions, moral support, and friendships with: Trevor Holster, Bill Pellowe, Kathleen Brown, Toshi Shiotsu, Rie Koizumi, Yo In'nami, Keiichi Minamikawa, Chris Chase, Quint Oga-Baldwin, Lexie Don, Darcy Delint, Mike Rogers, Neil McClelland, Dexter Da Silva, Robert Murphy, Curtis Kelly, Yuki Arita, and Jim Sick. I would also like to say a special thanks to Mike Linacre for answering some of my questions about Rasch analysis.

Thanks are due to my colleagues and student participants. Colleagues I wish to thank for discussions or help with data collection include: Art Nakano, Alan Williams, Phee Nugent, Mayumi Kido, Chiaki Hayashi, Jenifer Larson-Hall, August McInnis, and Jack Brajcich.

I would like to thank my family for all their love and support over the years. To my mom Tsutako, and dad David thanks for your love, support, and encouragement. My deepest thanks go to my wife, Noriko, I could not have done this dissertation without her support. She also was a great help with the Japanese portions of my study and constantly gave me helpful advice. My two daughters, Sara and Emma have always had a dad in graduate school and helped propel him along with love, laughter, and purpose. Thank you!

DEDICATION

Dedicated to my family, Noriko, Sara, and Emma.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The Background of the Issue

In an article on self-concept and achievement, Marsh and Craven (2006) called attention to the fact that “there is a revolution sweeping psychology, one that emphasizes a positive psychology and focuses on how healthy, normal and exceptional individuals can get the most from life” (p. 133). Positive psychology was founded as a movement within psychology when Martin Seligman decided to make it the theme of his presidency of the APA in 1998 (Seligman, 1999, 2002; Seligman, Steen, Park, & Peterson, 2005). This relatively new movement has generated a large literature introducing positive psychology to researchers and the public by key positive psychologists (e.g., Csikszentmihalyi & Csikszentmihalyi, 2006; Lopez & Snyder, 2009; Seligman, 2002, 2011); however, little relevant empirical research into the role of positive psychology in the field of second language (L2) education has been conducted.

At the same time, in the field of second language education, Dörnyei and Ushioda (2011) have noted that there has been “a gradual convergence of self theories and motivation theories in mainstream psychology” (p. 80). Dörnyei and colleagues have mostly focused on a possible future self that when contrasted with the present can guide a present self. Much of the work in positive psychology is focused on the present self. In positive psychology, as with humanistic psychology before it, the emphasis is on being authentic and true to oneself (Ryff & Singer, 2008; Schlegel & Hicks, 2011; Seligman, 2002; Sheldon, 2002). Elements of the future are often represented as goals. Importance is generally placed on how these goals are approached in the present rather than the discrepancy between the past and the present. For example, in hope theory, it is not the contrast with the future that is important but the belief that people can find ways of making progress toward their goals.

My personal interest in this area has a long history. In high school and university I would often read books in philosophy, psychology, theology, and science. Often I would go through cycles of being interested in psychology for a while then turn to philosophy for a while and then on to other areas just out of general broad interests or sometimes due to relationships with other students, classes I was taking, or faculty members. I remember reading many books by Sigmund Freud, Carl Jung, and Alfred Adler in my late teens. A few years later I read books by Joseph Campbell and then re-read many of the psychology books. I took a class in classics of philosophy and had a roommate who was well versed in classical philosophy so I read and we would talk about the pre-Socratics, and later Greek philosophers and how they later influence many theologians. These philosophers were early thinkers about what happiness and a flourishing life were all about.

During my college years for my bachelor of science degree, my main courses were in physics and chemistry because my plan at the time was to teach high school physics. My studies in psychology and philosophy were mostly outside of my main courses either in my own free time or in elective courses. Of course, psychology was touched on in the teacher education courses I took. My school required a capstone project done in the senior year and I did mine on the philosophy of science of Karl Popper.

During my graduate study, courses in psychology, individual differences, and educational psychology rekindled my long-held interests in psychology. The field of positive psychology began around the same time I started graduate study at the end of the 1990s. I had read books by Abraham Maslow, Carl Rogers, Rollo May, R. D. Laing and Ken Wilber, so I had some general knowledge of humanistic and transpersonal psychology. One of my language-focused interests is in language testing which uses quantitative analysis, so when I read about the new movement of psychology that was attempting to bring together the growth-oriented positive psychology through a more rigorous use of scientific methods, it resonated with my own personal interests.

One of the reasons I became a teacher is that I have been interested in helping people learn, develop, and grow as a person. Joseph Campbell used to advise his students to "follow your bliss" but it is not always clear how a student can be helped to do this. Positive psychology empirically studies what works in helping people grow, so there is a lot of overlap with my reasons for being an educator. I have similar feelings to something that Herbert Marsh said in an interview. He talked about learning about humanistic psychology but he said, "I wanted to combine this softer, more intuitive side of myself with my quantitative skills. Self-concept research seemed the ideal compromise. It is relevant and interesting to almost everyone" (Bembenutty, 2009, p. 538). When the field of positive psychology started, it was framed from the start as being different from humanistic psychology and the positive thinking movement because it used scientific methods and relied on "empirical research to understand people and the lives they lead" (Peterson & Seligman, 2004, p. 4).

Since the beginning of positive psychology much research has been done to identify character strengths, that is, constructs that help people function optimally. One weakness was the lack of organization of strengths, so various classification schemes have been proposed. Peterson and Seligman (2004) created the Values in Action (VIA) inventory of strengths that was composed of 24 strengths organized around six virtues. They define character strengths as being similar to personality, trait-like but more flexible as with "individual differences that are stable and general but also shaped by the individual's setting and thus capable of change" (p. 10). Linley (2009) offers another classification scheme with a larger number of 60 strengths and a broader definition of strengths that "is a pre-existing capacity for a particular way of behaving, thinking, or feeling that is authentic and energizing to the user, and enables optimal functioning, development and performance" (p. 9). Using strengths to help students learn and to live their lives more optimally aligns with some of my own goals as a teacher.

Developing a sense of agency, competence, learning, and enjoyment is an important part of education and unites the variables that are a part of the model that I propose. Underlying self-efficacy is a personal sense of agency and competence. The same is true of hope, although at a more general trait-level in my model. Agency, competence, learning, and enjoyment underlie domain-specific interest, passion for learning, L2 mastery goal orientation, curiosity, and is a large part of flourishing. A model uniting these positive psychology constructs could be put to use by teachers, learners, and others involved in learning environments but to clarify this, it is helpful to explain the statement of the problem, purposes, and theoretical overview.

Statement of the Problem

There are three main problems identified in creating this model of positive L2 self so that it can be empirically investigated. One problem is the validation of the components of a positive self. A second is the validation of the components of a positive L2 self. To investigate hypothesized relationships among these constructs and L2 motivational variables leads to a third problem: identifying the structural relationships among positive self-concept, positive L2 self, and motivational variables; and in an additional step that goes beyond self-reports, identifying the structural relationships among positive self-concept, positive L2 self, and second language proficiency.

The first problem addressed in this study is the construction of a composite latent constructs of general positive self-concept. Self-concepts can be grouped by level of generality. Some self-concepts are more general and stable and some are more specific and dynamic (Mercer, 2011, 2012). The first problem addressed is the more general level.

The second problem concerns the construction of a latent construct of a positive L2 self. There are many constructs that might contribute to a positive L2 self. There are also different theories of self-systems related to language learning (e.g., Dörnyei & Ushioda, 2011; Lau, Yeung, Jin, & Low, 1999; Marsh & Shavelson, 1985; Noels, 2009). The problem with a positive L2 self is that determinants of positive L2 self are yet unknown, as are the strengths of relationship among constructs contributing to an overall positive L2 self.

The third problem addressed in this study is the construction of a latent construct of L2 motivation. There is little empirical research on how constructs from positive psychology might be integrated with motivational constructs in the field of second language learning to form a positive L2 self. Of the many individual differences that influence a student to learn a second language, one of the most important is motivation (Dörnyei, 2005, 2009; Dörnyei & Skehan, 2003; Ellis, 2004

The fourth problem concerns the relationships among the levels of variables. Although a number of researchers have examined the hierarchical nature of self-concept (Lau, Yeung, Jin, & Low, 1999; Marsh & Shavelson, 1985; Yeung, Shui, Lau, McInerney, Russell-Bowie, & Suliman, 2000; Yeung & Wong, 2004), positive self-concept tends to be identified as self-esteem to the exclusion of other possible constructs, and a problem exists in that the relationships among positive self-concept, positive L2 self, and a measure of language proficiency. The ideal L2 self component of the self-system model created by Dörnyei (2005, 2009; Dörnyei & Ushioda, 2009, 2011) has received some empirical validation with self-report measures (Ryan, 2009; Taguchi, Magid, & Papi, 2009). This study addresses the problem of the unknown relationships among levels of variables and also included variables of objective measures of L2 proficiency.

Purposes and Significance of the Study

There were three main purposes of this study and one overarching purpose. The three main purposes were to: (a) construct and validate latent constructs of positive self-concept, L2 positive self, and L2 motivation; (b) explore the structural relationships among positive self-concept, positive L2 self, and L2 motivation; (c) explore the structural relationships among positive self-concept, positive L2 self, and L2 proficiency. An overarching purpose is to show how constructs from positive psychology can be integrated with constructs from second language learning motivation.

The first purpose of this study was the construction of a latent variable of a positive self with three core components of a positive self measured: flourishing, curiosity, and hope. A latent variable of a positive L2 self was constructed with components of interested-in-L2 self, a harmonious passion-for-L2-learning, and an L2 mastery goal orientation. A latent variable of L2 motivation was constructed using components of L2 speaking self-efficacy, L2 listening self-efficacy, and L2 reading self-efficacy.

The second main purpose concerned the structural relationships among the three levels of latent self-variables. The relationship paths are hypothesized to go from the more general to the more specific.

The third purpose was to determine the structural relationships among constructs of positive self-concept, positive L2 self, and L2 proficiency. The first two constructs are subjective measures based on self-reports. The construct of L2 proficiency was based on objective tests of language skills.

An overarching aim of this study is to show how constructs from positive psychology can be integrated with constructs from second language learning motivation. Although the two sets of constructs have intuitive similarities and connections, this study is the first time they have that they have been made explicit. One major area of significance for this study is that knowledge of the relationships among components of positive L2 self will contribute to an emerging area within second language motivation theory that combines concepts of the self with second language learning motivation (Dörnyei & Ushioda, 2009, 2011; Gregersen & MacIntyre, 2014; Mercer, 2011; Mercer & Williams, 2014). This will also lead to a better understanding of motivational constructs related to classroom learning situations and specific language skills and tasks. Many of the suggestions made in the early 1990s to expand the language learning motivational research agenda were not significantly followed up with research to become part of established second language motivation theory; this model could be a first step in opening again the motivational research agenda. The positive L2 self model and empirical research adds to the knowledge and understanding of current models of the self-system and L2 self.

The second area of significance of this study is that a demonstration of relationships between broad personality-type traits or global self-concepts with learning and beliefs about learning will help teachers and researchers understand how these beliefs might be integrated. Much of current research in second language motivation is highly fragmented in that isolated variables are studied without reference to how they might contribute to other research or to an organized whole. In addition, the abundance of motivation variables can be confusing to practitioners because of a lack of reference to level of generality is not given. This study provides a framework to organize past and future research. This study will also contribute to fields of expertise, discipline and disciplinary knowledge approaches to education, and educational psychology. More specifically, this study will contribute to understanding how language is learned in academic contexts. This is important for theories and research in language policy, curriculum design, materials design, and pedagogical practice.

This study also provides a much-needed demonstration of the practical applications of positive psychology. Even though the field of positive psychology is rapidly growing, there have been few applications in educational contexts that incorporate more than one construct. This study will show how multiple positive psychology constructs relate to each other. This study will contribute to the field of positive psychology by showing an application to education.

Theoretical Perspective

The theoretical perspective is based in part on the review and theoretical perspective outlined by Shavelson, Hubner, and Stanton (1976) where they reviewed self-concept studies and measures used in educational contexts and proposed a model of how the studies fit together. They defined self-concept as "a person's perception of himself" (p. 411). In addition, self-concept can be described as "organized, multifaceted, hierarchical, stable, developmental, evaluative, differentiable" (p. 411). They created a model that took account of all these features. In their hierarchy, there is a unitary general self-concept that is composed of lower order components that are less stable and more situation specific. Thus, a person might have a positive or negative academic self-concept that contributes to their general self-concept. Even if a person has a negative academic self-concept they can have non-academic self-concepts such as a physical self-concept that are strong and this contributes to their overall general self-concept. For example, a person might have a poor academic ability but be good at a particular sport and this contributes to their general self-concept.

In studies of self-concept, there might be positive or negative contributions within a hierarchy that leads to a more global self. To take academic self-concept as an example again, a person might have a positive academic self-concept even though they have a negative math self-concept because it might be another domain that contributes to the academic-self concept, for instance, a positive second language learning self-concept. The academic domain, in this instance, learning a second language, might be influenced by even more specific (positive) motivations or lack of (negative) motivation or "demotivations" (Kikuchi, 2013). Theoretically, it should be possible to examine the positive side of the self-concept hierarchy, to identify dimensions of a positive global self, to identify dimensions of a positive L2 self, and L2 motivational variables.

There are many constructs that use the term *self*, for example: self-concept, self-esteem, L2 self, and self-efficacy. These terms can be confusing if the hierarchy and multidimensionality of self-constructs are not kept in mind. This study is concerned with three levels of self: a global level, a domain level, and a situational level self (see Table 1). The global level self has no domain (other than the self); for example, a curious self is curious about many things. The domain level self has a specified domain; for example, a positive L2 self has the domain of second language learning. A situational level self is concerned with a particular situation or task within a domain, not the domain in general; for example, reading self-efficacy relates to the ones competence to do specific reading tasks. To keep these terms clear, self-concept, self, positive self will always refer to the global level; domain level self is always labeled with the domain; and situation level self is always labeled with a specific skill.

Table 1. *Levels, References, Examples*

Levels	In reference to	Examples
Global	Self	Curious self
Domain	Specific domain	Positive L2 self
Situational	Tasks	L2 Reading self-efficacy

The theoretical perspective also draws on the relatively new field of positive psychology (e.g., Csikszentmihalyi & Csikszentmihalyi, 2006; Diener & Biswas-Diener, 2008; Lopez, Pedrotti, & Snyder, 2014; Seligman, 2011). The theoretical perspective for the domain level motivational variables is based on theories of: achievement goal theory (e.g., Ames & Archer, 1988; Dweck & Leggett, 1988; Kaplan, Middleton, Urdan, & Midgley, 2002; Schunk, Pintrich, & Meece, 2008), interest (e.g., Hidi, 1990; Hidi & Renninger, 2006; Kashdan, 2009; Silvia, 2006), and self-efficacy (e.g., Bandura, 1977, 1986, 1997; Bandura & Schunk, 1981).

The many different theories of achievement goals, personal goal achievement, self-concept, and learning orientations date from the time of Aristotle (1974) to recently developed theories (McMahon, 2006; Waterman, 2013). Even a quick glance at educational psychology and motivation literature provides a reader with an overwhelming number of theories and claims, much of it confusing, some of it seemingly contradictory. In order to show the relationships among various theories and to organize them, this study takes the perspective that theories need to be compared at multiple levels of specificity. As of this writing there has been no proposal of a construct of positive L2 self. However, there have been suggestions of factors that might contribute to a self-system. In the field of second or foreign language teaching as part of a self-system model developed by Dörnyei (2005, 2009; Dörnyei & Ushioda, 2009, 2011) the concept of an ideal L2 self perhaps comes the closest, however, while that concept is concerned more with an ideal possible future self, the construct in this study a positive L2 self has more to do with interest, learning, challenge, enjoyment, and positive affect in the present self.

The Audience for the Study

The first audience for this study is second language acquisition (SLA) researchers investigating individual differences. The relationships among components of positive self-concept are significant to researchers working in areas of emotional, cognitive, and interpersonal approaches to self and identity. This study contributes to fields of self-concept and personality by showing how stable dispositions that are trait-like relate to other stable dispositions and how they relate to a narrower domain-specific self.

The second audience for this study is researchers in applied positive psychology for education. This study is significant to the theorists, researchers, and teachers in the fields of positive psychology, educational psychology, second language motivation, and second language learning. This study is also significant to researchers doing interdisciplinary work in these areas.

The third audience for this study is curriculum and program designers and administrators. In other words, it might interest people who apply research to practical problems, such as people who do: educational administration, educational leadership, curriculum design, materials development, language teacher training, and language teaching. It is theoretically possible to substitute other abilities in the place of language learning, so future audiences might include researchers and practitioners in other disciplines such as math, dance, physical education, and history.

Finally, the fourth audience for this study would be second language teachers and students. At the level of positive L2 self, this part of the study is most relevant to people specifically engaged in language learning issues such as: syllabus design, lesson planning, action research, and teacher assistance.

Delimitations

This study is delimited in a number of different ways. The site is a private institution that provides education from pre-school, through junior high school, high school, junior college, and university, to graduate school levels. This study was conducted with female participants in a two-year junior college and four-year university, so it is limited by location, gender, and age. A random sample of students was not feasible; instead, the participants were limited to the nine teachers and the students in their classes that were willing to participate. I did not investigate all possible positive psychological constructs, so there might be additional constructs that could be added or substituted in the model presented. This is a cross-sectional study, so it is unknown how the variables might change over time.

Based on the empirical data, the findings of this study should be generalizable to female student samples of similar age and proficiency level in Japan. Based on the literature review, if the results match findings of previous research, then this suggests that the results might be generalizable to other populations. The methodology of this study might also replicated by expanding to other populations and similar constructs as long as the specificity levels are maintained.

Organization of the Study

Chapter 2, Review of the Literature, is divided into five main sections: (a) positive psychology, (b) positive self-concept, (c) positive L2 self, (d), motivational constructs (e) and structural relationships involving self-related constructs and a measure of proficiency. Chapter 3, Methods, is divided into four main sections: participants, instrumentation, procedures, and the analyses employed in this study. In Chapter 4, Preliminary Analysis: Instrument Validation Evidence, I present validity evidence based on internal structure for the main instruments. Chapter 5, Preliminary Analysis: Peripheral Instrument Validation Evidence, concerns validity evidence based on internal structure for the supporting peripheral instruments used for convergent and divergent validity evidence in the following chapter. In Chapter 6, Preliminary Analysis: Convergent and Divergent Validity Evidence, I present external validity evidence for the main instruments through convergent and divergent relationships with the supporting instruments. Chapter 7, Results, has three sections in which the results of the five research questions are presented. In Chapter 8, Discussion, I interpret and discuss the findings of the research questions, and Chapter 9, Conclusion, is divided into four sections, a Summary of the Findings, Limitations, Suggestions for Further Research, and Final Comments.

Key Terminology

Terminology from the field of positive and educational psychology, and second language motivation are used in this study. The following terms are defined as follows:

Ideal L2 Self: A construct developed by Dörnyei (2005, 2009) based on the concepts of "possible selves" (Markus & Nurius, 1986) and "imagined community" (Norton, 2001). The ideal L2 self, states Dörnyei (2009b, p. 29) "is a powerful motivator to learn the L2 because of the desire to reduce the discrepancy between our actual and ideal selves."

Positive domain-specific self-concept: A positive domain-specific self-concept refers to a person's positive identity for a particular domain. This study does not examine negative aspects of the self such as anxiety. People can develop an identity in many domains, such as, sports, music, or academic fields (Bracken, 1996, 2009). Positive domain-specific self-concept is considered a middle-level abstraction as explained in more detail in the literature review.

Self-concept: A person's perspective of their own self is their self-concept. It is an identity-like concept that can refer to the global-self or self-as-a-whole, or it can refer to particular domains, or it can be even more specific as in relation to certain tasks or situations. In this study, if used alone, it refers to global self-concept, more specific self-concepts are used with more specific labels.

Self-Efficacy: A type of self-concept that is highly specific about one's competence to accomplish a particular task. As defined by Bandura (1997), self-efficacy refers to "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 3).

CHAPTER 2

REVIEW OF THE LITERATURE

Self-concept research, educational psychology, and L2 motivation are complex areas of study. Combining these areas into the present single interdisciplinary study makes it more complex. In this literature review, studies are described and explicated from multiple areas with general overviews that in turn narrow to the particular context of this study. Although the various disciplinary areas are complex, by elaborating the hierarchical nature of the generality to specificity continuum among specific research traditions, an understanding can develop so that initially complex and confusing details are understood. The various parts can then be put together into an integrated whole that clarifies much of the confusion that exists in L2 self and L2 motivation studies. The review of the literature begins with a brief overview Japan's educational environment and of positive psychology. In the second section literature related to positive self-concept and positive L2 self is reviewed. In the third section literature on motivational self-constructs, specifically, self-efficacy is reviewed. In the fourth section literature on structural relationships among these self-levels is reviewed. In the fifth section literature directly related to L2 motivation and learning is reviewed.

Japan's Educational Environment

Foreign language motivation has often been seen in a negative light in Japan. Students are often characterized as having low motivation or being demotivated. The situation in Japanese universities has been called, a "motivational wasteland" (Berwick & Ross, 1989, p. 207). Ushioda (2013, p. 6) points out that "Japan leads the field" in studies of L2 demotivation. In the Japanese educational context students are often represented as being reticent, afraid of making mistakes, or generally demotivated (McVeigh, 2002; Nakata, 2006; Sugimoto, 2010). Social anxiety and the phenomena of *hikikomori* or self-seclusion due to anxiety is recognized as a growing problem among Japanese youths affecting not only language learning but schooling (Furlong, 2008; Kaneko, 2006; Teo, 2010).

There are many reasons why Japanese university students are said to lack motivation. Intrinsic motivation for learning a foreign language fades as pressures of high school and passing university entrance exams replaces it with extrinsic motivation. Once the last exams are passed and students enter university there is no longer intrinsic or extrinsic motivation (Berwick & Ross, 1989; Sugimoto, 2010). University life has been portrayed as something of a break between the rigors of high school and the rigors of adult working life (Goodman, 2003; Sugimoto, 2010). Business and government employers tend to hire people based on the rank of the university rather than school major or student grades. Once students enter university there is no necessity to study hard and grades are largely based on attendance not on content learned (Makarova & Rogers, 2004; Sugimoto, 2010).

There are also gender differences in how education and motivation is perceived in Japan (McVeigh, 1997; Okano, 2009; Okano & Tsuchiya, 1999; Fujimura-Fanselow, 1995). This stems from traditional gender-related social norms and cultural expectations. The majority of female students in Japan study in the humanities, social sciences, home sciences and education. Gender-based expectations start from the early years of schooling and continue after university graduation in jobs or career opportunities. Education and motivation at university for female students is based on the symbolic value of higher social status and cultural sophistication rather than the practical value for males of better employment possibilities (Okano & Tsuchiya, 1999). In the context of a private women's university, Da Silva and McInerney (2008) stress the complexity and variation among motivations. They found that female students have a wide range of motivations and that many of the young women they studied were highly motivated. Thus researchers need to be sensitive to differences among students and not lump them into common stereotypes of Japanese university students.

Positive Psychology

Positive psychology began as a movement when Martin Seligman was elected president of the American Psychological Association (APA). As part of the theme to his presidency he advocated that psychology needed to refocus on a positive psychology (Seligman, 1999). This meant that in addition to a concern with helping people with psychological deficits to overcome their problems to become normal, psychologists also needed to research and understand how people could live optimally (Peterson & Seligman, 2004; Seligman, 2002, 2011; Seligman & Csikszentmihalyi, 2000). Just as some types of medical practices are used to correct for some injury or illness, for example, a doctor might treat a broken bone or a surgeon might cut out a cancerous tumor to return a person to a previous state, there are other medical practices that are concerned with optimizing a person's physical state or wellness, such as getting the right balance of nutrition or physical exercise. Psychology, Seligman claimed, was dominated by theories and research that looked at negatives, deficits that need remediation rather than positives that helped people grow and reach their full potential. Even his own work that made him a well-known psychologist that focused on learned helplessness was such an example.

Seligman (2002) proposed that positive psychology could be understood as centered around authentic happiness that was composed of three pillars: positive emotions or the pleasant life; positive character traits or the engaged life; and positive institutions or the meaningful life. These three realms can be measured empirically and thus studied scientifically. They can also be learned so that positive changes can be made in individuals. Just as with a particular mental illness a person can be made to feel normal, a person can increase their use of character strengths and increase their resilience, positive emotion, sense of meaning, and purpose (Seligman, 2002, 2009). In recent years, Seligman has shifted his focus from happiness to well-being and has revised his model to include five pillars: positive emotion, engagement, relationships, meaning, and accomplishment (PERMA; Seligman, 2011). With the shift in focus comes additional empirical research supporting his views and a stronger emphasis on education and achievement.

Since the field of psychology began, there have been psychologists and theories that shared a concern for psychological growth, for example, the psychologists Jung (1933, 1957, 1964), Maslow (1954, 1955, 1968), and Rogers (1961, 1980) and theories of humanistic psychology, but much of this work relied on cases and anecdotal evidence with few quantitative empirical studies that would make the psychological constructs more generalizable. Seligman emphasized that positive psychology should be based on scientific methods and empirical data that would make research results more replicable and cumulative. Perhaps because of this scientific emphasis, the field of positive psychology has seen the proliferation of numerous scales to measure different positive psychological constructs. In addition, constructs and scales from research predating the positive psychology movement that have been subjected to psychometrically sound research have been incorporated into positive psychology.

Positive Self-Concept

People can view themselves on various dimensions either negatively or positively. For example, a person can hold beliefs that they are physically attractive, socially adept, and of low intelligence; or a person can hold beliefs that they are of academically gifted, highly athletic, but of poor musical ability. A person can also hold general beliefs that they have a high self-regard and are well respected by others or domain-specific beliefs such as that they are great at math, or very specific beliefs such as they are a competent world-class squash player but not good at other sports.

A confusing point about self-concept is that researchers often do not specify the level of specificity and in relation to what context when using the term. One example of a self-concept is self-efficacy. Self-efficacy is a self-concept, however, as originally conceived, it is quite specific to a narrow task or activity and one that relates to one's belief that they are competent to do a task in a specific domain, for example, the belief that one can do their math homework on the multiplication tables in the double digits. Self-esteem is also a self-concept but this would be a self-concept at a global level that does not refer to any specific domain and would refer to a person's overall sense of being a competent person. As can be seen here, self-concept about one's competence can be used at three different levels of specificity but when working in specific research traditions this is often ignored and when using different terms they become even more difficult to understand how they can be applied in practice.

There are many terms and theories that are related to self-concept. For example, identity and self have many overlapping points in common, although, much of identity research is concerned with one's social identity and the many different roles and positions one can take as an individual in society. Also, there is much research in social psychology that studies self-perception biases and errors. I am using the term self-concept to refer to one's self-beliefs that might or might not be based on perceptions that are biased or erroneous in different ways.

To avoid confusion, I use term positive self-concept or positive self to refer only to the very general global self and only to refer to a positive dimension. There are many competing but related theories about what should be considered to contribute to a person's overall positive global functioning. Even though they differ in some ways they are still considered to be constructs related to positive psychology. The trait-like or personality-like nature of a positive self relates to the dispositional aspects of self, that is, there is some longitudinal stability or cross-situational consistency in behavior that can be described. Of course, these descriptions are general tendencies and not determiners of behavior, or not necessarily permanent, because they can be shaped by individual development, social relationships, and environmental factors.

As mentioned in Chapter 1, many terms that are used to express general positive self-concept and these are often related to particular researchers and their research programs, so for example: life satisfaction and subjective well-being is associated with Diener and colleagues (Diener & Biswas-Diener, 2008) and subjective happiness with Lyubomirsky (2007; Lyubomirsky & Lepper, 1999). There is much overlap and many differences with these particular theories. For the purposes of this study, except where otherwise noted, positive self or positive self-concept is used at a level of generality that is compatible with both of them.

Positive Self Model Constructs

The constructs used for modeling are all selected within the context of an academic learning environment. Learning can be considered Janus-faced, that is, facing forward or backward in time. Constructs like curiosity, hope, flourishing, interest, passion, mastery goal orientation, and self-efficacy are situated in the present but are oriented to the future. Constructs such as self-esteem, subjective happiness, positive social relationships, satisfaction in life, positive and negative trait affect are situated in the present but are oriented to the past. Ideal L2 self and intended learning effort are based on the idea of discrepancies between some future state and present state. In positive psychology the emphasis is on being authentic and true to oneself (Ryff & Singer, 2008; Schlegel & Hicks, 2011; Seligman, 2002; Sheldon, 2002), rather than an emphasis on reducing the discrepancy between present and a future ideal. Although, constructs oriented toward the past and the future correlate, the setting of a learning context with participants beginning to live life on their own determines the orientation toward the future.

Flourishing

Flourishing as a psychological construct means being mentally healthy. Flourishing, like mental illness, is a collection of symptoms, that is, some observable sign of an unobservable underlying state that persists over time. Flourishing individuals have shown the highest levels of psychosocial functioning in a number of studies (Keyes, 2002, 2007; Reschly, Huebner, Appleton, & Antaramian, 2008; Ryff & Singer, 1998).

Another perspective on flourishing comes from Seligman (2011). Seligman (2002) promoted a version of flourishing where happiness was central. In his 2011, reworking of positive psychology, he advocated for an updated model that focuses on well-being composed of five elements of: positive emotion (of which happiness and life-satisfaction are aspects), engagement, relationships, meaning, and achievement (PERMA). For Seligman (2011), the target of positive psychology is flourishing, that is, well-being through having high levels of PERMA in life. Positive emotions consist of emotions such as happiness, joy, amusement, interest, and gratitude. Engagement refers to being absorbed, and finding flow, that is, using strengths to meet challenges in life. Relationships have to with having and maintaining positive personal relationships in life. Meaning refers to having purpose in life, that is, meaning that serves something greater and beyond themself. In Japanese, this can be termed *ikigai* (生きがい) (Seligman, 2011; Mathews, 1996; Mathews & Izquierdo, 2009). Achievement concerns accomplishing things in life, to have mastery, and have competence in life.

Other theories of psychological flourishing overlap considerably with Seligman's PERMA model (2011). For example, Ryff and Singer (1998, 2008) included dimensions of: self-acceptance, purpose in life, environmental mastery, positive relationships, personal growth, and autonomy. Some theories are more parsimonious with fewer elements, for example, Deci and Ryan (2000) proposed that differing dimensions of well-being can be subsumed by three basic psychological needs of autonomy, relatedness, and competence. Other theories are more complex and include more elements, for example, Keyes (2007) proposed thirteen dimensions that can be loosely grouped into three categories: positive emotions: positive affect and avowed quality of life; positive psychological functioning: self-acceptance, personal growth, purpose in life, environmental mastery, autonomy, and positive relations with others; and positive social functioning: social acceptance, social actualization, social contribution, social coherence, and social integration. Diener et al. (2010, p. 144) created a flourishing scale, used in this study, that includes the "major aspects" of these theories.

Curiosity

Curiosity is a trait-level construct that, unlike interest in this study, is not focused on an object or skill and is distinctly different from enjoyment and happiness. That is, like other global self-concepts, the "object" is the self. It might be better to consider the self as an "object" a "subject". This makes it clear that the variables in this study are all "subjective" except for the "objective" measure of L2 proficiency. Curiosity as used in this study refers to "recognizing, embracing, and seeking out knowledge and new experiences" (Kashdan et al., 2009, p. 988). In their study on the development of a curiosity measure, Kashdan et al. found that curiosity correlated positively with various other positive measures such as openness to experience, happiness, personal growth, autonomy, positive relations with others, and purpose in life. Curious people tend look for opportunities to acquire knowledge and pursue new experiences. Curiosity helps learners to seek to fill in knowledge gaps, recognize potential learning material, and seek new learning situations thus leading to increased achievement and competence (Kashdan, 2004, 2009).

Curiosity has been shown to have positive relationships to both well-being and learning. Kashdan, Rose, and Fincham (2004) suggest that curiosity leads to personal growth through an orientation to stimuli that are: novel and challenging, rewarding, and flow-like, in addition, through assimilation or accommodation that integrates novel experiences. They found curiosity to be associated with hope and well-being. In another study Kashdan and Yuen (2007) found that when the school environment was supportive of growth and learning, higher levels of curiosity was demonstrated to be associated with higher scores on national achievement exams and school grades. Von strum, Hell, and Chammorro-Premuzic (2011) conducted a meta-analysis and found that curiosity had as much influence on academic achievement as intelligence. In brain imaging studies, curiosity has been shown to enhance learning by improving memory by consolidating new information (Kang et al., 2009). Curiosity also activates areas in the brain associated with rewards systems so that learning new information can create stimulus for further learning, that is, “prime a hunger for knowledge” (p. 971).

Hope

The hope construct is composed of elements of clearly defining goals, thinking about ways to achieve those goals, and motivating oneself to act toward goals. Hope can be characterized and measured of as a trait or state. In this study, hope is measured at the trait level.

Hope is composed of two subcomponents that act toward goals, agency or agentic thinking and pathways or pathway thinking. Agency refers to the belief that one has the ability to initiate, act, persist, and exert effort toward valued goals. It is the belief that one has volition and is in control of making progress toward goals. Sometimes agentic thinking is called willpower. Pathways refers to the belief that one can find a way or multiple ways, even in the face of obstacles, toward a goal. Sometimes pathways thinking is called *waypower*.

Hope has been shown to have effects on academic achievement in a number of studies (Lopez, 2013; Snyder, 1994, 2000). For example, Snyder, Shorey et al. (2002) found that hope predicted academic achievement in college.

Hope has been associated with well-being and learning in a number of studies. Curry Snyder, Cook, Ruby, and Rehm (1997) found that hope in college students predicted athletic performance beyond training, academic ability, and global self worth. Chang (1998) found in his sample of college students that hope had a positive influence on well-being. Ciarrochi, Heaven, and Davies (2007) tested hope, self-esteem, and attributional style for effects on academic achievement and well-being and found that hope had a strongest effect in predicting school grades in high school and was the only variable to have predictive utility across all outcome measures. Schmid, Phelps, Kiely, Napolitano, Boyd, and Lerner (2011) similarly found that hope was the best predictor of entering a trajectory of positive youth development, outperforming self-regulatory skills as predictors.

Positive Self Instrument Validation Constructs

Self-Esteem

Self-esteem is a component of self-concept that is a global evaluation of oneself. It is a value judgment and feeling of how favorably one regards past accomplishments and current skills. Rosenberg (1979) defined self-esteem as "a positive or negative attitude toward a particular object, namely, the self . . . . The individual simply feels that he is a person of worth; he respects himself for what he is" (pp. 30-31). Rosenberg (1965) developed a global self-esteem scale that is the most widely used and has seen the most psychometric and validation research done with it (Byrne, 1996; Gray-Little, Williams, & Hancock, 1997). Tafarodi and Swann (1995) claim that the Rosenberg scale is used in over a quarter of the research on self-esteem. "The appeal of the Rosenberg scale derives from both its theoretical and its practical attributes" (Gray-Little et al., 1997, p. 444). Byrne (1996) in her review of self-concept measures across the life-span notes that it measures global self-esteem or a general self-concept unlike many multidimensional models, in other words, it is true to the "unidimensional model"

(p. 15). Mruk (2006, p. 4) has called the Rosenberg scale the “gold standard” in self-esteem research and argues that it should be considered a positive psychological variable. It was included in this study because of its wide use in past studies.

Studies on the effect of global self-esteem on academic achievement are inconsistent or even negative (Baumeister, Campbell, Krueger, & Vohs, 2003). Self-esteem has been linked to many other negative effects, such as maintaining self-esteem by blaming others for personal shortcomings, egotistical thinking and behavior, self-defensiveness, narcissism, and bullying (Baumeister, Campbell, Krueger, & Vohs, 2005; Baumeister, Smart, & Boden, 1996; Crocker & Park, 2003, 2004). Marsh and Craven (2006) generally accept Baumeister et al.’s findings; however they point out that most of the studies looked at were very old, did not include multidimensional perspectives and excluded academic self-concept research. In research done by Marsh and colleagues (some of it using the same data as Baumeister) strong support was found for academic self-concept having a causal effect on academic achievement (Marsh, 1990; Marsh, 1992; Marsh & Craven, 2006). In a meta-analysis by Valentine, DuBois, and Cooper (2004) they found that effects for self-beliefs on achievement were small but when self-beliefs and achievement were matched by domain the effect sizes were larger. Similar results were found by Hansford and Hattie (1982) and Hattie (2009).

The variable effect of self-esteem can be explained by how it develops in an individual. The early studies in the 1960s were based on “found” or described self-esteem. In this case, highly competent people who were helpful towards others had high self-esteem. People who were not competent with poor social relationship skills were said to suffer low self-esteem. By the late 1970s self-esteem was established as a desired self-concept and parents and educators tried to instill it in school children whether they had developed particular competences or skills. In the 1990s researchers such as Baumeister and colleagues (2003) found that instilled self-esteem correlated with negative outcomes and not academic achievement. Students raised to have high self-esteem without being competent in a particular domain or without particular skills instead felt entitled and behaved narcissistically (Twenge, 2006, 2008; Twenge & Campbell, 2009).

What seems clear, if we can take self-esteem as an example, is that global self-concepts have weak, if any, relationships to an individual outcome set in a particular situation and time. This makes sense because global self-concept has no domain and is quite distal from a particular outcome. However, there might be middle level or mediating domain level variables that have links both to a global self and a particular task or behavior in a domain. This is logical because global self-concepts might be arrived at from competence and value from many possible domains that might or might not match up with objects in the domain under scrutiny. If an activity is linked to domain self beliefs, then, activity, domain, and self can be linked together even though the link between any particular activity and global self is weak.

Satisfaction with Life

Satisfaction with life is a subjective judgment of general life satisfaction. Pavot and Diener (1993) found that the satisfaction with life scale had good convergent validity with other measures of subjective well-being, such as self-esteem, interviewer ratings, informant reports, and the Fordyce Global Happiness Scale.

Satisfaction with life has been associated with well-being and academic achievement in a number of studies. Lyubomirsky, King, and Diener (2005) found in a large scale meta-analysis that satisfaction with life was related to positive outcomes such as satisfying personal relationships, and superior mental and physical health. In a review of the literature on life satisfaction in young people, Proctor, Linley, and Maltby (2009) found relationships with hope, coping behavior, self-efficacy, interpersonal relationships, and mental and physical health. Life satisfaction was associated with hope and satisfying personal relationships in other studies (Gilman, Dooley, & Florell, 2006; Gilman & Huebner, 2006), and it has also been identified as indicators of school adjustment and academic achievement. In a study with college students Frisch et al. (2005) found satisfaction with life was related to hope and that satisfaction with life predicted college retention. Satisfaction in life in students has also been found to have relationships with academic self-efficacy, perceived academic competence, and academic achievement (Huebner, Gilman, & Laughlin, 1999; Leung & Leung, 1992; Suldo & Huebner, 2006; Suldo, Riley, & Shaffer, 2006).

Positive Affect

Positive affect is often used as one of the components of subjective well-being (Diener, Suh, Lucas, & Smith, 1999). Positive and negative affect either have weakly positive, weakly negative, or no correlation with each other. Positive affect can be trait-like or state-like. Simply put, positive emotions allow a person to open up to new experiences and add to one's mental and social resources (Fredrickson, 1998, 2001). Trait-like positive affect has been found to precede success in multiple life domains (Lyubomirsky, King, & Diener, 2005).

Affective traits are different from affective states such as specific emotions in several ways. Affective traits differ from emotional states in that they are more diffuse, that is, they lack the intensity of states. Traits lack specific objects, whereas emotional states are about something. Affective traits last for a longer time period, emotions last for seconds while trait emotion can last years. They also lack direct physiological response, while states can rapidly prepare the body for action or communication. Finally, they are undifferentiated being basically positive or negative while there are many different emotions (Ekman, 1992; Sherer, Schorr, & Johnstone, 2001; Silvia & Warburton, 2006; Watson, 2000).

Because emotional states can quickly change, it is possible to experience both positive and negative states one after another in a relatively short period of time. For example, in a horror movie one might experience fear and then relief in a matter of seconds or minutes. Some types of positive and negative emotional states might exist together, such as, fear and happiness when on a roller coaster. Some positive emotional states cannot exist together with its polar opposite, especially extreme states such as ecstasy or panic. For example, extreme states of very happy and very unhappy (Silvia & Warburton, 2006; Watson, 2000).

Affective traits are personality-like constructs that are based on the structure of emotional experiences in terms of systems of positive affectivity and negative affectivity (Watson, 2000; Watson & Tellegen, 1985; Watson, Wiese, Vaidya, & Tellegen, 1999). Positive and negative trait affects are different dimensions, that is, they are not on the same continuum. Unlike states, where many emotions are bipolar with two different states at opposite poles, (e.g. the incompatibility of simultaneous comfort and fear), traits while generally negatively related, they can be positively related or not related at all (Diener, 2013; Diener & Emmons, 1985; Watson, 2000; Watson & Tellegen, 1985). For example, someone who is generally moderately cheerful can also be moderately nervous.

In the field of positive psychology positive affect has a number of desirable outcomes. In the absence of harmful stimulus people generally experience a mild background affect that helps them engage with the world and become involved in activities (Diener & Diener, 1996; Diener, Kanazawa, Suh, & Oishi, 2014; Lucas, Diener, Grob, Suh, & Shao, 2000). Positive affect is associated with: health, social relationships, resilience, planning, creativity, developing skills, and learning (Diener et al., 2014; Diener & Seligman, 2002; Fredrickson, 2001, 2009).

In the field of second language learning and teaching, positive affect has not been studied as much as negative affect (Arnold & Brown, 1999; Bown & White, 2010; Imai, 2010). MacIntyre and Gregersen (2012) suggest that positive affect leads to L2 learning and call for more research of the positive to balance the mostly negative affect studies in the past. Dewaele and MacIntyre (2014) found that participants in their study experienced higher levels of foreign language enjoyment than foreign language classroom anxiety and found them to be independent constructs. They also found that participants with higher levels of foreign language mastery had increased levels of foreign language enjoyment. Schumann (1997, 2001; Schumann & Wood, 2004) from a cognitive-neuroscience perspective contended that patterns of stimulus appraisal lead to positive affect and motivation that drives language learning. Positive affect serves as the basis for language learning and also memory systems that maintain knowledge and ability. Positive affect also leads to cognitions that encourage learners to approach, expend effort, and attended to potential learning stimuli; in his terms, positive appraisals drives learners to forage for new knowledge and skills (Schumann, 2001).

Negative Affect

The absence of negative affect is often used as one of the components of subjective well-being (Diener, Suh, Lucas, & Smith, 1999). Negative affect can be trait-like or state-like. Negative emotions block a person to extraneous experiences to constrain attention to the threat at hand. Negative emotions (Fredrickson, 2001) tend to protect what can be called deficiency needs (Maslow, 1943, 1955) and divert attention away from growth needs.

Negative affect can lead to adaptive or maladaptive behavior. Negative affect can signal that a problem needs to be solved or there is danger to be avoided or that some life circumstance has created a loss (Diener et al., 2014; Nesse & Ellsworth, 2009; Seligman, 2011). When negative affect is frequent and prolonged people can suffer depression and anxiety (Watson, Clark, & Carey, 1988). Negative affect has been associated with: poor physical health, low motivation to engage with daily activities, sleep deprivation, and poor social relationships (Diener et al., 2014; Beck & Koenig, 1996).

There are more studies of negative affect than positive affect in the field of second language learning and the most studied construct is anxiety (Bown & White, 2010; Dewaele & MacIntyre, 2014; Iwai, 2010). According to Dewaele and MacIntyre (2014) the first measures of language-specific anxiety were a part of Gardner's (1985, 2010) Attitude-Motivation Test Battery and the measure used the most in research is the Foreign Language Classroom Anxiety Scale developed by Horwitz, Horwitz and Cope (1986).

Foreign language anxiety (FLA) has been associated with a number of undesirable outcomes. MacIntyre and Gardner (1991, 1994) found that FLA disrupted language learning, production, and retention of new material. Foreign language anxiety can disrupt learning processes and classroom dynamics, and it also interferes with brain functioning so that memory is impaired and less learning occurs (Arnold & Brown, 1999). Studies have found that FLA has a negative affect on language learning (Dewaele, 2002; Gregersen & MacIntyre, 2014; Horwitz, 2001, 2010).

Subjective Happiness

Subjective happiness is an overall subjective assessment of a person's direct experience of being happy. Because global subjective happiness is a person's personal judgment of how they generally feel, it can be measured through self-reports. This is a more direct approach than similar concepts like subjective well-being (Diener, Suh, Lucas, & Smith, 1999) or psychological well-being (Keyes, Shmotkin, & Ryff, 2002) where multiple measures are used. Trait-like levels of happiness can also be thought of as experiences of frequent positive emotions and infrequent negative emotions. In a meta-analysis of the literature, Lyubomirsky, King, and Diener, (2005), found that happiness precedes successful outcomes across the cross-sectional, longitudinal, and experimental literature. On the other hand, frequent successful outcomes over time can lead to happiness.

Subjective happiness leads to many benefits in life. Happiness has been found to be associated with both mental health (Howell, Kern, & Lyubomirsky, 2007; Pressman & Cohen, 2005; Keyes, 2005) and physical health (Lyubomirsky, King, & Diener, 2005). Happiness has also been associated with better job performance (Wright, 2005, 2010) and work productivity (Zelenski, Murphy, & Jenkins, 2008). Happiness has been found to be associated with friendships and positive relationships (Cooper, Okamura, & Gurka, 1992; Diener & Seligman, 2002).

Positive Social Relationships

Positive social relationships reflect the degree of fulfillment of the basic need that humans have for associating with other humans. In many theories of well-being, there is a dimension that stresses the importance of positive social relationships (e.g., belongingness, Baumeister & Leary, 1995; belongingness needs, Maslow, 1943; relatedness, Deci & Ryan, 2002). In a meta-analysis of relationships of personality traits to subjective well-being (SWB) variables, DeNeve and Cooper (1998) found that extraversion and agreeableness were positively related to their SWB measure. These traits are associated with elements of sociability. Their study found that "positive affect stems primarily from our connections with others" (p. 220).

In many theories of motivation an individual's relationships with others plays an important part in the well-being of that individual. Ryff and Singer (2000) for example, noted, "Across time and settings, people everywhere have subscribed to the view that close, meaningful ties to others is an essential feature of what it means to be fully human" (p. 31). Belongingness and affection for others is featured as an important need in his theory of motivation. In self-determination theory (Deci & Ryan, 2000), the need for relatedness to others must be met as a condition for psychological growth and well-being.

Grit

The achievement of a long-term goal requires degrees of persistence and passion in pursuing the goal. Grit is constant persistence over time toward an abstract, superordinate goal, that is, a self or an identity-type goal (Gollwitzer & Kirchhof, 1998; Gollwitzer & Wicklund, 1985). Grit differs from persistent effort in this study because it is more dispositional and it lacks a specific object or skill and is not associated with one. Grit also differs from passion in this study because it is more dispositional and lacks a specific object or skill. Duckworth, Peterson, Matthews, and Kelly (2007) found that higher levels of grit are related to higher levels of academic achievement. The grit construct consists of two trait-level components of perseverance and passion for long-term goals (Duckworth et al., 2007; Duckworth & Quinn, 2009). Perseverance for long-term goals means maintaining continuity in terms of effort over years of time. Passion for long-term goals means maintaining continuity in terms of interest over years of time. Together these components of grit can be considered maintaining stamina toward long-term challenging goals.

Grit has some similarities and some differences with self-control (Duckworth & Gross, 2014). Grit and self-control are both related to aligning intentions with actions and both are related to benefits accrued through long-term effort as opposed to momentary benefits and effort. Self-control relates to more immediate or proximal goals. In a particular situation a person can have competing goals, for example, a person might have to choose between finishing homework or checking their social media. Self-control is required to avoid the temptation to disrupt doing homework. People exercise self-control over specific goals or objects. Grit relates to more abstract, distal, identity-type superordinate goals. A conflict with identity-type goals for example, might result in a person abandoning their identity as a highly ranked athlete to become a knowledgeable academic or vice versa. The superordinate identity-type goals over long periods of time relate to a person's general sense of self.

Hopelessness in Achievement

Hopelessness is a person's negative expectancy about self. Hopelessness is a factor in a variety of pathological conditions such as depression, suicide, schizophrenia, and physical illness (Beck, Weissman, Lester, & Trexler, 1974; Minkoff, Bergman, Beck, & Beck, 1973). Beck et al. (1974) developed a 20-item checklist that measured hopelessness. A Japanese version of the scale developed by Beck was created by Takahira (1998) and was adapted for this study.

Hopelessness in achievement has been attributed to a number of different causes. It is not brought about by an absence of goals but by a lack of belief that the goals will be achieved and a lack of plans to make them happen (Vincent, Boddana, & MacLeod, 2004; MacLeod & Conway, 2007; Melges & Bowlby, 1969). Negative outcomes such as past failures to achieve a goal can lead to attributions of helplessness; that is, agency and effort are noncontiguous with desired outcomes. Learned helplessness can then lead to more stable beliefs of hopelessness, which is associated with depression, passivity, and lack of motivation (Abramson, Metalsky, & Alloy, 1989; Abramson, Seligman, & Teasdale, 1978; Dweck & Wortman, 1982; Seligman, 1975).

Hopelessness of Interpersonal Relations

Relationships with other people are a basic human need (Baumeister, 2005; Baumeister & Leary, 1995; Deci & Ryan, 2002). The lack of interpersonal relationships can also contribute toward negative expectations about the future. The study in Japan by Takahira (1998) pointed out that the scale developed by Beck et al. (1974) was related to hopelessness in achievement. Takahira extended the construct of hopelessness to include a dimension of hopelessness of interpersonal relations and developed an extended scale. She found positive correlations with depression and negative correlations with self-esteem that showed high discriminant and convergent validity.

Hopelessness in relationships has been associated with a number of outcomes. A sense of hopelessness in relationships leads to avoidance, withdrawing of social support and positive affect (Miller & Roloff, 2006; Miller, Roloff, & Reznik, 2014). Hopelessness in relationships can make the individuals involved to view disputes as intractable and further relationship as irreconcilable (Pruitt & Olczak, 1995). Hopelessness in relationships has also been shown to have a negative association with interdependent happiness (Hitokoto & Uchida, 2015).

Positive L2 Self

As mentioned above, self-concept can refer to different levels of specificity. If we consider a global self-construct as a higher-level, at a lower level one might have an academic self-concept, that is, a self-concept about one's academic self. An even lower level or more specific academic self-concept would be about one's self in relation to specific academic domains, for example, a person might have a strong math self-concept but weak history self-concept. One could also have a negative domain self-concept, for example, if they were poor at math they might have a negative math self-concept. More specifically, positive and negative refer to the affective valence one has to a specific self-concept. This study proposes that there is a positive second language self-concept that I am calling a positive L2 self.

Components of positive L2 self are composed of L2 domain level dispositional constructs that are positively related to both well-being and second language learning. Constructs at this level are specific to the academic domain or academic language-learning domain, but they are more general than classroom situations and specific language skills and tasks.

For the purpose of understanding aspects of the psychology of the language learner relating to positive psychology and learning three core aspects of a positive L2 self are elaborated: an interested-in-L2 self, a harmonious passion for L2 learning, and mastery L2 goal orientation. For the purpose of instrument validation by providing convergent evidence other positive constructs that are also related to a positive L2 self are elaborated, namely: prosociality goals and an ideal L2 self.

Although prosociality goals are not explicitly linked to the L2 domain, as mentioned in relation to flourishing, positive social relationships are important for psychological well-being. Prosociality goals are theorized to have similar relationships for the academic social domain.

Positive L2 Self Model Constructs

Interested-in-L2 Self

Interested-in-L2 self is a domain-specific mid-level self-concept that can be defined as the disposition to find learning an L2 interesting and enjoyable. It is a consequence of believing that one is competent in the L2, and experiencing repeated positive experiences of discovering novel aspects of the language and successfully learning them. The interested-in-L2 self-construct differs from trait-level interest, also known as curiosity, in that trait curiosity does not necessarily have an object unless one considers the self as an "object." At this level of specificity, discovering novel aspects of the language and enjoyment exist together. This is unlike more specific interest states that do not co-occur with enjoyment. At the state-specific level, interest comes first, triggering learning, and enjoyment comes from having learned. It is only at the mid-level that interest has a domain and is diffuse enough to overlap with enjoyment and be interpreted as a unitary construct. It is also similar to the construct of flow, but interested-in-L2 self is a longer-term, more general cognitive and affective structure that produces states of flow. Hunter and Csikszentmihalyi (2003) found that for adolescents there was a strong association between interest and well-being.

Harmonious Passion for L2 Learning

The harmonious passion for L2 learning construct has similarities to but is different from interested-in-L2 self. Passion is defined as a strong inclination toward activities that are liked or loved. Where interest theory developed over time from the "bottom-up" based on decades of empirical research, passion theory was created "top-down" from self-determination theory. The model developed by Vallerand and colleagues (2003) posited two types of passions, a more self-determined harmonious passion and a more self-uncontrolled obsessive passion. Harmonious passions are associated with adaptive behaviors and obsessive passions with maladaptive behaviors. Passions differ from interests in that it has these two types, harmonious passion and obsessive passion, but also in that passions are valued and they are activities in which time and energy are spent. Interests might or might not be valued and the time and energy are unspecified. Also, as in self-determination theory, harmonious passions are developed under conditions of autonomy, positive relationships, and competence. Vallerand et al. (2007) found that harmonious passion predicted mastery goals, which, in turn led to deliberate practice and higher performance. Harmonious passion was also found to be related to subjective well-being. In this study, the passion for L2 learning construct is short for harmonious passion for L2 learning as obsessive passion is not part of this study.

Mastery L2 Goal Orientation

Mastery goals, which are also known as learning goals, are based on goal orientation theory or achievement goal theory (Dweck & Leggett, 1988; Elliot, 2005; Kaplan, Middleton, Urdan, & Midgley, 2002) and have to do with building competence. Mastery goals are defined by the purpose or orientation toward absolute gains in learning within an individual. Mastery goal orientation is also called task or learning goal orientation, and involves an orientation towards mastery of a task or learning domain (Anderman & Wolters, 2006; Maehr, 1984; Meese, Anderman, & Anderman, 2006). The focus is on learners "concerned with increasing their competence" (Dweck & Leggett, 1988, p. 256). Outcomes are measured as growth from self-comparisons of previous abilities with gained abilities. The second main type of orientation is known as performance goal orientation (also called relative, ego-involved, or competitive goal orientation), in which the focus is on demonstrating competence relative to the competence of others. Outcomes are measured as normative comparisons relative to the abilities of in the identified group, such as in a classroom or school. Performance goal orientation is manifest when "individuals are concerned with gaining favorable judgments of their competence" (p. 256). Kaplan and Maehr (1999) found that mastery goal orientations were positively related to well-being measures and academic achievement. Woodrow (2006) found that mastery goal orientations correlated with speaking proficiency as measured by a section of the IELTS.

Based on the goal orientation literature, a mastery goal orientation that is associated with self-improvement, interest, effort, learning, and self-efficacy can contribute toward a positive self. A mastery goal orientation toward learning another language is an aspect of a positive L2 self. On the other hand, a performance goal orientation would not be an aspect of a positive L2 self and is not measured in this study.

Positive L2 Self Instrument Validation Constructs

Prosociality Goals

Prosociality goals as used in this study refer to positive social relationship behavior in academic contexts, that is, more specifically, behavior oriented toward enhancing positive relationships with fellow students and classmates in the school and class settings. Although not specifically connected with L2 learning, prosociality goals can be expected to show a relationship with a positive L2 self in academic settings. Wentzel (1993) found that prosocial behavior predicted academic achievement. Positive social relationships are implicit in passion for L2 learning and in mastery goal orientation because these are theorized to be in contexts of adaptive or healthy types of relationships rather than the maladaptive types of relationships that develop in more compulsive passion or in performance goal orientations.

This special type of prosocial content goal in educational settings has been the focus of research by Wentzel into what she has called social goals (1989, 1991, 1999, 2005). Goal content refers to what students are trying to achieve (Massey, Gebhardt, Garnefski, 2008; Wentzel, 2000). A social goal references the content of having positive relationships with other students in an academic context. Wentzel (1999, 2000, 2005) has pointed out that students can have social relationship goals, for example, to gain approval from others, develop personal relationships, or cooperate with teachers or students. In addition to goals for gaining knowledge, students can have other goals that they are pursing at the same time, such as, being responsible and dependable, sharing knowledge, or making friends. These goals can also conflict, depending on, for example, the context, learning goals, or classroom climate. Academic goals and social goals can work together in a number of different ways, for instance, social goals leading to academic goals or vice versa, or they could work together so that the student achieves both social and academic goals at the same time.

Ideal L2 Self

The ideal L2 self is part of Dörnyei's (2009) theory of the L2 self system, the other parts being ought-to L2 self and L2 learning experiences. The ideal L2 self, stated Dörnyei (2009b), "is a powerful motivator to learn the L2 because of the desire to reduce the discrepancy between our actual and ideal selves" (p. 29). The ideal L2 self contrasts with ought-to L2 self , which is a type of extrinsic motivation based on beliefs about what ought to be done due to expectations of others “and to avoid possible negative outcomes” (p. 29). The ideal L2 self is not as problematic as the avoidance-focused ought-to L2 self in that it is more approach- or promotion-focused but it has some complications, in that there are two types of ideal selves, an adaptive type and a maladaptive type. In the social psychology literature, discrepancies between the ideal and actual self can also lead to depression and even suicide and discrepancies with the ought-to self can lead to anxiety; however, for Dörnyei, the ideal L2 self is something like a target. Because studies of the ideal L2 self do not follow the self-system theory as explained by Dörnyei in that they leave out the discrepancies and attempt to measure only an ideal vision, the measures used unintentionally tap into a construct closer to positive future time perspective or outcome expectancy. In this study, ideal L2 self is explicitly used as an outcome expectancy. Ryan (2008) found that ideal L2 self positively correlated with intended learning effort. The ideal L2 self is a peripheral variable used for convergent validation purposes in this study. Its inclusion also affords a deeper understanding of how variables of a positive L2 self relate to a key construct in Dörnyei’s self-system.

Math Self-Concept

Math self-concept is one of the dimensions of academic self-concept (Marsh & Craven, 2006; Marsh, Craven, & McInerney, 2008). Students develop specific subject area self-concepts such as math, history, or English and these specific subject self-concepts contribute toward an overall academic self-concept. Math self-concept is more general than math self-efficacy because math self-concept references the academic domain of math and beliefs are more general, such as the belief that one is good at math (Marsh, 1991). Math self-efficacy references specific tasks in math, such as being able to multiply single digits or being able to find the volume of a cube (Ferla, Valcke, & Cai, 2009).

The relationship of math self-concept to specific math achievement outcomes is mediated by math self-efficacy. As previously mentioned, because of the specificity matching principle (Swann, Chang-Schneider, & McClarty, 2007) a specific outcome belief is likely to have a strong relationship with a specific outcome. For example, students who believe that they can calculate the volume of a cube are more likely to be able to actually calculate cubic volume than those students who believe they cannot. Students who have a variety of math self-efficacy beliefs are likely to have higher math self-concept. Math is unrelated to the model of positive L2 self developed in this study except for the hierarchical nature of self-beliefs. Math self-concept is included here to provide divergent validity evidence for the language related items.

L2 Motivational Constructs for Model

Self-Efficacy Theory

Self-efficacy theory is a central part of Bandura's (1986, 1997, 2001) social cognitive theory of human behavior, which relates an individuals actions and cognitions with environmental influences. As Bandura (2006) stated, “People create social systems, and these systems, in turn, organize and influence people’s lives” (p. 164). As a part of social cognitive theory, competence perceptions and control beliefs are important agency components of human development and change. Self-efficacy belongs to the category of outcome expectancy beliefs, which has long been a part of motivational psychology (Atkinson, 1957, 1964), and which continue to play an important role in contemporary motivational psychology (e.g., Eccles, 1983; Eccles & Wigfield, 2002; Wigfield & Eccles, 2002). In self-efficacy theory, what is important is not the outcome per se, but the belief that one is competent to carry out the necessary processes to complete the outcome (Bandura, 1997).

Perceived self-efficacy beliefs are formed from information gathered from a variety of sources in given domains (Bandura, 1997). These sources include enactive mastery experiences that provide feedback about one’s capabilities, vicarious experiences that provide comparative feedback of the attainment of others, verbal persuasion and social influences that one has certain capabilities, and physiological symptoms and affective states that provide feedback that one is capable to do a task. Once formed these self-efficacy beliefs

> influence the courses of action people choose to pursue, how much effort they put forth in given endeavors, how long they will persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-hindering or self-aiding, how much stress and depression

they experience in coping with taxing environmental demands, and the level of accomplishments they realize. (p. 3)

Self-efficacy has been researched in thousands of studies and pre-dates the positive psychology movement but it can be now considered as a construct in positive psychology (Bandura, 1977, 1997, 2008; Maddux, 2002). It is used as a motivational variable in this study because self-efficacy's proximity to motivated behavior and achievement. L2 self-efficacy has been shown to have a relationship with L2 achievement and other motivational variables (Mills, 2014). Pajares (2001) found that self-efficacy was positively related to positive psychology variables of optimism, authenticity, self-inviting, and other-inviting. Inviting refers to the messages people send to themselves and others which has a role in developing the beliefs people have of themselves. Self and other inviting is a process where people develop their own potential and the potential of others.

Speaking Self-Efficacy

Self-efficacy can be task or domain-specific, that is, it can refer to a particular task that is immediately present or a particular academic domain. When self-efficacy is more general in nature, it becomes similar to the construct of confidence. Speaking self-efficacy as used here refers to being capable of successfully engaging in speaking performances at different levels of difficulty.

Burrows (2009) created a self-efficacy scale that was designed to be used by students that studied abroad. Although the scale was created for a future study, he provided an example of a completed instrument. Burrows (2009) noted that "research in self-efficacy has been relatively limited in the foreign language field" (p. 9).

Mills (2009) created scales that were aligned with the American Council on the Teaching of Foreign Languages (ACTFL) Standards for Foreign Language Learning. These standards were grouped into five goal areas of: Communication, Cultures, Connections, Comparisons, and Communities. Included in her scales were items that addressed L2 speaking self-efficacy. In her study of project-based learning (PBL), students showed gains in items of L2 self-efficacy.

Listening Self-Efficacy

Listening self-efficacy as used here refers to the belief in being capable of successfully listening and understanding at different levels to different sources of spoken language. Mills, Pajares, and Herron (2006) found that listening self-efficacy was associated with listening proficiency in the female participants of their study. Mills (2009) found that a curriculum based on PBL improved L2 listening self-efficacy as measured by items on L2 listening self-efficacy.

Graham (2006, 2007) investigated the association between L2 listening strategies and L2 listening self-efficacy. She found that low ability students had attributional beliefs that led to poor strategies and diminished self-efficacy. Graham (2007) found that listening strategy training improved student listening performance and L2 listening self-efficacy. Graham and Macaro (2008) did a similar study that also found that listening strategy training increased listening proficiency and L2 listening self-efficacy with stronger effects on student ability to comprehend details and opinions.

Reading Self-Efficacy

Reading self-efficacy as used here refers to the belief in being capable of successfully reading and understanding written texts at different levels for sources differing in levels of difficulty. Mills, Pajares, and Herron (2006) found that reading self-efficacy was positively associated with reading proficiency.

Lake (2014) found that students who read extensively with easy graded readers gained L2 reading self-efficacy while those who used graded readers but did not read extensively showed no gains. In addition, gains in L2 reading self-efficacy were shown to have a relationship with gains in a positive reading self as measured by an L2 reading interest measure.

L2 Motivational Constructs for Instrument Validation

Intended Learning Effort

Intended learning effort is a construct used by Ryan (2008, 2009) that was related to intentions of student efforts to learn and possible intended future efforts both inside and outside the classroom. Ryan (2008, 2009) found a high correlation between an ideal L2 self scale and an intended learning effort scale.

In past ideal L2 self research, intended learning effort has often been used as an outcome variable but there are some problems with using it this way. Both ideal L2 self and intended learning effort are used to describe some desired future so it is understandable they correlate well with each other. However, future intended learning effort, as are many intentions, is often far removed from any course of action. The emphasis in positive psychology is on being authentic and true to oneself in the present (Ryff & Singer, 2008; Seligman, 2002; Sheldon, 2002) rather than intentions or ideals in the future. In addition, future intended learning effort is easy to confound with wishful thinking. Positive psychology also "does not rely on wishful thinking, faith, self-deception, fads, or hand waving" (Seligman & Csikszentmihalyi, 2000, p. 7). Rather than focusing on future ideals and intentions, motivation is stronger when framed in terms of the present, that is, the "journey" not the "destination."

Persistent Effort at L2 Learning

Persistent effort at L2 learning is defined as the amount of time and frequency one spends studying the L2 and persisting in the face of obstacles or difficulties. Rather than referring to future intended effort, this construct relates to beliefs about present actual persistent effort. Persistent effort in the form of deliberate practice has been found to be positively related to levels of expertise (Ericsson, Krampe, & Tesch-Romer, 1993).

Persistent effort is a feature of many theories of achievement. Dweck and colleagues (e.g., Dweck & Leggett, 1988, Dweck, 2000, 2006) have linked implicit incremental theories of intelligence to learning goal orientations and persistent effort. Novices can become experts through years of persistent effort in theories of expertise development (Ericsson & Charness, 1994; Ericsson, Charness, Feltovich, & Hoffman, 2006; Ericsson, Krampe, & Tesch-Romer, 1993). Persistent effort of practicing language skills leads to procedural knowledge and automatization (DeKeyser, 2007). Persistent effort is also a part of grit theory. While grit is a broad construct, persistent effort in specialized domains are necessary for mastery to develop (Duckworth & Gross, 2014). Persistent effort of practicing language skills leads to procedural knowledge, automatization, and fluency (DeKeyser, 2007).

Summary of Constructs

To summarize and to see the relationships among the constructs in this study see Table 2. Constructs are grouped into three categories: Positive self related constructs, which are general and have no specific domain; positive L2 self related constructs, which, within the domain of second language learning, can be considered a middle level between the more general self constructs and the more specific task or activity related constructs; and L2 motivation related constructs, which are both domain-specific and related to specific tasks or activities.

Table 2. *Categories of Constructs Used in this Study by Level*

Modeled constructs	Positive relationships (Convergent validity)	Negative or no relationships (Divergent validity)
Positive Self: flourishing, curiosity, hope	Self-esteem, satisfaction in life, positive affect, subjective happiness, positive feeling, negative feeling, positive social relationships, grit	Negative affect, negative feeling, hopelessness in achievement, hopelessness in relationships
Positive L2 Self: interested-in-L2 self, harmonious passion for L2 learning, mastery L2 goal orientation	Prosociality goals, ideal L2 self	Math self-concept
L2 Motivational Constructs: speaking self-efficacy, listening self-efficacy, reading self-efficacy	Intended learning effort, persistent effort at L2 learning	

Structural Relations

The relationship between global self-concepts and academic achievement and the causal ordering are reviewed in the next section. Structural relationships are conceptualized as being a part of frameworks in many areas of psychology. Structural relations are reviewed in the following sections: self research, social psychology and personality theory perspectives, goal setting theory, and achievement goal theory.

Self Research

There has been some criticism of the global self-esteem construct in the last few years that the relationship to important outcome variables such as academic achievement are very weak or nonexistent (Baumeister, Campbell, Krueger, &

Vohs, 2003; Crocker & Park, 2004). In relation to global self-esteem, Baumeister, et al. (2003) claimed that "with the exception of the link to happiness, most of the effects are weak to modest. Self-esteem is thus not a major predictor or cause of almost anything (again, with the possible exception of happiness)" (p. 37). They further claimed that self-esteem might lead to antisocial behavior such as violence and delinquency. Self-esteem has been linked to many other negative effects, such as maintaining it by blaming others for personal shortcomings, egotistical thinking and behavior, self-defensiveness, narcissism, and bullying (Baumeister, et al., 2003, 2005; Baumeister, Smart, & Boden, 1996; Crocker & Park, 2003, 2004). Marsh and Craven (2006) generally accept Baumeister et al.'s (2003) findings; however they pointed out that most of the studies looked at were very old, did not include multidimensional perspectives and excluded academic self-concept research. Under these conditions it seems reasonable that correlations would be weak at best with academic performance. Research done by Marsh and colleagues, some of it using the same data as Baumeister, found strong support for academic self-concept having a causal effect on academic achievement (Marsh, 1990; Marsh, 1992; Marsh & Craven, 2006). In a meta-analysis, Valentine, DuBois, and Cooper (2004) found that effects for self-beliefs on achievement were small but when self-beliefs and achievement were matched by domain the effect sizes were larger. Similar results were found by Hansford and Hattie (1982) and Hattie (2009).

Crocker and colleagues (Crocker & Park, 2003, 2004; Crocker & Wolfe, 2001) have argued that researchers need to consider contingent self-esteem, that is, domains that a person has staked self-esteem on "so that person's view of his or her value or worth depends on perceived successes or failures or adherence to self-standards in that domain" (Crocker & Wolfe, 2001, p. 594). Most studies focus on level of self-esteem but not on why self-esteem is being pursued. In cases of unstable self-esteem or a desire for positive evaluations by others, pursuing self-esteem might be harmful.

In an article that took a broad look at self-concept and some of the criticisms of it, Swann, Chang-Schneider, and McClarty (2007) argued that self-concepts and outcomes need some sort of contextualization. They point out that broad attitudes and traits were critiqued decades ago for not predicting specific behavior. Now, attitude and trait researchers still use these broad personality constructs such as attitudes and traits with the understanding that there might be many mediator or moderator variables in between the broad construct and any particular behavior. Also, it is now understood that the strength, accessibility, and relevance of the attitudes and traits have an effect on how they predict behavior. In the case of self-views, meta-cognitive aspects such as strength of the self-view can bolster predictive validity. Strength of self-view might be indicated by importance, certainty, clarity, extremity, accessibility, temporal stability, or goal-relatedness.

Swann et al. (2007) also pointed out that attitude and trait researchers use the *specificity matching* principle that "holds that the specificity of predictors and criteria should be matched" (p. 87). Any particular outcome can be caused by a multitude of factors and the further away a predictor variable becomes the less predictive it is. To show relationships among variables, researchers need to consider the specificity among the variables with the understanding that there are stronger relationships with variables of similar levels of specificity and less of a relationship as the specificity differs. For example, global self-esteem, academic self-concept, and math self-concept should have increasing relationships with an outcome such as math achievement. To sum up the research on the relationship of global self-concept to specific academic achievement it seems fair to say that research shows that there is little or no relationship. However, if a middle level construct or intermediate variable between the global self-concept and specific academic achievement is taken into account then there is a strong relationship between the middle level domain variable, for example, academic self-concept and specific academic achievement.

This study is partly based on a theoretical perspective drawn from positive self-concepts (Bracken, 1996, 2009; Bracken & Lambrecht, 2003). In a theoretical perspective developed by Bracken, people get input about a variety of domain-specific self-concepts through personal evaluations or through the evaluations of others. Self-concepts are hierarchically structured with global self-concept encompassing all aspects of the self. Multiple domain-specific self-concepts can be considered multiple dimensions of the self. For Bracken, these multiple dimensions are academic, social, affect, competence, physical, and family. Any combination of these dimensions might contribute to the global self-concept. Bracken (2009) points out that for interventions it is easier to target "individual context-specific domains for change than it is to attempt to improve a person's global or overall self-concept" (p. 100).

Self-concept researchers have taken different positions on the directionality of the relationship of more general variables such as academic self-concept and academic achievement, that is, does academic achievement cause academic self-concept in a bottom up approach or does academic self-concept cause academic achievement in a top down approach (Marsh & Craven, 2006). A self-enhancement model predicts that the direction is from self-concept to achievement whereas a skill-development model predicts that the direction is from achievement to self-concept. Marsh and colleagues (Marsh & Craven, 2006; Marsh & Martin, 2011) favor a reciprocal effects model where the directionality is in both directions. In other words, academic achievement affects self-concept and self-concept affects achievement.

Another consideration when taking directionality into account might be age of development, that is, at younger ages many self-views might not have been developed so that directionality flows from developing competence toward self-views. Also, there might be differing strengths of directionality associated with age-related developments so when directionality flows in both directions, the relative strengths can change with age. In addition, there might be differing rates of change at different developmental stages. For example, math self-concept might develop faster for an elementary school student than for an adult or history self-concept might produce greater academic achievement in a high-school student than a junior high school student. Research by Harter (1981, 2012) on how the self is constructed links development of cognitive and motivational structures to messages received by children in their educational and social environment. These structures then might serve as motivational orientations or goals in more mature or future situations.

Social Psychology and Personality Theory Perspectives

As mentioned previously, positive L2 self is a domain-specific level of specificity, as opposed to more general higher level self or even more specific state level self, that is centered on learning English in an academic context. In the proposed model, the positive L2 self is an intervening or mediating construct between an independent construct of positive self and a dependent or outcome construct of positive English learning motivation. This domain-specific level can be considered a middle level of abstraction because, in the words of Mischel and Morf (2003), “it is at this middle level that trait prototypes involve specific, contextualized representations of the self and others” (p. 35). For this study, positive L2 self is a middle level between the more general self and more specific, less stable motivational variables. This middle level has been called personal strivings (Emmons, 1986, 1999), personal projects (Little, 1983; Little, Salmela-Aro, & Phillips, 2007; Romero, Villar, Luengo, & Gomez-Fraguela, 2009), and personal goals (Salmela-Aro, Aunola, & Nurmi, 2007; Sheldon, 2002, 2004; Sheldon & Elliot, 1999). Other concepts include current concerns (Klinger, 1975), identity goals (Gollwitzer, 1987), life goals (Nurmi, 1991, 1992), developmental goals (Heckhausen, 1999; Heckhausen & Heckhausen, 2008). Again, when looked at cross-sectionally the direction of causality is from the general trait level to the more stable domain disposition to the more dynamic and situational level of classroom context, learning processes, and specific tasks.

Goal-Setting Theory

Goal-setting theory was developed over about four decades based on work done by Locke and Latham (Locke, 1968; Locke & Latham, 1990, 2006, 2013). They are organizational psychologists who primarily do research on motivation and performance in organizational or work-related tasks (Latham, 2007). Their work on goal-setting theory "focused on the relationship between conscious performance goals and level of task performance" (Locke & Latham, 2002, p. 705). Specificity, difficulty, and proximity are important in goal setting theory and the attainment of achievement (Locke & Latham, 1990, 2002, 2006). More specific and proximal goals reduce ambiguity in cognizing achievement strategies and constrain the type of behaviors needed for successful outcomes thus helping to more precisely regulate performance. Also, attention, energy and direction can be more focused on the task and distractions can be more easily identified so that they can be avoided or overcome. Specific goals help in understanding and measuring goal progress and sustaining effort to completion rather than a slow winding down because of uncertainty over whether enough progress has been made. Clearer goals help in knowing when a goal is attained and that is time to move on to the next goal. More challenging goals lead to greater performance outcomes and require more effort, energy, frequency of activities, and persistence. Easy goals might not even generate enough energy to overcome initial inertia to engage action, difficult but attainable goals make involvement more worthwhile. Being able to make progress on an attainable goal generates commitment. Additionally, the degree to which goals are self-set might influence the importance one places in the goal.

According to Locke and Latham (2002) goals also indirectly affect performance "by leading to the arousal, discovery, and/or use of task-relevant knowledge and strategies" (p. 707). Knowledge and strategies that are relevant are used automatically to achieve the goal. If automatized skills do not exist, then people borrow from previously used skills that were used in related contexts to use in the new context. Those with higher self-efficacy develop more effective strategies than people with lower self-efficacy (Locke & Latham, 2002, 2006).

Zimmerman (2008) has directly linked goal-setting to self-regulation in his three phase model of self-regulation. For Zimmerman, self-regulation can be broken down into different phases: forethought, performance, and self-reflection. In the forethought phase, before effort has been expended, two categories are important: task analysis and self-motivation beliefs. As part of task analysis, goal setting is one of the key components of self-regulation. In the later performance phase, goal setting influences performance through processes of self-control and self-observation. Self-control methods include the use of task strategies, attention focusing, and self-instruction. Self-observation methods include metacognitive monitoring and self-recording. In the self-reflection phase, self-judgment and self-reaction are important processes. Goal setting leads to better self-evaluation because then standards exist to evaluate progress. Goal setting might lead to better causal attributions of needed effort and persistence to reach a goal. Attributions are more likely to be made about controllable causes such as effort or use of learning strategies. For self-reaction, goal setting influences self-satisfaction and adaptive inferences. Students link self-satisfaction and positive effect with progress toward learning goals and skills attained. Goal setting makes for better adaptive inferences toward future goals rather than negative defensive inferences to protect them from future disappointment. Finally, goal setting aids in self-reflections that better prepare them for subsequent forethought phases.

Achievement Goal Theory

Achievement goal theory represents an integrated pattern of beliefs one has when pursuing a goal or goal orientations (Pintrich, 2000a, 2000b). It is “currently one of the most active areas of research on student motivation in academic settings” (Pintrich, Conley, & Kempler, 2003, p. 319).In contrast to goal-setting theories and goal content theories, as mentioned above, where the focus is the goal or how to self-regulate for completion, achievement goal theory is concerned with what are called goal orientation theories. That is, with goal orientations the focus is not only on the content (what the goal is) or method or setting (how to get the goal) but also on the meaning or purpose (why achievement goal behavior is being engaged) and how it is perceived and evaluated (Ames & Archer, 1988; Dweck, 1992; Dweck & Leggett, 1988; Kaplan et al., 2002). Much of the research done in this area has been in school settings and has been concerned with different patterns of motivation, the quality of student learning, and levels of achievement.

There are two main orientations studied in achievement goal theory; the terms used here are mastery goal orientation and performance goal orientation. Mastery goals have also been called task-involved, learning, or task-focused goals, but with roughly the same meaning. Similarly, performance goals have been called ego-involved or ability-focused goals, with roughly the same meaning (Ames & Archer, 1988; Dweck & Leggitt, 1988; Maehr & Midgley, 1991; Nicholls, 1984). Mastery goals, like the term implies, have to do with mastery learning, that is, improving oneself from one point in time to another point in time where learning material is mastered. The person seeks to make progress or grow in competence or ability. The improvement is in relation to a previous level of reference and learning is what has been achieved or gained in reference to one's self. In measurement theory this would be called an absolute learning orientation. Performance goals have to do with demonstrating or displaying that one has competence or ability relative to others. The person seeks to show that they are better than others or avoid the appearance of failure. Performance does not depend on what one could do before but is in relation to how others perform. In measurement theory this has been called a relative ability orientation.

These different goal orientations show how many different cognitions and behaviors can be related. The orientations have both more temporary situational component and more enduring personal component (Kaplan & Maehr, 2007; Kaplan et al., 2002). Students can enter an academic setting with orientations that are partly based on their particular configurations of goal contents or differing importance or value they place on goals at different levels of goal hierarchies. Also, they are influenced by differing socialization processes in their personal histories. Their current academic situation, school, department, and classrooms also have an influence and are predictive of the types of personal goal orientations adopted (Anderman & Wolters, 2006). If the different organizational levels are not congruent they might send mixed signals to students. For example, if the school sends messages that encourage a performance orientation, the department sends messages that encourage learning orientations, and teachers in different classrooms send mixed messages, then the situational component can be confusing. Teachers also vary in their goal orientations and send messages through their teaching practices about their own preferences for goal orientation (Perry, Turner, & Meyer, 2006).

One of the earliest achievement goal theorists was Carol Dweck (e.g., Dweck & Leggett, 1988). In her work, she has linked general self-theories with more specific achievement goal theories. Specifically, she proposes that people hold entity theories or incremental theories about intelligence and ability that in turn affect more specific goal orientations and achievement motivation. Implicit incremental theories and the associated learning goal orientations are also called growth mindsets. In the field of second language learning this has been explored most fully by Sarah Mercer and colleagues (Mercer, 2011a, 2011b, 2012a; Mercer & Ryan, 2010). Much of her research was based on case studies or interviews so it is difficult to generalize her findings, although they tend to support quantitative findings in educational psychology. Woodrow (2006, 2008, 2012), has done qualitative studies that has looked at achievement goal orientations and found that mastery goal orientation (in her study called task goal orientation) correlated with achievement as measured by speaking proficiency as measured by the IELTS while performance goal orientations did not.

From the above perspectives of goal hierarchies and goal setting theories people can have goals that range from higher-level goals that are more abstract and distal to the lower level goals that are concrete and proximal to behavior. In a top-down process that is similar in manner to the way that traits affect states in personality theory, more stable higher-level goals form systems of goals that cascade into more dynamic lower level goals that direct behavior. In a bottom-up process, the feedback provided by behavior might then affect the lower level goals that might then, over time, work back up into higher-level goals. That is, behavior that results in repeated successes on specific goals might encourage an individual to attempt a broader goal in the same domain. For example, if a student successfully understands simple sentences in the L2, they might try to understand a short passage. Over brief amounts of time, the direction is from the higher-level goals to the lower level goals, although over longer periods or developmentally formative stages of time the direction might also be from the lower to higher-levels of goals.

Section Summary

To summarize this section on structural relations, various frameworks in psychology, namely self research or research on the self, social psychology and personality theory perspectives, goal setting theory, and achievement goal theory illustrate the need to place specific variables into proper structural relationships. Examples included: global to academic self-concept to domain-specific levels of specificity in self research, such as global self-esteem to academic self-concept to self-efficacy of a specific skill; broad personality traits to middle level traits in social psychology and personality theory, such as personality traits to personal goals, strivings, and projects; distal to proximal goals in goal-setting theory; and self-beliefs or mindsets to goal orientations to achievement in achievement goal theory, such as having incremental theories or a growth mindset to a learning goal orientation to specific academic achievement.

To summarize the structural relations research in a diagram, Figure 1 shows structural relationships in an abstract general case. For the structural relationships in this study, Figures 2 and 3 show the theoretical relationships of the latent variables.

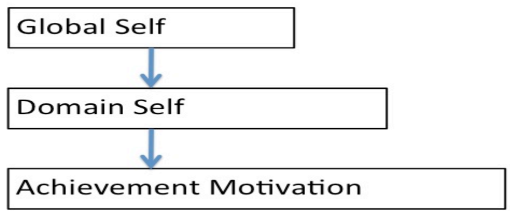

Figure 1. General relationships among levels of self and motivation.

Figure 1 shows the most abstract case where the structural relations are directed from the global, to domain, to particular motivations. The global level is general in that it relates to the whole self and is relatively stable and trait-like. The middle level is less general and relates to a particular domain or individual interest in life. The domain level relates to a relatively enduring disposition in a particular field or sphere of activity. The achievement motivation level is highly specific to particular tasks.

The content and structural relations in this most abstract case would include positive and negative dimensions at all three levels. For an example of a negative dimensions, a generally anxious person would have as part of their global self-concept, some degree of trait anxiety. Another person who is not generally anxious, can have anxiety in a specific domain, for example, learning an L2. A person could be lacking in achievement motivation for a certain task for example, by becoming anxious from the thought of reading a book in the L2.

Figure 2 shows an application of the more abstract case of Figure 1. In Figure 2, the structural relations are similarly directed from global, to domain, to particular motivations but the content is more specific because only the positive dimensions are modeled and a particular domain, L2, is specified.

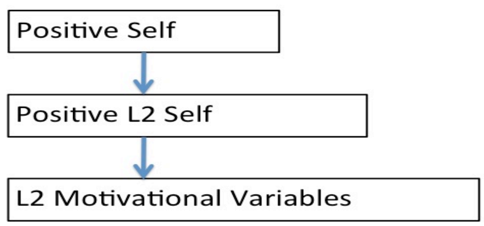

Figure 2. Relationships among levels of positive selves and motivation.

Figure 3 shows an application of the more abstract case of Figure 1 with a substitution of L2 proficiency. In this study, self-report variables were used to create the modeled latent variables. One potential objection for the use of self-reports at all levels is that there might be a method effect where the variables share variance due to the similar method of measurement. To counter this potential objection, an objective L2 measure of proficiency was introduced to substitute for the self-reported motivational variables.

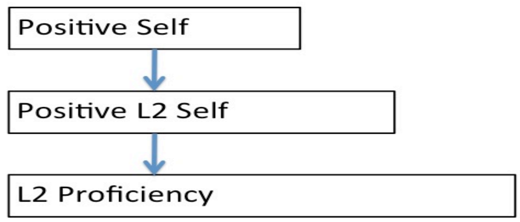

Figure 3. Relationships among levels of positive selves and L2 proficiency.

L2 Self and Motivation

In the early 1990s, there were calls to expand the research agenda that was originally set by the theory and research work done by Lambert and Gardner from the 1950s and 1960s (Crookes & Schmidt, 1991; Oxford & Shearin, 1994). Dörnyei has done more than anyone in answering this call. While originally not questioning the importance of the work done by Lambert and Gardner, he has built on the more generalizable aspects of their work and created and introduced to the field of second language learning his own motivational self system that is much more comprehensive and takes into account various theories that have developed in cognitive and educational psychology.

The L2 domain level variable that has been developed by Dörnyei (2005, 2009b) and colleagues by building on the concept of *possible selves* (Markus & Nurius, 1986) and *imagined community* (Norton, 2001) conceptualizes an *ideal language self* (Dörnyei, 2005; Ryan, 2005). This idea is further developed in recent work centered on identities and the L2 motivational self-system (Dörnyei &

Ushioda, 2009). The ideal L2 self, states Dörnyei (2009b) "is a powerful motivator to learn the L2 because of the desire to reduce the discrepancy between our actual and ideal selves" (p. 29). Other components of the self-system include an ought-to L2 self that involves more extrinsic motivation and L2 learning experience that involves elements of interactions within a learning situation.

One potential problem with the L2 motivational self-system is the prominent role that ought-to L2 self plays. In educational psychology, beliefs and behavior that one ought to meet extrinsic expectations and avoid negative outcomes is associated with weak motivation, short-term effects, and poor outcomes. Avoidance motivation is often related to attributing poor outcomes to forces beyond individual control, defensive strategies to lower the negativity of the outcome, self-handicapping, and avoidance of help-seeking behavior (Elliot, 2008; Elliot & Dweck, 2005; Kaplan et al., 2002; Urdan, Ryan, Anderman, & Gheen, 2002). Another problem is that as in self-discrepancy theory, the ought-to self has been linked in psychological studies to depression, anxiety, distress and mental disorders (Cornette, Strauman, Abramson, & Busch, 2009; Higgins, Klein, & Strauman, 1985; Strauman & Higgins, 1988).

The ideal L2 self is also problematic. Although the ideal L2 self is an approach rather than avoidance focused type of motivation also known as promotion-focused motivation, it still has some complications. Constructs of ideals can be similar constructs of perfectionism. Constructs of ideals then are problematic in the same ways as constructs of perfectionism, in that they can manifest as either adaptive types or maladaptive types (Blatt, 1995; Hewitt & Flett, 1991; Nugent, 2000). For example, maladaptive perfectionists have high ideals that are often difficult to meet leading to dysphoria. Maladaptive perfectionism is related to being afraid of making mistakes, hesitations in making decisions and acting on them, and a host of pathological and mental health problems.

As these examples show, some constructs have a positive or adaptive type and a negative or maladaptive type. In the Japanese educational context where students are often represented as being reticent, afraid of making mistakes, or generally demotivated, then it might be that for some students the discrepancy between ideal and actual self is so great that it does not lead to motivation but withdrawal. In discrepancy theory, for some people the ought-to self causes anxiety and ideal self causes depression (Higgins, 1987, 2012) and these can lead to emotional and social withdrawal or thoughts of suicide. In Japan, the phenomenon of *hikikomori* or self-seclusion due to anxiety is recognized as a growing problem (Furlong, 2008; Teo, 2010). Suicide and attempted suicide rates, due in part to depression, in Japan are among the highest in the world and have remained a problem in recent years (Hidaka, Operario, Takenaka, Omori, Ichikawa, & Shirasaka, 2008; Nakao & Takeuchi, 2006).

Researchers and teachers need to be aware of possible negative effects of discrepancies between present and future selves. For researchers, in the context of Japan, if for some students an ideal L2 self or ought-to L2 self is maladaptive then this might get conflated with an adaptive type and this suggests that it needs to be studied more before promoting it without qualification. Teachers need to be aware that self-discrepancies can in some situations be maladaptive and they need to be careful when advising students.

Another possibility in relation to measures of ideal L2 self is that the items are not being interpreted as “ideal” but as something that does not create maladaptive-tendencies; that is, the measure might to some extent be mislabeled. In other words, items used on ideal self instruments are not measuring discrepancies but vague goals. In this case of mislabeling, a future looking self-guide such as an ideal L2 self might have a relationship with a positive L2 self. Dörnyei’s terminology is used here with the understanding that “ideal” might not be the best label, that is, as operationalized in the ideal L2 self scale might underrepresent the ideal construct. The ideal L2 self might represent a type of goal or expectancy outcome.

Rather than an ideal future L2 self, this study elaborates a present-time oriented positive L2 self and a strong contributor to this construct is an interested-in-L2 self. Domain level interest is central to the construct of a positive L2 self because the psychological construct of interest has been demonstrated to be important in positive psychology (e.g., Fredrickson, 2001; Kashdan & Silvia, 2009) and learning theories (e.g., Dewey, 1913; Hidi, 1990; Hidi & Renninger, 2006; Silvia, 2006). In addition, the psychological construct of interest is a component of many theories of motivation. Some examples include: contemporary expectancy-value theories of motivation such as that developed by Eccles and her colleagues (e.g., Eccles & Wigfield, 2002; Wigfield & Eccles, 2002), especially as part of the value construct; social-cognitive theory and self-efficacy theory (e.g., Bandura, 1986, 1997), especially as part of internal personal determinants of behavior and self-efficacy; flow theory and autotelic activities (e.g., Csikszentmihalyi, 1990; Egbert, 2003), especially as a part of individual engagement (e.g., Hunter & Csikszentmihalyi, 2003; Shernoff, 2013; Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003); goal orientation theory (e.g., Dweck & Leggett, 1988; Pintrich, 2000a, 2000b), especially as a part of mastery beliefs or learning goal orientations; and self-determination theory (e.g., Deci & Ryan, 1985, 2002; Ryan & Deci, 2000), especially as related to intrinsic motivation. Despite the central importance of interest to theories of motivation, it has been understudied in L2 learning and in the Japanese context.

Section Summary

The constructs at the various levels and the structure to be tested in this study can be summarized in the schematic shown in Figure 4. Central to the model is a middle-level construct of a positive L2 self, with a higher-level positive self, and lower level motivational construct. The model assumes that a positive L2 self partially mediates the relationship between a positive self and L2 motivation. In other words this study investigates relationships of positive self to both a positive L2 self and L2 motivation. The three latent variables are all related to a positive self-concept, albeit at differing levels of specificity. Positive self and L2 motivation differs the most in terms of specificity, so this relationship should be weaker.

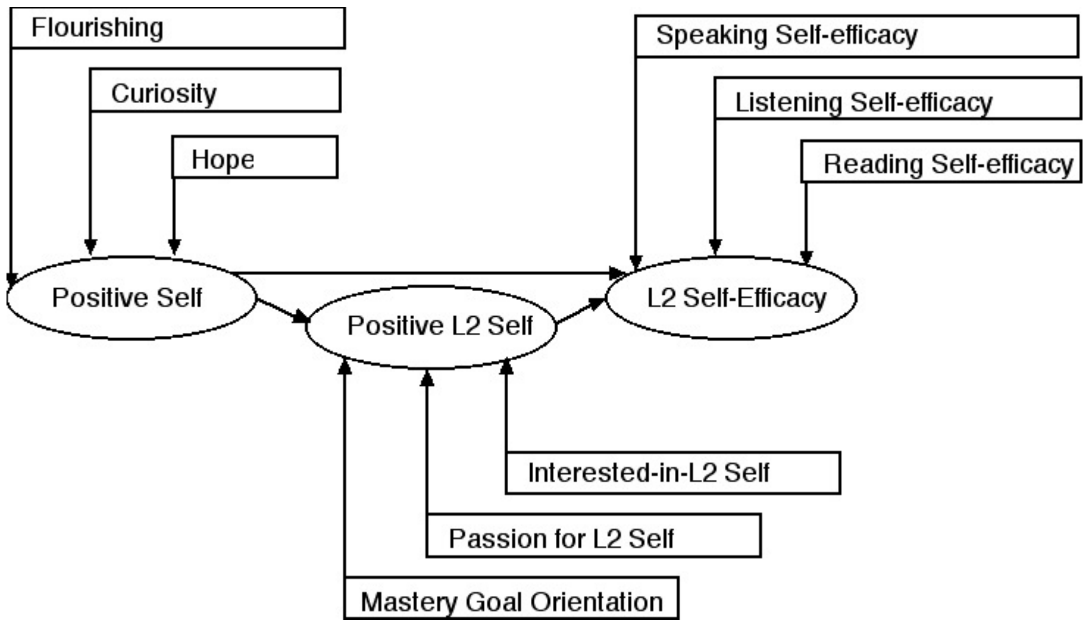

Figure 4. Model of positive selves and motivation.

Also tested is an alternative model of positive selves and motivation. The model of a middle-level construct of a positive L2 self, with a higher-level positive self, and lower level motivational construct that fully mediates the relationship between positive self and L2 motivation. The schematic in Figure 5 summarizes these relationships with one less path.

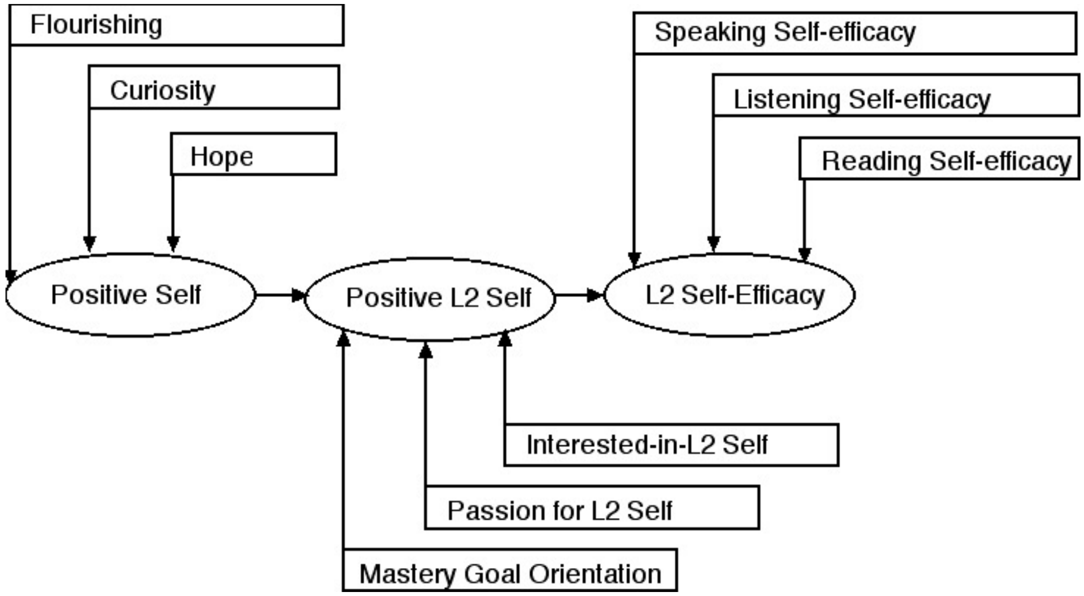

Figure 5. Alternative model of positive selves and motivation.

Putting together the positive self constructs at the two levels and L2 proficiency the structure to be tested can be summarized in the schematic shown in Figure 6. Again, central to the model is a middle-level construct of a positive L2 self, with a higher-level positive self, and a more unstable outcome variable of L2 proficiency. Positive L2 self mediates the relationship between positive self and L2 language proficiency as suggested by the literature. Positive self can be developed by being competent in many different domains so there is no reason for a strong relationship to L2 proficiency beyond a relationship to a positive L2 self.

There are two different L2 proficiency measures used in this study as part of models, TOEIC and TOEIC Bridge. For the models with proficiency measures the number of participants is much less because they are split. However, this allows for a cross-validation study. If equivalence is found between the models with two different proficiency measures, this strengths the validity evidence for the models.

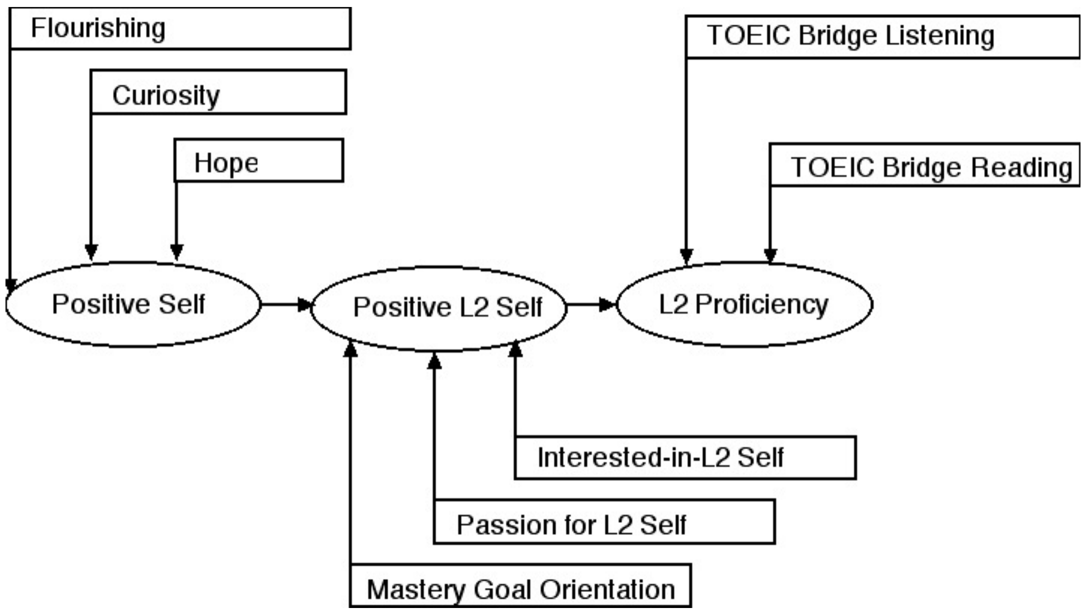

Figure 6. Model of positive selves and L2 proficiency.

This study also investigates an alternative model if the relationship to L2 proficiency is partial, that is, if positive self shows relationships to both a positive L2 self and L2 proficiency. The schematic in Figure 7 summarizes these relationships.

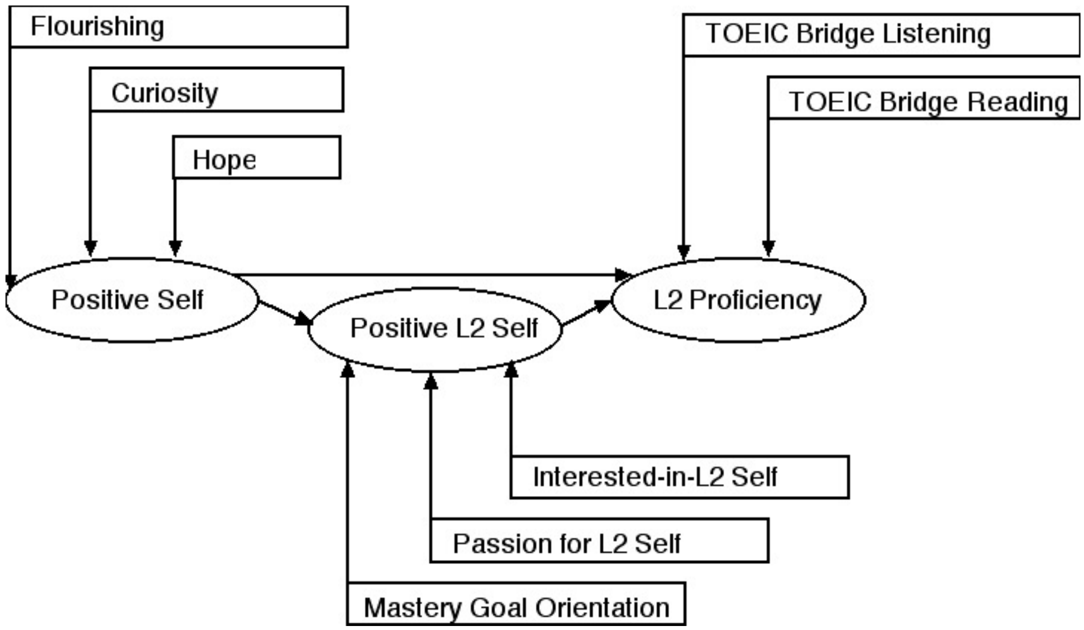

Figure 7. Alternative model of positive selves and L2 proficiency.

Gaps in the Literature

The general gap is that no empirical research has attempted to link global or domain level constructs from positive psychology to L2 motivation or L2 achievement. In the field of second language studies some recent research on self-concept has been done by Mercer (2011a, 2011b, 2012); although, this research is largely descriptive or done through qualitative explorations using case study or interviews. Dörnyei and colleagues (Dörnyei & Ushioda, 2009, 2011) have elaborated an L2 self-system guided by a future self and empirical work has been done in various contexts (Csizer & Kormos, 2009; Ryan, 2009; Taguchi et al., 2009), however this research does not take the present or past self into account and is driven in large part by social and cultural factors.

The first specific gap addressed in this study is the lack of a composite construct at a global level of positive self-concept in an academic-learning context. Individual constructs have been created in the field of positive psychology that are future-oriented and are related to learning or acquiring novelty. This study is the first to examine such a composite construct and relate it to a positive domain level construct.

The second gap concerns a lack of a composite construct in the domain of L2 learning in an academic-learning context. Individual constructs have been created for mid-level domain self-concepts. This study created individual constructs to examine the construction of a composite construct of a positive L2 self and how it relates to both a global-level positive self-concept and a more specific composite motivational construct. As with the global construct, the positive L2 self construct is future-oriented and related to learning. In addition, the relationship among composite positive L2 self and objective measures of L2 proficiency was examined. Components of a positive L2 self are an interested-in-L2 self, a passion for L2 self, and a learning goal orientation. The construct of an interested-in-L2 self developed from a positive psychology perspective has yet to be created and a measure for this construct was designed for this study. As Eidswick (2010) stated, "It is surprising then that so few studies have been conducted to investigate the presence of interest on second language (L2) learning" (p. 150). The construct of a passion for L2 self has not been developed so a measure for this construct was created for this study.

A third gap consisted of the lack of a composite motivational construct composed of L2 self-efficacy constructs. Individual constructs have been created for domain skills that assessed self-efficacy for specific L2 tasks however, new scales were constructed for this study. The composite motivational construct included components of L2 speaking self-efficacy, L2 reading self-efficacy, and L2 listening self-efficacy.

A fourth gap is the lack of incorporation of separate self-related constructs into an integrated structural model. In recent years there has been a renewed interest in applying self-psychology to language learning (Dörnyei, 2005, 2009; Mercer, 2011; Mercer, Ryan, & Williams, 2012). Studies in positive psychology, educational psychology, and L2 motivation often study an isolated variable or a small set of variables that are correlated with each other. However, few studies have elaborated an integrated structural model. No study has yet to elaborate a model at multiple levels containing constructs from these separate fields and empirically tested it.

Purposes of the Study

The main purpose of this study is to explore empirically some core constructs of positive psychology by testing a structural model of the causal relationships among constructs of global self-concept, L2 domain-specific self-concept, and L2 proficiency. In order to do that, it is first necessary to ascertain the measurable components of each construct. All relevant constructs are introduced below, starting with the two macro-level composite constructs, and continuing with the directly measured component constructs.

The first purpose of this study is the construction of latent variables of at different levels of specificity. A global positive self in an academic learning context, was constructed composed of three components of that facilitate learning, flourishing, curiosity, and hope. A latent variable of positive L2 self that facilitates language learning was constructed with components of interested-in L2 self, harmonious passion for L2 learning, and an L2 mastery goal orientation. A latent L2 motivational variable of L2 self-efficacy was constructed with components of L2 speaking self-efficacy, L2 listening self-efficacy, and L2 reading self-efficacy. An attention to levels of specificity is common to hierarchical models of self (Marsh & Craven, 2006; Marsh & Shavelson, 1985). These levels form part of the self-system, that is, a motivated self-relevant meaning system that informs, constrains and guides interpretations of experience, goals, and self-regulation (Swann & Bosson, 2010).

The second purpose was to determine the structural relationships among the levels of specificity of the latent self-variables. The direction of the relationship paths are hypothesized to move from the general to a middle-level to the specific. It is important to determine the structural relationships because past studies on the effect of global self-concepts on academic achievement are inconsistent or even negative (Baumeister, Campbell, Krueger, & Vohs, 2003; Pullmann & Allik, 2008; Valentine, Dubois, & Cooper, 2004).

The third purpose concerned the structural relations among positive self-concept, positive L2 self, and L2 proficiency. The composite constructs of positive self-concept and positive L2 self were composed of variables using self-reports. To create a model that goes beyond self-reports, another model was created that uses an objective measure of L2 proficiency. The reason to use an objective measure was to ensure that relationships do not exist due to a method-effect where all measures used the same type of instruments. In other words, the objective measure can confirm that relationships exist independently of the type of variable used.

Research Questions

The research questions that guide this study are:

1. To what degree can a composite construct of Positive Self-concept be constructed?
2. To what degree can a composite construct of Positive L2 Self be constructed?
3. To what degree can a composite construct of L2 Motivation be constructed?
4. To what extent does Positive Self-concept affect L2 Motivation with Positive L2 Self a mediator?
5. To what extent does Positive Self-concept affect L2 Proficiency with Positive L2 Self a mediator?

CHAPTER 3

METHODS

In this chapter the participants and educational context are described, then the instruments and survey design are described, and finally, the procedures of Rasch analysis, confirmatory factor analysis, and structural equation modeling are explained. The instruments are described in sections organized by level of specificity, from general, global self-concepts, to an L2 domain-specific level, to a highly specific motivational level. The procedures for Rasch analysis are described for items and scales. Confirmatory factor analysis procedures are described as a prelude to the procedures for structural equation modeling. Finally, cross-validation using structural equation modeling is described.

Participants

The participants in this study are 539 first- and second-year Japanese female students in a two-year private college and a four-year private university in western Japan. Most of the students are 18-20 (M = 19.2; SD = 1.3) years old, an age that can be characterized as the developmental period known as emerging adulthood (Arnett, 2000, 2004). The selection of these students was based on a convenience sampling from an institution in which the researcher had access to an adequate pool of participants. However, as wide as possible a range of teachers and students were asked to participate to more broadly sample from within the institution. The participants are in English-speaking classes taught by nine different teachers. The students came from a variety of majors including: Contemporary Culture, Expressive Arts, Psychology, Early Childhood Development, and English. Some of the students took a TOEIC test and some took a TOEIC Bridge test depending on the faculty and year in school. Second-year students took the TOEIC and first-year students took the TOEIC Bridge. TOEIC stands for Test of English for International Communication and is a test of English proficiency for people whose native language is not English. The TOEIC Bridge is an easier version that measures lower levels of proficiency with scores ranging from 20 to 180. The mean TOEIC score was 382.5 with a standard deviation of 96.3. The mean TOEIC Bridge score was 122.7 with a standard deviation of 18.8. The students that took the TOEIC were of higher ability on average than the students who took the TOEIC Bridge. The students were native speakers of Japanese in intact classes of approximately 15 to 20 students per class. As is typical for Japanese students, before entering college they had studied English for three years in junior high school and three years in high school.

Educational Context

The site where this research was conducted is a private women's university in southwestern Japan. The school was founded by Christian missionaries in the late 1800s. All students take courses called first-year English in two classes that meet twice a week, that is, one course over four days a week. These are the only courses that are common to all students. In addition to these classes, there are a number of compulsory and elective courses that vary among student major fields or tracks within a major. Most students come from middle to upper-middle class backgrounds. Most students have graduated from local high schools in the same prefecture.

Validation for the instruments was done in two parts. In the first part, internal validation was done though Rasch analysis of items and scales. In the second part, validation external to the items and scales was done though validation of variables situated in a nomological network with other related but different variables (Cronbach & Meehl, 1955; Loevinger, 1957). Strong positive relationships showed by high positive correlations was taken as convergent validity evidence while inverse or negative relationships showed by negative correlations was taken as divergent validity evidence.

Instrumentation

The main instruments used in this study is a 187-item self-report questionnaire comprised of the 23 scales described below, and L2 proficiency measures, the TOEIC Bridge and TOEIC. The 187 items were divided into two parts. All response options were labeled because this tends to improve reliability and validity (Krosnick, 1999; Krosnick & Fabrigar, 1997; Krosnick & Presser, 2010; Tourangeau, Cooper, & Conrad, 2007). Part one contained 28 items consisting of words relating to affect and feeling with six response options: *1* =

Never, 2 = Seldom, 3 = Sometimes, 4 = Often, 5 = Almost always, and 6 = Always. Part two contained 159-items composed of various scales with six response options: *1 = Definitely not true of me, 2 = Not true of me, 3 = Slightly not true of me, 4 = Slightly true of me, 5 = True of me, and 6 = Definitely true of me*. Although this questionnaire is on the longish-side, Dörnyei and Taguchi (2010, p. 12) suggest that more than four to six pages or over half an hour is too long. Given that this questionnaire was four pages and took 25 minutes to complete, give or take five minutes, it was deemed to be an acceptable length. Some of the instruments used in this study are well established in the general psychology literature and have been used in hundreds or even thousands of studies. Others were created for this study. The established scales come from a variety of formats in both English and Japanese with differing response options; however, for this study all the scales were given in Japanese with six response options. An experienced English teacher who is a native-speaker of Japanese with a master's degree in TESOL translated the English items into Japanese. Another different experienced English teacher who is a native-speaker of Japanese also with a master's degree in TESOL back translated all the items back into English (Brislin, 1970). I checked the back translated with reference to the originals and determined the items to be acceptable. In addition, this study goes beyond self-report by using English language proficiency scores as measured by the TOEIC Bridge and TOEIC as an outcome measures.

The instruments of this study were used to develop a model that shows structural relationships with a positive L2 self. In the proposed model, the positive L2 self is an intervening or mediating variable between a positive self independent variable and a dependent or outcome variable of language proficiency as measured by TOEIC Bridge and TOEIC scores. As mentioned previously, this model has three levels of specificity. The positive self-concept is a general latent construct that is global and trait-like. The second level contains a latent construct that is also dispositional but that is more specific in that it centers on a construct of a positive academic learning of English. This latent construct is more specific in that it relates to positive beliefs about the L2, particular classroom beliefs, learning behaviors, and cognitive beliefs about learning an L2. The third level of specificity is self-efficacy beliefs about specific language tasks in the specific skills of reading, speaking, and listening.

Measured Variables Related to Positive Self-Concept

Components of positive self-concept are broad constructs related to well-being, happiness, and positive personal growth. Three scales were used to model positive self-concept: Curiosity and Exploration Inventory-II, Hope, and Flourishing. For the purpose of instrument validation by providing convergent validity evidence, other positive psychological functioning and well-being constructs that are also related to a global positive self are elaborated, namely: Self-Esteem, Subjective Happiness, Positive Social Relationships, Satisfaction in Life, and Positive trait affect. Again for the purpose of instrument validation, in this case, divergent evidence, three constructs are elaborated: Hopelessness in Achievement, Hopelessness of Interpersonal Relations, and Negative trait affect.

The three measures used for modeling are all selected within the context of an academic learning environment. Learning can be considered Janus-faced, that is, facing forward or backward in time. Constructs like curiosity, hope, flourishing, interest, passion, mastery goal orientation, and self-efficacy are situated in the present but are oriented to the future. Constructs such as self-esteem, subjective happiness, positive social relationships, satisfaction in life, positive and negative trait affect are situated in the present but are oriented to the past. Ideal L2 self and intended learning effort are based on the idea of discrepancies between some future state and present state. In positive psychology the emphasis is on being authentic and true to oneself (Ryff & Singer, 2008; Schlegel & Hicks, 2011; Seligman, 2002; Sheldon, 2002), rather than an emphasis on reducing the discrepancy between present and a future ideal. Although, constructs oriented toward the past and the future correlate, the setting of a learning context with participants beginning to live life on their own determines the orientation toward the past. As mentioned in Chapter 2, there are a large number of measures at the trait-level in positive psychology. This is because researchers were looking for stable measures such as, personality traits, attitudes, character strengths, and virtues that could describe positive aspects of life.

Curiosity and Exploration Inventory-II

The Curiosity and Exploration Inventory (CEI; Kashdan, Rose, & Fincham, 2004) was an earlier version of the Curiosity and Exploration Inventory-II (CEI-II; Kashdan et al., 2009) a scale designed to measure trait curiosity. The response options are on a 6-point scale ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The CEI-II contains five items measuring a dimension of curiosity about the motive to seek out new knowledge and experiences (e.g., *I actively seek as much information as I can in new situations* and five items that measure a dimension of curiosity about a general willingness to embrace the novel, uncertain, and unpredictable in life (e.g., *I am the type of person who really enjoys the uncertainty of everyday life*). Kashdan et al., reported alpha reliabilities of .85 and .86. They also suggested that because the two dimensions strongly correlate that the ten items be used together. They also did an IRT analysis with all ten items and found that the items covered a good range of trait curiosity with the scale centered on the midpoint, as is common with self-report measures. For the version used with this sample the Rasch person reliability was .82, the Rasch person separation was 2.10, the Rasch item reliability was .99, and the Rasch item separation was 12.11. The Curiosity and Exploration Inventory-II scale is in Appendix B.

Hope Scale

The Hope Scale (Snyder, Harris, et al., 1991) is an 8-item scale that measures trait level hope. The hope construct consists of two factors. Four items reflect agentic thinking about one's goals or "willpower" (e.g., *I meet the goals that I set for myself*) and four items reflect a "pathways" thinking about the ways to achieve goals or "waypower" (e.g., *There are a lot of ways around a problem*). Six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. Also, piloted were two items about growth hope in case they were needed (e.g., *I feel hopeful about being a better person than I am now*) and one item concerning hope for the future (*I feel hopeful about the future*); these were adopted from a study by Tong, Fredrickson, Chang, and Lim (2010). It is unknown if these additional items are factors of agency, pathways, or a new factor. In order to be able to make comparison with past studies of hope, these additional items were not used. Alpha reliabilities have ranged from .74 to .88. For the version used with this sample the Rasch person reliability was .81, the Rasch person separation was 2.10, the Rasch item reliability was .98, and the Rasch item separation was 6.79. The Hope Scale is in Appendix C.

Flourishing Scale

The flourishing scale (Diener, Wirtz, et al., 2010) consists of eight items describing aspects of positive functioning and human flourishing (e.g., *I actively* *contribute to the happiness and well-being of others*). Six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The alpha reliability reported was .87. For the version used with this sample the Rasch person reliability was .75, the Rasch person separation was 1.72, the Rasch item reliability was .99, and the Rasch item separation was 12.43. The Flourishing Scale is in Appendix D.

Positive Self Instruments for Validation

This section describes variable used for validation evidence. These variables are peripheral to the main study where variables are used to construct latent variables and a structural model.

Self-Esteem Scale

The Self-Esteem Scale is based on the Rosenberg Self-Esteem Scale (RSE) (Rosenberg, 1965). The original RSE is a 10-item scale of global self-esteem typically administered with 4-, 5-, or 7-point response options ranging from *Strongly disagree* to *Strongly agree* (e.g., *On the whole, I am satisfied with myself*); the RSE is the most widely used and has received the most psychometric and empirical validation of any self-esteem measure (Byrne, 1996; Gray-Little, Williams, & Hancock, 1997; Mruk, 2006). A Japanese version was created with the five negatively worded items re-worded so that they would not need to be reverse-scored and also to avoid the possible statistical problems that can occur with negatively worded items. Re-wording was done by eliminating or changing the negative words to positive. Negatively worded items often create a separate additional dimension and have lower item discriminations (Barnette, 2000; DiStefano & Motl, 2006; Horan, DiStefano, & Motl, 2003; Rodebaugh, Woods, & Heimber, 2007). When this was done, two items resembled other items. *At times I think I am no good at all* when reversed in Japanese, was similar to *I feel that I have a number of good qualities*. *I wish I could have more respect for myself* when reversed in Japanese, was similar to *I often feel I have much to be proud of*. Because of the similarity, these two items that were reversed were deleted from the scale. After piloting, one item was removed because of a poor item-total correlation of .33 and a poor infit mean square value in the Rasch analysis of 1.45, and a poor outfit mean square value of 1.50. Self-esteem is an overall global self-estimation so items should have high correlations (Gray-Little, Williams, & Hancock, 1997; Robins, Hendin, & Trzesniewski, 2001) An item-measure correlation of .33 relative to the inter-measure correlations of the other items .66 to .75 is much lower. Item misfit was also much greater relative to the other items. Seven items were retained for the final version. Previous studies have suggested that items can be reduced from the RSE while retaining good psychometric qualities (e.g., Gray-Little et al., 1997; Robins, Hendin, & Trzesniewski, 2001). Six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. Alpha reliabilities for the RSE in past research have ranged from .72 to .88 (Gray-Little et al., 1997). For the version used with this sample the Rasch person reliability was .82, Rasch person separation was 2.11, Rasch item reliability was .96, and Rasch item separation was 5.15. The Self-Esteem Scale is in Appendix E.

Satisfaction With Life Scale

The Satisfaction with Life Scale (SWLS) (Diener, Emmons, Larsen, & Griffin, 1985) contains five items that measure overall life satisfaction (e.g., *I am satisfied with my life*). Participants choose one of six response options ranged that from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The items do not refer to any specific life domains or cultural influences but only how a person consciously judges general life satisfaction. SWLS is widely used with measures of positive and negative affect to measure subjective well-being (Diener, 1994; Diener et al., 1999; Lucas, Diener, et al., 1996). The SWLS has been used in many cultures and translated into many different languages (Kuppens, Realo, & Diener, 2008). Alpha reliability for the SWLS has ranged from .79 to .89 (Pavot & Diener, 1993). For the version used with this sample the Rasch person reliability was .74, the Rasch person separation was 1.70, the Rasch item reliability was .99, and the Rasch item separation was 12.95. The Satisfaction with Life Scale is in Appendix F.

Positive and Negative Affect Scales

The Positive and Negative Affect Scales are based on items from the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS contains ten descriptor words of positive affect and ten descriptor words of negative affect (e.g., *enthusiastic* is a positive descriptor word; and *scared* is a negative descriptor word). The scale can be used for measuring affective states or traits depending on the time frame specified, for example, a short time frame for states, to what extent *You feel this way right now, that is, at the present moment* to a more trait-like time frame, *You generally feel this way, that is, how you feel on the average*. The respondents choose from six response options: *1 never, 2 seldom, 3 sometimes, 4 often, 5 almost always, and 6 always*. The two scales correlate either not at all or slightly negatively. The PANAS is one of the most widely used affect scales and has been translated into many languages and for a variety of types of studies (Schmukle, Egloff, & Burns, 2002). The Positive and Negative Affect scales used were adapted from a shorter Japanese version of the PANAS developed by Sato and Yoshida (2001). Watson, Clark, and Tellegen (1988) reported reliabilities for the two scales that ranged from .84 to .90. In the version used in this study the Rasch person reliability was .77 for positive affect, Rasch person separation of 1.83, Rasch item reliability of .99, and Rasch item separation of 10.06. For negative affect the Rasch person reliability was .81, the Rasch person separation was 2.04, the Rasch item reliability was .99, and the Rasch item separation was 11.42. The Positive and Negative Affect scales are in Appendix G.

Positive and Negative Feelings

The Scale of Positive and Negative Experience (SPANE) was designed to measure subjective positive and negative feelings. Six items measure positive feelings and six measure negative feelings. Where the PANAS items have high affect arousal, the SPANE spans more levels of arousal. Three items of each scale reference general feelings (e.g., *positive*, *negative*) and three items of each scale are more specific about feelings (e.g., *joyful*, *sad*; Diener, Wirtz, Toz, Kim-Prieto, Choi, Oishi, & Biswas-Diener, 2010). For this study, six response options ranged from: *1 never, 2 seldom, 3 sometimes, 4 often, 5 almost always, and 6 always*... Reported alpha reliabilities ranged from .81 to .89 (Diener et al., 20100. The Rasch person reliability for this study was .82 for positive feelings, Rasch person separation 2.14, Rasch item reliability .99, and Rasch item separation 8.85. For negative feelings the Rasch person reliability was .76, Rasch person separation was 1.80, Rasch item reliability was .88, and Rash item separation was 2.71. The Positive and Negative Experience Scales are shown in Appendix H.

Subjective Happiness Scale

The Subjective Happiness Scale uses three items based on the Subjective Happiness Scale (SHS; Lyubomirsky & Lepper, 1999). The SHS contains four items with one item reversed scored (e.g., *In general, I consider myself*: with response options ranging from *not a very happy person* to *a very happy person*). When the negative wording was changed and translated into Japanese, one item resembled another item so this item was dropped. Although the results for the three piloted items were low but acceptable (For the piloted 3-item version the alpha reliability was .71), in a second piloting, four other items tapping into the construct of subjective happiness were created and piloted. This resulted in a total of seven items. Six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. Alpha reliabilities have been reported to range from .79 to .94. Test-retest reliability in time frames from three weeks to one year ranged from .55 to .90 (Lyubomirsky & Lepper, 1999). For the scale with seven items the Rasch person reliability was .86, the Rasch person separation was 2.46, the Rasch item reliability was .99, and the Rasch item separation was 9.76. The Subjective Happiness Scale is in Appendix I.

Positive Social Relationships Scale

The seven items on the Positive Social Relationships Scale were adapted from three sources. The first source was four items from the positive relations with others subscale of Ryff's psychological well-being scale (e.g., *I enjoy personal and mutual conversations with family members and friends*. Ryff & Keyes, 1995) items five and six were adapted from Keyes social well-being scale (e.g., *You believe that* *people are kind* was changed to *I believe that people are kind* to fit the other items in the questionnaire. Keyes, 1998). The last item was created based on the theory that good social relationships are a component of well-being. Six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. For the version used here the Rasch person reliability was .71, the Rasch person separation was 1.57, the Rasch item separation was .99, and the Rasch item separation was 12.71. The Positive Social Relationships Scale is in Appendix J.

Grit Scale

The Grit Scale is a 9-item measure of perseverance and passion for long-term goals adapted from Duckworth et al., (2007). The construct of grit is composed of two factors. The first four items measure a factor of consistency of interests that were all reverse scored in the original and reworded for this study (e.g., *I can maintain my interest in topics for years*) and the last five items measure a factor of perseverance of effort (e.g., *I finish whatever I begin*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. Reported alpha reliabilities ranged from .77 to .85. For this version the Rasch person reliability was .81, the Rasch person separation was 2.04, the Rasch item reliability was .99, and the Rasch item separation was 10.53. The Grit Scale is shown in Appendix K.

Achievement Hopelessness Scale

The Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974) is a 20-item scale with true or false response options (e.g., *Things just won't work out the way I want them to* or reverse scored: *I look forward to the future with hope and enthusiasm*). The BHS was developed with a sample of hospitalized patients who had made recent suicide attempts. It is used in this study to provide divergent validity evidence. The scale contains eleven items measuring hopelessness and nine items measuring hope that are reverse scored and it has a reported alpha reliability of .93. An item response theory analysis of the BHS found that it measures a single dimension of hopelessness (Young, Halper, Clark, Scheftner, & Fawcett, 1992). Six items were adapted for use in this study. The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The Achievement Hopelessness Scale had a Rasch person reliability of .72, Rasch person separation of .1.59, Rasch item reliability of .98, and a Rasch item separation of 8.03. The Achievement Hopelessness Scale is in Appendix L.

Relationship Hopelessness Scale

A study in Japan extended the Beck Hopelessness Scale by recasting the items in terms of interpersonal relations (e.g., *I don't think my relationships with my friends will become what I want*; Takahira, 1998). Six items were adapted for this study. The six response options ranged from *Definitely not true of me, Not true* *of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The Relationship Hopelessness Scale had a Rasch person reliability of .68, Rasch person separation of 1.45, Rasch item reliability of .99, and a Rasch item separation of 9.32. The Relationship Hopelessness Scale is in Appendix M.

Positive L2 Self Instruments for Model

Three instruments provide measures that are used to construct a latent variable of positive L2 self. They are Interest in L2 Self Scale, Passion for L2 Learning Scale, and the Mastery Goal Orientation Scale. In addition, they are used as part of a structural model.

Interest-in-L2 Self Scale

Based on the literature on interest as a psychological construct (Berlyne, 1949, 1960; Dewey, 1913; Harackiewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008; Hidi, 1990; Hidi & Renninger, 2006) eight items were created (e.g., *English is an interesting field of study*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*.. The Rasch person reliability was .87, the Rasch person separation was 2.59, the Rasch item reliability was .99, and the Rasch item separation was 9.38. The Interested-in-L2 Self Scale is in Appendix N.

Passion for L2 Learning Scale

Items from the Harmonious Passion Scale (Vallerand, Blanchard, Mageau, Koestner, Ratelle, Léonard, & Gagné, 2003) were adapted for English learning (e.g., *I am passionate about learning English*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The seven-item Passion for L2 Learning Scale measures only harmonious passion; obsessive passions are not part of this study. The Rasch person reliability for the scale was .87, the Rasch person separation was 2.62, the Rasch item reliability was .99, and the Rasch item separation was 11.92. The Passion for L2 Learning Scale is in Appendix O.

Mastery Goal Orientation Scale

The eight items on the Mastery Goal Orientation Scale were adapted from three sources. Three items were adopted from the Mastery-Approach Goal Subscale of the Achievement Goal Questionnaire-Revised (Elliot & Murayama, 2008; e.g., *My aim is to completely master the material presented in this class*); one item was adapted from the original Mastery-Approach Goal Subscale of the Achievement Goal Questionnaire (Elliot & McGregor, 2001; *I want to learn as much as possible*) and four items were created based on the achievement goal literature for this study (e. g., *I like learning difficult things in this class*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The alpha reliability for the 2008 study was .84, for the 2001 study, it was .88. For the version used with this sample the Rasch person reliability was .85, the Rasch person separation was 2.35, the Rasch item reliability was .99, and the Rasch item separation was 13.66. The Mastery Goal Orientation Scale is in Appendix P.

Positive L2 Self Instruments for Validation

Three instruments are for measures used for validation evidence of the main variables used to construct a latent variable and for use in a structural model. They are Ideal L2 Self, Prosociality Goals, and Math Self-Concept. These are peripheral variables to the main variables in that they are used in the preliminary analysis to show either convergent or divergent relationships with the main variables.

Ideal L2 Self

The Ideal L2 Self scale was designed to measure "the clarity and intensity of learners' visions of themselves as users of the language" (Ryan, 2008, p. 147; e.g., *If my dreams come true, I will use English effectively in the future*). As mentioned in Chapter 2 it is a peripheral variable in this study to confirm convergent validity. It is a 6-item scale with six response options that ranged that from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The reported alpha reliabilities ranged from .82 to .87. For this sample the Rasch person reliability was .84, the Rasch person separation was 2.26, the Rasch item reliability was .99, and the Rasch item separation was 11.90. The Ideal L2 Self Scale is in Appendix Q.

Prosociality Goals

Prosociality goals, as used in this study to confirm convergent validity, refer to positive social relationship behavior in academic contexts, that is, more specifically, positive relationships with fellow students and classmates in the school and class settings. Wentzel (1993) found that prosocial behavior predicted academic achievement. The items one, two, and three in the scale used in this study was based on the Prosocial Subscale of the Social Goals Scale (Wentzel, 1993), item four from the peer relationships subscale of the social-emotional learning scale (Coryn, Spybrook, Evergreen, & Blinkiewicz, 2009), and items five, six, and seven from the social relatedness subscale of a need satisfaction questionnaire (Kunter, Baumert, & Koller, 2007). Seven items are in this scale. An example item is *I cooperate with my classmates to learn new things*. The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. For the version used in this sample the Rasch person reliability was .76, the Rasch person separation was 1.78, the Rasch item reliability was 1.00, and the Rasch item separation was 17.11. The Prosociality Goals Scale is in Appendix R.

Math Self-Concept

The math self-concept scale is a subscale of the Academic Self Description Questionnaire II (Marsh, 1990, 1992) designed to measure self-concept in the academic domain of math. Math self-concept is included in this study to provide evidence of divergent validity. It has eight items with six response options that ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me* (e.g., *I learn things quickly in mathematics classes*). Reported alpha reliabilities are usually around .90 or better. For a version used with this sample the Rasch person reliability was .85, the Rasch person separation was 2.42, the Rasch item reliability was .97, and the Rasch item separation was 5.54. The Math Self-Concept scale is in Appendix S.

Motivational Instruments for Model

The three motivational instruments are for measures of self-efficacy for different L2 skills. They are the Speaking Self-Efficacy Scale, the Listening Self-efficacy Scale, and the Reading Self-Efficacy Scale. These are the main motivational variables that are used to construct a latent variable and for a structural model.

Self-efficacy is the belief that one has the capabilities in a situation to enact a course of action to attain a given level of performance (Bandura, 1977, 1986, 1997). It is the judgment that one has the means to do a task successfully. Expectancy beliefs might be different than efficacy beliefs because they are more focused on the end state and the possibility of an end being reached. Confidence differs from self-efficacy in being more generalized across skills, levels, and situations. Self-efficacy as used here refers to being capable of successfully engaging in language skill performances at different levels of difficulty. Pajares (2001) found that self-efficacy was positively related to positive psychology variables, such as, scales of optimism and authenticity.

Speaking Self-Efficacy Scale

This scale is designed to measure self-efficacy in speaking. Three items from the Speaking Self-Efficacy scale were adapted from Burrows (2009; e.g., *I can speak English to order a meal in a restaurant*) and six items were created for this study (e.g., *I can give a speech in English*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. Because the Burrows (2009) study was about the theory and design of a scale there were no previously reported alpha reliabilities. For the version used with this sample the Rasch person reliability was .87, the Rasch person separation was 2.64, the Rasch item reliability was 1.00, and the Rasch item separation was 17.73. The Speaking Self-Efficacy Scale is in Appendix T.

Listening Self-Efficacy Scale

Listening self-efficacy as used here refers to the belief in being capable of successfully listening and understanding at different levels to different sources of spoken language. Two items from the Listening Self-Efficacy scale were adapted from Mills, Pajares, and Herron (2006) and seven were created for this study (e.g., *I can understand the main ideas when listening to English songs*. The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. In the Mills, Pajares, and Herron (2006) study, reported alpha reliability was .97. For the version used with this sample the Rasch person reliability was .87, the Rasch person separation was 2.55, the Rasch item reliability was 1.00, and the Rasch item separation was 15.09. The Listening Self-Efficacy Scale is in Appendix U.

Reading Self-Efficacy Scale

Reading self-efficacy as used here refers to the belief in being capable of successfully reading and understanding written texts at different levels for sources differing in levels of difficulty. Two items from the Reading Self-Efficacy scale were adapted from Mills, Pajares, and Herron (2006) and five were created for this study (e.g., *I can read and understand a menu in English*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. In the Mills, Pajares, and Herron (2006) study, reported alpha reliability was .95. For the version used with this sample the Rasch person reliability was .86, the Rasch person separation was 2.46, the Rasch item reliability was .99, and the Rasch item separation was 9.09. The Reading Self-Efficacy Scale is in Appendix V.

Motivational Instruments for Validation

Two motivational instruments are for measures used for validation evidence of the main variables used to construct a latent variable and for use in a structural model. The are the Intended Learning Effort Scale and the Persistent Effort at L2 Learning Scale. These are peripheral motivational variables to the main motivational variables in that they are used in the preliminary analysis to show convergent or divergent relationships with the main variables.

Intended Learning Effort Scale

Intended learning effort is a scale used by Ryan (2008) that was designed to measure perceptions of efforts to learn and possible intended future efforts. The Intended Learning Effort scale was designed to measure “both learners’ perceptions of their current efforts to learn and their possible intended future efforts” both inside and outside the classroom (Ryan, 2008, p. 147; e.g., *If an English course were offered in the future, I would like to take it*). In this study, it used to confirm convergent validity with the motivational variables that are part of the structural model. It is an eight item scale with six response options that ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me,* *True of me,* to *Definitely true of me*. The reported alpha reliabilities ranged from .81 to .86. The Rasch person reliability for this sample was .85, the Rasch person separation was 2.34, the Rasch item reliability was .99, and the Rasch item separation was 12.99. The Intended Learning Effort Scale is in Appendix W.

Persistent Effort at L2 Learning Scale

Persistent effort is defined as the amount of time and frequency one spends studying the L2 and persisting in the face of obstacles or difficulties. The Persistent Effort at L2 Learning Scale was created to measure participants' beliefs in how much they persist or keep at studying English. One aspect of persistence measured is time and frequency and another is not quitting because of discouragement or setbacks. In this study, it is used to confirm convergent validity with the motivational variables that are part of the structural model. Six items were created (e.g., *When I have a problem understanding English, I keep trying until I understand*). The six response options ranged from *Definitely not true of me, Not true of me, Slightly not true of me, Slightly true of me, True of me,* to *Definitely true of me*. The Rasch person reliability of this sample was .81, the Rasch item separation was 2.04, the Rasch item reliability was 1.00, and the Rasch item separation was 14.48. The Persistent Effort at L2 Learning Scale is in Appendix X.

Table 3 contains a list of the constructs used in this study. Shown are the modeled constructs and constructs used to show positive and negative relationships for convergent and divergent validity evidence.

Table 3. *Summary of Modeled Constructs and Constructs for Validity Evidence*

Modeled constructs	Positive relationships (Convergent validity)	Negative or no relationships (Divergent validity)
Positive Self: Flourishing, Curiosity, Hope	Self-esteem, Satisfaction in life, Positive affect, Subjective happiness, Positive feeling, Negative feeling, Positive social relationships, Grit	Negative affect, Negative feeling, Hopelessness in achievement, Hopelessness in relationships
Positive L2 Self: Interested-in-L2 self, Harmonious passion for L2 learning, Mastery L2 goal orientation	Prosociality goals, Ideal L2 self	Math self-concept
L2 Motivational Constructs: Speaking self-efficacy, Listening self-efficacy, Reading self-efficacy	Intended learning effort, Persistent effort at L2 learning	

Survey Design

In this study, I used a cross-sectional design where all scales are administered at the same time with the scale items randomly mixed together. In order to reduce response bias, that is, a bias to respond similarly to previous responses, scales that were unrelated or divergent were included. This approach was preferable to including negatively worded items, because they can add another dimension to a scale (e.g., DiStefano & Motl, 2006; Horan, DiStefano, & Motl, 2003; Motl & DiSefano, 2002). For example, rather than include negatively worded hope items that would need to be reverse scored, I included hopelessness items. These items then would serve two functions, i.e. preventing response bias and provide divergent validity evidence. The complete survey is in Appendix A.

Measured Variables Composing L2 Proficiency

English proficiency is the general ability to use the English language for communicative purposes. It is composed of oral and written dimensions that are measured in this study with either the TOEIC, a standardized test of business English proficiency, or a lower level test of listening and reading proficiency called TOEIC Bridge. The TOEIC is a two-hour multiple-choice test composed of one hundred listening questions and one hundred reading questions, each section has a reported reliability of .90 or above. The total scores are reported on a scale from 10 to 990. The TOEIC Bridge is a one-hour multiple-choice test composed of 50 listening questions and 50 reading questions, each section has a reported reliability of .85. The total scores are reported on a scale from 20 to 180.

Procedures

Teachers participating in this research were given the questionnaire to distribute to students to complete during class time. Nine teachers participating in the research were given the questionnaire to distribute to participating students to complete. The students were asked to participate in a research project by filling out a questionnaire during class time. The students were told that participation was voluntary, would not affect their grades, and promised that anonymity would be maintained. The average completion time was about 25 minutes, with some students finishing up within five minutes before or after the average.

The complete questionnaire was administered by teachers in their own classes at their own convenience toward the end of the 15-week second semester. Data from the completed questionnaires were manually entered into a spreadsheet. From the student identity numbers, TOEIC and TOEIC Bridge were individually looked up and entered with the other data. Sets of data for the different scales could then be imported in Winsteps (Linacre, 2011) for the Rasch analysis and then into Amos for the confirmatory factor analysis and structural equation model.

Rasch Analysis

Rasch Model Introduction

Rasch models have many beneficial characteristics: measurement invariance, the conversion of ordinal data into interval measures, and information about item and scale functioning, In classical test theory item and examinee statistics are sample dependent. In other words, statistics about groups of examinees change with differing groups. For example, for one group a researcher might get a reliability coefficient of .95 and for another group .65. Even though researchers talk in a short-handed fashion about test reliability, there is really no such thing as a reliable test, that is, in classical test theory reliability refers to the data produced by a particular group (Thompson, 2003). Item statistics also have the same characteristics of changing with differing samples. Also, in classical test theory, statistics about people change with differing items. This dependence on samples of items and people pose a problem for measurement. Imagine if using three different measures of length or the same measure over three occasions for a person's height gave readings of: 170 centimeters, then 150 centimeters, and then 180 centimeters. Measurement would be inconsistent and thus impossible. This was the problem George Rasch solved. Rasch developed a model based on a concept of specific objectivity that provides measurement invariance (Engelhard, 2013; Rasch 1960/1993; Wright, 1977). Measurement models and analysis using his insights are called Rasch models.

Rasch models provide item and examinee statistics that do not depend on any particular set of items or examinees, that is, they are sample independent. Invariance solves many measurement problems. With invariant measurement it becomes possible to monitor test or item quality and person or group measurement precision. Just as with measuring length, it is useful to evaluate the level of precision for different measurement purposes. For example, it might be satisfactory to measure the distance between cities in meters. For clothing, meters would be too imprecise and centimeters would be a better level of precision. Understanding Rasch models helps in test or scale construction (Wright & Masters, 1982; Wright & Stone, 1979). Solutions to other practical issues are relatively easy using Rasch analysis: item banking, test linking, measuring learning and development, placing students into similar ability groups, and making educational program decisions (Boone, Staver, & Yale, 2014; Engelhard, 2013; Rasch 1960/1993; Wright, 1977).

Raw data from Likert type scales return ordinal data; however, one of the assumptions of parametric statistics is that the data be interval in form. Interval measures can be constructed by transforming the ordinal data from the scales by applying the Rasch measurement model. This model is a stochastic or probabilistic model that is based on the probabilities that a person will endorse different items, or given an item, the probabilities that persons with different abilities will endorse that item. While the mathematics might seem complicated at first, further study reveals what Thorndike (1904) pointed out long ago about statistics, "There is, happily, nothing in the general principles of modern statistical theory but refined common sense, and little in the technique resulting from them that general intelligence can not readily master" (p. 1). The Rasch model is based on the conceptually simple idea that respondents with higher ability have increasingly higher probabilities of responding to more difficult to endorse items, while items with greater difficulty decreases the probability of being endorsed (Rasch, 1960/1993). This can be also be checked with item-measure correlations which show that higher observations correspond to more of the latent variable.

The Rasch model can be represented mathematically in different forms. The natural logarithm of the odds ratio in the Rasch model is modeled as the difference of a person's ability level from the item difficulty. In mathematical form the log odds dichotomous Rasch model can be given by:

$$\text{Logits} = \text{Log odds} = \ln[P_{ni}/(1 - P_{ni})] = \beta_n - \delta_I \qquad (1)$$

Where :

$\ln = \log_e$ = natural logarithm or logarithm to the base e

P_{ni} = probability of success for person n on item i

$1 - P_{ni}$ = probability of failure for person n on item i

β_n = person n's ability level and scale location

δ_i = item i's difficulty level or scale value

Natural logarithm of the odds or log odds are usually given in units called logits. The abilities of the respondents (also called persons) and difficulties of the items can then be mapped onto a scale with the same linear interval units. As can be seen by the formula, in the Rasch model items and abilities have specific objectivity, that is, they are invariant over specific items or specific persons. Item difficulties can be generalized beyond the sample and person abilities generalized beyond a particular set of items. For example, in the case of person ability β_1 and person ability β_2 for an item difficulty of level of δ_i, the difference is:

$$\ln[P_{i1}/(1 - P_{i1})] - \ln[P_{i2}/(1 - P_{i2})] = (\beta_1 - \delta_i) - (\beta_2 - \delta_i) = \beta_1 - \beta_2 \qquad (2)$$

The item difficulty drops out of the calculation so the difference in the ability of person 1 and person 2 is invariant and does not rely on a particular item or set of items. In the same way, it can be shown that the item difficulty does not rely on any particular person or set of person abilities.

Another form of the Rasch model is:

$$P_{ni} = e^{(\beta_n - \delta_i)}/(1 + e^{(\beta_n - \delta_i)}) \qquad (3)$$

Where:

P_{ni} = probability of success for person n on item i

β_n = person n's ability level and scale location

δ_i = item i's difficulty level or scale value

This is known as the simple logistic Rasch model. This formula is also known as the one-parameter logistic model by researchers from an item response theory perspective that might include additional parameters; the one parameter in this case being item difficulty. From the Rasch perspective though, person ability is also modeled, so that at least two parameters are included in a Rasch analysis. Even though the formula is the same for one form of the Rasch model and the one parameter model, one of the main differences between the terminology is that an a priori decision is made in Rasch analysis to examine if the data fits the model, where with the one parameter model, the model is fit to the data; if it does not, other models, such as the two- or three-parameter models, can be used with further modeling to fit the data. Unlike the log odds form of the Rasch model, the relationship of probability to item difficulty or person ability is nonlinear. The stretched s-shaped curves of these relationships are called item characteristic curves or item response function curves. They are monotonic ogive curves, that is, they function in one direction with increased probability as points move up the curve.

In addition to a dichotomous model, the Rasch family of models also includes polytomous models for when there are multiple response options such as on a Likert-style rating scale. In this case, items consist of ordered categories with steps or thresholds between categories, for example, an item with six categories would have five thresholds between them. In mathematical form the log odds polytomous rating scale Rasch model can be given by:

$$\text{Logits} = \text{Log odds} = \ln[P_{nik}/(1 - P_{nik})] = \beta_n - (\delta_i + \tau_k) \qquad (4)$$

Where :

ln = $\log_e$ = natural logarithm or logarithm to the base *e*

Pnik = probability of success for person *n* on item *i* at threshold *k*

1 – *Pnik* = probability of failure for person *n* on item *i*

β_n = person *n*'s ability level and scale location

δ_i = item *i*'s difficulty level or scale value

τ_k = threshold difficulty at the *k*th intersection of category boundaries

In the rating scale model the threshold difficulties at the boundaries of the categories are the same for all items on the scale in contrast to the partial credit model where steps between categories are allowed to vary. The number of parameters to estimate in the rating scale model is thus much less than the number for the partial credit model. The person response estimate can be calculated by a combination of the item locations and the threshold difficulties.

The probability form of the rating scale polytomous Rasch model can be given by:

$$P_{nik} = e^{(\beta_n - (\delta_i + \tau_k)} / (1 + e^{(\beta_n - (\delta_i + \tau_k)}) \quad (5)$$

Where:

P_{nik} = probability of success for person *n* on item *i* at threshold *k*

β_n = person *n*'s ability level and scale location

δ_i = item *i*'s difficulty level or scale value

τ_k = threshold difficulty at the *k*th intersection of category boundaries

The mathematical forms of the Rasch model show that it is possible to calculate estimates of person abilities that do not depend on a particular set of items and thresholds, and that it is possible to calculate estimates of scale values and thresholds that do not depend on a particular set of person abilities. This specific objectivity and the additivity for fundamental measurement gives Rasch analysts insights, such as individuals' responses to items, how the person is situated relative to the group, items' contributions to the measure, how information is organized throughout the measure, and how the group is distributed relative to the measure.

Rasch Fit Statistics

When measuring something well, there is a paradox that the more precise researchers are measuring something, the more they should expect their measurement to contain some amount of error. In addition if they are aiming at measurement "truth," they might get less error but the measurement is less useful. For example, say I am measuring the heights of a group of people. I could be quite accurate with no error if I were to gauge heights within a meter of being accurate. If I were to gauge the heights with more precision, say, an estimate on a centimeter scale, my measurements might contain more error, especially with heights that are around the mid-centimeter mark. If I were to be even more precise, to say, the millimeter level, my measurements might often be in error even though they are quite precise. At the precision of the meter level, measurement might be "true" or "correct," but the measurement would not be very useful. At the level of millimeter, measurement might be in error but it could be useful. Buying clothes at the meter level would be a poor fit to their body; buying tailored clothes, even if in error at the millimeter or centimeter level would fit their body well.

Real data differ from the theoretical, mathematical Rasch model as with any type of measurement. However, because both the data and model are known it is possible to calculate differences between theoretical values and actual data. These differences can then be summarized over items or respondents indicating how well the data fit the model. The Rasch model is a probabilistic model so that data can deviate from the model by either deviating in one direction by not being probabilistic "enough" or in another direction as being too random. In other words, data can deviate from the ideal probability (that models a certain amount of probabilistic variation) as being too ordered or absolute (lacking unpredictability or lacking stochasticity) or it can deviate by having too much unknown variance or "noise" (randomness). Data that overfit the model contain fewer probabilistic responses than predicted, and sometimes this is referred to as being deterministic or Guttman-like. Data that underfit the model contain more random responses than predicted. Model-data fit values are derived from residual difference from expected values and actual values in respect to the measurement scale.

If the expected value is subtracted from the actual value the result is the score residual.

$$y_{ni} = x_{ni} - E_{ni} \quad (6)$$

A standardized residual can be calculated by dividing by the square root of the response variance or standard deviation of the actual score responses.

$$z_{ni} = x_{ni} - E_{ni}/[p_{ni}(1 - p_{ni})]^{1/2} = y_{ni}/SD_{xni} \qquad (7)$$

A fit statistic can be calculated by averaging the standardized residual variance for either items or persons.

$$U_i = \text{sum of } z_{ni}^2 \text{ / N for } n = 1 \text{ to } N\text{, or,} \qquad (8)$$

$$U_i = \text{sum of (residual}^2\text{/information)/N} \qquad (9)$$

This is the unweighted mean square fit statistic that is commonly called outfit mean square. This statistic is sensitive to unexpected responses that are relatively distant from the person's or item's measure; that is, a few unexpected responses far from the person or item scale location can cause misfit. Outfit is short for outlier sensitive fit.

A way to diminish the effect of distant unexpected residuals is to weigh nearby residuals so that they have more influence on fit. The squared residual can be weighed by its variance W_{ni}.

$$v_i = \text{sum of } z_{ni}^2 \text{ multiplied by } W_{ni} \text{ / the sum of } W_{ni} \text{ for } n = 1 \text{ to } N. \qquad (10)$$

The variance can also be considered information so:

$$v_i = \text{sum of ((residual}^2\text{/information)*information)/sum of information} \qquad (11)$$

or,

$$v_i = \text{average ((standardized residuals)}^2\text{*information)} \qquad (12)$$

This is called a weighted mean square or infit mean square. Infit is short for information weighted fit. When item and person values are close the individual variance W_{ni} is larger and when they are far apart the variance decreases lessening the impact on the infit mean square.

Infit and outfit mean squares are chi-square fit statistics divided by degrees of freedom that follow chi-square distributions, that is, they are not symmetrical around a mean but are positive values from 0 to infinity with an expected value of 1.0. However, infit mean squares and outfit mean squares can be transformed so that they can also be reported as standardized t values. For smaller *N*-sizes mean square fit statistics can be misleading by showing misfit due to the smaller sample. Standardized *t* fit statistics can be a better gauge of fit. The values are analogous to *z* scores in that they have an expected mean of 0 and values over plus or minus 2.0 are considered to be misfitting as they correspond to *p* values > .05. However, just as mean squares can be misleading for small samples, *t* fit statistics can be misleading for large samples. An *N*-size of 300 is suggested as a maximum value to evaluate misfit when using *t* values (Linacre, 2011, p. 515). As with most statistics, there are no absolute values that serve as a cutoff point. Instead when determining fit there are a number of points to consider so that items that might appear to misfit should not be carelessly discarded or items are retained merely because they appear to fit. Linacre (p. 514) suggested that mean square values between .5 and 1.5 indicate good item fit for rating scales.

PCA of Item Residuals

In addition to providing information about how data fit the Rasch model, item residuals can provide information relating to measure dimensionality. When measuring a single construct, it is expected that item variances relate to the one measure and that additional variance, if detected, is error. By definition random variance is error, or to put it another way, additional item variance is uncorrelated after extracting the variance due to the measure. An assumption in Rasch analysis, as in most statistical models, is that measures are unidimensional (Gustafsson & Aberg-Bengtsson, 2012). Patterns in the residuals suggest that additional dimensions might exist in the data. One way to detect patterns is to do a principal components analysis (PCA) of the residuals.

Another assumption in Rasch analysis is that items are independent (Henning, 1989; 1992). In order for item measures to be additive toward a total score, each item must be statistically independent to function probabilistically. Dependence among sets of items suggests that additional dimensions are being measured (Wainer & Thissen, 1996; Yen, 1984, 1993). Dependence can also result from an item having an influence on another item (Jiao, Wang, & Kamata, 2007). If this is the case then these items will behave more deterministically than expected (Henning, 1989; 1992). Dependency of items also suggests a degree of redundancy in sampling thus lowering construct representation and artificially increasing reliability. Excessive overfit can indicate dependence and correlations among item residuals can also indicate dependence.

If items are measuring on a different dimension beyond that explained by the measure this might be seen as unaccounted variance of the measure. A PCA of the item residuals is done to see if there is any patterning in the residuals that suggests an additional dimension. If item residuals have a strong component loading, that is, some item residuals are highly correlated, then this suggests the scale might not be unidimensional and further examination of the items might be necessary. Principal component eigenvalues can be rounded to a whole number that represents the number of items. Linacre (2011) suggested guidelines for detecting the existence of additional dimensions in the data:

- Variance explained by the measure should be at least 50%
- Eigenvalue units of unexplained variance in the first contrast should be less than 3.0
- Percentage of unexplained variance in the first contrast should be less than 10.0%.

Reliability and Precision

In classical test theory (CTT), an important characteristic of measurement is reliability, that is, how consistently can a group be measured. In CTT observed score variance is equal to true score variance plus error variance and the reliability is the proportion of true score variance to error variance. Adding good items increases reliability because true score variance increase at a faster rate than error variance. When a range of items from very easy to very difficult is given to large samples that have abilities from very low to very high, the patterns become more fixed. For example, the low ability people get the very easy items correct and the very high ability people get the easy up to the difficult items correct. The patterns become fixed because the low ability people are getting the difficult items incorrect and the high ability people are getting the easy items correct. Obviously, however, when different people are taking a test this changes the variance observed so that reliability in CTT refers to the test scores of a particular group and not the test itself (Thompson, 2003).

For an individual member of a group, it is possible to calculate the standard error of measurement (SEM). Just as there is measurement error associated with the group scores, the SEM is the associated measurement around a particular individual score. The score is calculated by SEM = $SD\sqrt{1 - r}$, where SD is the standard deviation of the scores and r = reliability. As can be seen from the formula, there is only one SEM for all the items in the test because it is based only on the standard deviation and the reliability. Therefore, SEMs are not consistent because they vary with different groups because standard deviations and reliabilities can vary.

From a Rasch analysis perspective or an item response theory perspective, what is important is the precision of the test at important points of difficulty for the construct being measured and precision in determining person ability. As Thissen and Orlando (2001) stated, "reliability is frequently not a useful characteristic of an IRT scale-scored test" (p. 117). Consistency or reliability for the group is not aimed at directly but is something like a by-product of precision. To gain precision, there needs to be a gain in information at the level difficulty or ability under scrutiny. (Information is sometimes referred to as information function, information curve or Fisher's information (Embretson & Reise, 2000; Thissen & Orlando, 2001)). In CTT reliability describes group consistency, SEM describes consistency for an individual regardless of where the information in a test is located, Daniel (1999) pointed out the contrast, “... IRT makes clear, in a way that reliability does not, that a test usually is more accurate for some members of a group than for others" (p. 50).

In the Rasch model, the maximum information of an item is at the level of difficulty of the item. Item information is based on the person and item probabilities that are based on the person ability (β) and item difficulty (δ). The item information function (IIF) can be described mathematically in different ways. IIF is equal to the derivative of the probability at a particular difficulty level squared, divided by the probability of getting the item correct multiplied by the probability of getting it incorrect, or more simply, it can also be calculated by multiplying the probability of getting an item correct (P) by the probability of getting it incorrect (Q) (Doran, 2005; Wright & Stone, 1979):

$$I_i = P \times Q \tag{13}$$

$$Q = 1 - P \tag{14}$$

Which then gives us:

$$I_i = P \times (1 - P) \tag{15}$$

The test information function (TIF) at a particular difficulty level is the sum of the item information over all items for a specific difficulty level.

Standard Error of the Estimate (SEE) are analogous to the SEM in CTT. They are the inverse of the square root of the test information function, or:

$$SEE = 1 / (\text{TIF})^{1/2} \tag{16}$$

In fact, with Rasch analysis when the difficulty level and the ability level are equal, the probability of getting the item correct versus getting it incorrect are equal at .5 and the item information is at a maximum of .25 (.5 x (1 - .5) = .5 x .5 = .25). This ease of calculation might be one reason for the claim (Ostini & Nering, 2006) that "test information is rarely employed in the Rasch measurement literature where the issue of measurement precision is subordinated to issues surrounding measurement validity" (p. 30). However, Luo and Andrich (2005) stated that, "Information functions are central in understanding the range in which a scale may be useful" (p. 324). Even when raw scores are ultimately used, constructing test items using item information functions creates measures that give better precision and reliability than randomly choosing from an item pool (Davey & Pitoniak, 2006; Wendler & Walker, 2006). Because item information is near the same person location on the logit scale, a visual inspection of the Wright map (Wilson, 2005) indicates if researchers have an adequate match between items and persons. Researchers inspect the Wright map for person and item range and distribution, floor and ceiling effects, and spacing throughout the map.

Rasch person reliability in Rasch analysis is analogous to Cronbach alpha reliability where the former is calculated based on the Rasch measures where the latter is based on the raw scores. Extreme scores cannot be precisely estimated so these are eliminated or adjusted in Rasch analysis while they are included in calculating reliability with raw scores. In addition, missing scores can create artificially lower scores giving greater range to the raw scores, again inflating reliability. Rasch analysis can impute missing scores for Rasch reliability calculations so that it is usually a more conservative and, thus, more generalizable coefficient than alpha reliability.

Rasch item reliability is analogous to person reliability. For example, when person reliability is low, it suggests that there is a restricted range of person ability or not enough items. When item reliability is low, it suggests that that the range of item difficulties is low or there are not enough respondents. In both cases, poorly targeted persons or items also lowers reliability suggesting that better targeting is needed.

Rasch analysis also calculates a separation index that is related to reliability (Wright, 1996). As with reliability, separation involves true scores, observed scores, and errors, and is calculated by:

Separation = true score standard deviation / error standard deviation (17)

= observed score adjusted standard deviation (18)

/ root-mean-square standard error (RMSE)

Separation coefficient^2 = true score variance / error variance (19)

This can be interpreted as the signal to noise ratio.

Since Observed score variance = true score variance + error variance, and,

reliability = true score variance / observed score variance, then:

Separation coefficient = square root (reliability / (1 – reliability)); or, (20)

Reliability = Separation coefficient^2 / (1 + Separation coefficient^2) (21)

The Rasch separation index like the reliability index can be calculated for both respondents and items. While the reliability as a ratio of true score variance to observed score variance can be difficult to conceptualize, the separation index when squared is more easily conceptualized with its squared values being a signal to noise ratio (Brown & Hudson, 2002; Linacre, 2011). Initially used in communication systems, as explained by Cronbach and Gleser (1964), "Signal to noise ratio compares the strength of the transmission to the strength of the interference" (p. 468). Some researchers suggest that it is easy to understand because it shows the power of the measurement in relation to error. Brennan and Kane (1977) state that signal to noise ratio "is an intuitively appealing and easily interpretable index of the quality of a measurement procedure" (p. 610).

Rating Scale Effectiveness

In addition to investigating individual item statistics, it is possible to investigate scale statistics. As with item statistics, scale guidelines vary, so consideration of the purposes and stakes of the scale should be made. Guidelines considered here are based on Linacre (1999, 2002). Rating scales are more effective if:

- There are at least 10 observations per category.

More than 10 might be necessary but this could be considered a minimum number. This is necessary so that estimates can be precisely calculated and for scale measurement stability.

- There are regular distributions of observations.

Observations should not fluctuate widely between categories, as this would suggest that respondents are interpreting some categories differently. Category distributions should be understandable in terms of what is being measured. If sharp variations are discovered, steps such as collapsing categories might need to be taken.

- Average measures advance monotonically with categories.

This means that average person measures should be higher as categories increase in number. This simply means that in order to be functioning as a rating scale higher values should increase with higher categories or the scale is not working.

- Outfit mean squares are less than 2.0.

This refers to the category fit, not item fit. The interpretation is similar though with outfit values too high, there is too much noise in that category.

- Step calibrations advance.

Steps are the points where categories meet. They are sometimes referred to as thresholds, step calibrations, or tau's. As with average measures in categories, steps need to advance with increasing higher values for an interpretation that the scale is measuring something of increasing value.

- Steps difficulties advance by 1.4 logits with few scale categories and 1 logit with more scale categories.

Three category items should have steps separated by about 1.4 logits for optimum scaling. For items with more categories a rule-of-thumb is that 1 logit indicates good separation. When this criterion is met, the thresholds can be considered to contain information similar to dichotomous items. This means that with optimal separation between steps more information is provided by each item. These separations are not a strict requirement and for six categories it is actually less than 1. Technically speaking, as seen in Table 4, for a six-category scale, threshold distances or threshold advances should be .92 or .69 or greater. When the scale is long enough, these distances between thresholds are not required and might not be helpful (Linacre, 2002). For example, if a seven-item scale has less than optimum distances and gives the information of 30 binary items instead of 35, and is otherwise functioning well, then it might not be helpful to have narrower threshold distances.

The decomposition of rating scales into equivalent binary items has been investigated by Huynh (1994, 1996). He calculated the distance between thresholds for rating scales to be equivalent to the information in binary items.

$$\text{Distance} \geq \tau_{k+1} - \tau_k \geq \ln\left[1 + (n+1)/k(n-k)\right] \tag{22}$$

Where τ_k = threshold instances

$k = 1, \ldots, n\text{-}1$

n = thresholds

For example, when there are three categories, two thresholds or $n = 2$, and a single distance or we could write that the distance is equal to or greater than ln[1 + (3)/1(1)] = ln [1 + 3/1] = ln 4 = 1.3863 ≈ 1.39 ≈ 1.4, as previously noted. Another example for six categories, five thresholds, and four distances, between thresholds one and two needs to be equal or greater than ln[1 + 6/(1(4)) = ln(5/2) ≈ .916. Between thresholds two and three, the distance needs to be equal or greater than ln[1 + 6/2(3)] = ln2 ≈ .693. For thresholds three and four, the distance is again equal to ln2 ≈ .693. As can be seen in Table 4 the distances are different and decreasing for increasing numbers of categories and symmetrical. As previously mentioned, distances between thresholds might not be a helpful guideline and are different for different threshold positions, so Linacre (2006) suggested that a good rule-of-thumb is for thresholds to "advance by one-logit" (p. 1052) as mentioned above. The distances between thresholds need to be checked to see if collapsing categories is warranted. In some cases, collapsing improves scale functioning and sometimes it does not.

Table 4. *Values Thresholds Advance for Dichotomous Equivalents*

Cat.	Thresh.	Number of distances between thresholds	Threshold Distances 1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7
3	2	1	1.39					
4	3	2	1.10	1.10				
5	4	3	0.98	0.81	0.98			
6	5	4	0.92	0.69	0.69	0.92		
7	6	5	0.88	0.63	0.58	0.63	0.88	
8	7	6	0.85	0.59	0.51	0.51	0.59	0.85

Note. Cat. = Categories; Thresh. = Thresholds.

- Step difficulties advance by less than 5.0 logits

If steps are too far apart it disperses the information provided by the steps. Precision is lost between steps. If steps are too far apart then individual items cannot optimally contribute to scale information. Respondents in the middle of the steps have no matching items, so they have more measurement error. This effect is similar to reducing the number of items.

Confirmatory Factor Analysis and Structural Equation Modeling

Confirmatory factor analysis (CFA) is a statistical analysis that shows relationships among latent or measured variables. In CFA, researchers have an a priori theory about how variables relate to each other. A model is constructed based on the theory and then the model can be tested against the data. The model is not really “confirmed” in the sense of being true or even the best model. If the data fit the model, this shows that the model is a plausible one and can show the strengths of the relationships among the variables.

Like CTT where observed score variance is composed of true score variance and error variance, in CFA variables are assumed to be composed of variance influenced by the construct being measured and error variance. In CFA, however, it is possible to detect correlations among errors. This would show that the errors are not randomly generated but that some underlying factor is shared that is not accounted for by the measure.

As with the Rasch model, CFA generates a set of expected values that can be compared with actual values. The values in CFA are expected correlations or covariances that can then be compared with the actual correlations or covariances to produce residuals. The smaller the residual values the better the data fit the CFA model, in other words, goodness of fit increases.

There are a number of CFA fit indexes. This study uses the Chi-Squared test, Akaike Information Criterion (AIC), the comparative-fit index (CFI), and the root-mean-square of approximation (RMSEA) as recommended by Byrne (2010) and Schumacker and Lomax (2010).

The chi-squared test in an absolute fit index that compares the difference between the expected and actual covariance matrices. An acceptable value for the chi-squared test is a non-significant value ($p > .05$). One problem with the chi-squared test is that it is sensitive to sample and model size. The AIC is a parsimony-based fit index that indicates model fit and model parsimony. There are no absolute values for model fit but smaller values indicate a better model. The RMSEA is not sensitive to model size. Various values are suggested for model fit with RMSEA < .05 indicating close fit .05 < RMSEA < .08 indicating adequate fit, .08 < RMSEA < .10 mediocre fit, and, .01 < RMSEA as poor fit (Browne & Cudeck, 1993; Byrne, 2010; Kaplan, 2000; Schumacker & Lomax, 2010). In addition, the 90% confidence intervals were reported. Model comparisons were not the primary objectives of this study. The research questions were concerned with examining relationships among the latent variables. Model fit is necessary to interpret a structural model but the regression weights were the primary concern. Values between .05 and .10 were considered acceptable for this study. The CFI is based on the comparison of the model with the data while adjusting for sample size. CFI > than .95 indicate good fit and greater than .90 considered acceptable model fit (Hu & Bentler, 1999). The $\Delta CFI \leq .01$, the difference in CFI values, was used as the criterion for decision making among models and for the invariance testing in the cross-validation study.

Cross-Validation

Cross-validation in structural equation modeling analysis where once a model is specified with a sample, a second sample is tested with the same model. Cross-validation is also called invariance testing or equivalence testing (Byrne, 2010). There were two different proficiency groups in this study and the CFA and SEM were done separately for the different groups. One group had taken the TOEIC Bridge (n = 221) and one had taken the TOEIC (n = 275). Theoretically, because both tests measure second language proficiency, a cross-validation study can show equivalence of causal structures. Cross-validation provides some evidence that the model is more generalizable than for a single study. This provides a stronger argument for research question 5 relating positive self-concept and positive L2 self with L2 proficiency.

There are different levels of invariance (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). Measurement variables, paths, and factorial structure are tested to see if they replicate in different groups. Configural invariance is in the measurement model step where measured variable and latent constructs have the same patterns of loadings. Weak factorial invariance also called metric invariance or pattern invariance is an additional check to test that relative factor loadings across groups are invariant. Strong factorial invariance adds the structural covariances or factorial invariance. Finally, strict factorial invariance means that the structure and measurement residuals are invariant. Byrne (2010) noted that the last step is rarely met in practice.

Assumptions about CFA and SEM

There are several assumptions that must be met before conducting a CFA (Kline, 2011; Tabachnick & Fiddell, 2007). Assumptions of reliability and interval level measurement were treated in the instrument validation chapter. Other assumptions of CFA and SEM are presented here.

Assumptions about sample size and missing data: A rule-of-thumb mentioned by Kline (2011) for an ideal minimum sample size is a ratio of 20:1 of participants to number of parameters estimated while a ratio of 10:1 might be considered the minimum acceptable. Kline also noted that sample sizes of 200 are often considered a minimum number. In either case, this study met the assumptions of adequate sample size. In Rasch analysis, isolated cases of missing data can still generate precise estimates of person measures, consequently there were no missing data.

Assumptions about univariate and multivariate normality; Normality statistics were previously calculated in the instrument validation chapter and found to have met normality requirements. In addition graphs were visually inspected and found to be normally distributed.

Assumptions about outliers: Outliers are extreme values that can distort the analysis. The data were examined for both univariate and multivariate statistical outliers. Univariate outliers generally have a *z*-score greater than 3.29 but this depends on sample size. Large samples can naturally have a few of these values. Values greater that 3.29 were brought in 3.0 using *z*-scores. This is a conservative value. The number of outliers for the modeled variables are given in Table 5.

Table 5. *Number of Outliers > 3.29 brought in to 3.0*

Measure	Number of outliers
Curious	6
Flourishing	6
Hope: Agency	5
Hope: Pathways	8
L2 Interest	4
L2 Passion	10
Mastery orientation	6
L2 listening SE	8
L2 reading SE	10
L2 speaking	9

Note. Out of a total *N* = 539; SE = Self-Efficacy

Multivariate outliers were examined by calculating the Mahalanobis distance (D^2). Twenty-four multivariate outliers were found using the procedure in Bryne (2010) and removed from the analysis for a remainder of 514 cases. It was not possible to determine the cause for all the multivariate outliers. In many cases, the multivariate outliers included one or more univariate outliers, that is, the multivariate outliers often included a case that had variance curtailed by bringing in an univariate outlier to three standard deviations. Many of the univariate outliers in turn seemed to be caused by an extreme response style, that is, often choosing the first or last categories and avoiding the middle categories.

Another assumption is that relationships among variables are linear: This was checked by inspections of bivariate scatterplots of the variables. Inspection revealed that the scatterplots were sufficiently linear.

Assumptions of multicollinearity and singularity must also be checked: Multicollinearity refers to extremely high correlations between variables and singularity refers to perfect correlations. This can be checked by examining the correlation matrix. No instances of multicollinearity or singularity were found.

Summary

Four analyses are conducted: Rasch analyses, correlational analyses, confirmatory factor analysis (CFA), and structural equation modeling (SEM). First, a Rasch rating scale analysis was carried out for item and scale measures. A principal components analysis (PCA) of the Rasch measure residuals was also carried out for the scales for the three different levels. When poorly fitting were found they were dropped. When criterion suggested by Linacre (1999; 2002) were not met, categories were collapsed if they improved rating scale functioning.

A second stage was a correlational analysis done for global, L2 domain, and motivational measured variables with the modeled variables and peripheral variables. This analysis was done to provide evidence of convergent and divergent validity of the measures used.

At a third stage of this study, CFA was used to determine the relationship of the latent variables with the measured variables and CFA model fit. A fourth stage of the analysis was to run the full structural model composed of the two latent constructs and the objective outcome measure of proficiency scores.

CHAPTER 4

PRELIMINARY ANALYSIS: VALIDITY EVIDENCE

The preliminary analysis for this study was done in three stages: main instruments for modeling, peripheral instruments, and relationships among main and peripheral variables. In the first stage the main instruments are examined for validity evidence of the individual scales and items. In other words, evidence is examined to show whether or not the items fit the Rasch model and whether the scale structure fits the Rasch model. This is done to ensure that the scale is indeed capable of actually measuring respondents (Bond & Fox, 2015; Boone, Staver, & Yale, 2014; Engelhard, 2013; Wilson, 2005). In the second stage (Chapter 5) the same analysis is done for the peripheral but related instruments. In the first two stages, the analyses are focused on internal characteristics of the scale; that is, do the items work together to create an individual scale to measure an intended dimension. The third stage analysis (Chapter 6) is focused on providing validity evidence that is external to the scale (Cronbach & Meehl, 1955; Loevinger, 1957). The relationships among the main instruments and peripheral instruments are examined for convergent and divergent validity evidence.

In this chapter the preliminary analyses of the main instruments are described. Rasch analysis of the items is done and then analysis of the scale as a whole. Rasch analysis for this study was done with *Winsteps* software (Linacre, 2011). Item analysis includes descriptions of fit statistics, item difficulties, item standard errors, and item-person map as depicted in the Wright maps. Scale analyses include principle components analysis (PCA) of the Rasch residuals for descriptions of unidimensionality given by variance explained by the measures, eigenvalues of unexplained variance in the first contrast, and percentage of unexplained variance in the first contrast. Scale structure is examined through category observation counts and ascending orders of endorsability, and separation of thresholds. Guidelines (explained in Chapter 3) are compared with the empirical values obtained from the Rasch analyses. In addition, Rasch person reliability and separation, as well as traditional descriptive statistics are given.

Modeled Positive Self Variables

Curiosity and Exploration Inventory II

The Rasch analysis item statistics are presented in Table 6. As can be seen, fit statistics range from a lower value of .85 to an upper value of 1.25. This is well within the targeted range of 0.5 to 1.5 suggested by Linacre (2011). Item standard errors of .04 and .05 shows that there are enough respondents in the measurement range to provide fairly precise item measures. The item-measure correlations show that items are contributing to the scale measure. The Wright map, or item-person map, in Figure 10 shows that items and persons were well targeted although items as a whole were slightly easy to endorse for this group of students. There were a few extreme cases but given the large sample size of over 500 this was to be expected. These item statistics suggest that the items are functioning properly and contributing to the overall scale.

Table 6. *Rasch Descriptive Statistics for Curiosity and Exploration Inventory Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
6	.82	.04	1.19	1.19	.58
3	.76	.04	1.04	1.04	.56
7	.56	.04	.96	.95	.64
8	.28	.04	1.11	1.10	.67
1	.01	.04	.76	.77	.65
2	-.19	.04	.90	.89	.68
9	-.33	.04	.90	.89	.65
10	-.41	.05	1.17	1.25	.63
4	-.74	.05	.88	.85	.67
5	-.78	.05	1.07	1.05	.57

The results of the unidimensionality analysis in Table 7 showed that the variance explained by the measure was 49.5%. This just misses the criterion of 50% suggesting some dimensionality in the data. As explained in literature review section, this scale measures the stretching and embracing of new knowledge and experiences so very minor additional dimensions are to be expected. The eigenvalue units of unexplained variance in the first contrast was 1.7, well below the 3.0 criterion (Linacre, 2011). The percentage of unexplained variance in the first contrast was 8.8%, below the 10% criterion. These values suggest that even though there might be additional dimensionality in the data that it is not enough to harm effective measurement.

```
------------------------------------------------------------------
More curious students|  More difficult to endorse items
    4             .  +
                     |
                     |
                     |
                  .  |
                     |
    3             .  +
                     |
                  .  |
                  .  |
                     |
                  . T|
    2            .#  +
                 .#  |
               .###  |
                 .#  |
               .###  |
              ##### S|T
    1          ####  +
           .#######  |  C6:  Like doing frightening things
         .#########  |  C3:  At best doing something challenging
              .####  |S C7:  Always looking to challenge myself
        ########### M|  C8:  Prefer excitingly unpredictable jobs
             ######  |
    0  ############  +M C1:  Seek information in new situations
         .#########  |  C2:  Enjoys uncertainty of life
         .#########  |  C9:  Frequently challenge myself
                 .# S|S C10: Embrace unknown people, events, places
              .####  |  C4:  Looking for new things/experiences
                 .#  |  C5:  Challenges help grow and learn
   -1          .###  +
                 .#  |T
                  . T|
                  .  |
                  #  |
                  .  |
   -2                +
                     |
                     |
                     |
                  .  |
Less curious students|  Easier to endorse items
------------------------------------------------------------------
```

Figure 8. Wright map of Curiosity and Exploration Inventory. Each “#” represents 5 persons. Each “.” represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

Table 7. *Unidimensionality Analysis for Positive Self Variables*

Variable	Variance explained by measures %	Eigenvalue of the 1st contrast	Unexplained variance in 1st contrast %
Curiosity	49.5	1.7	8.8
Curiosity (collapsed)	50.4	1.7	8.8
Flourishing	49.1	1.5	9.6
Hope (combined)	49.8	1.6	10.1
Hope (pathways)	60.8	1.5	15.1
Hope (agency)	55.3	1.6	18.1

The classical statistics for this scale were a mean of .33 logits with a standard deviation of .87. As showed in Table 11 skewness and kurtosis values have acceptable values. The Rasch person reliability was .81 and person separation index was 2.06. The Rasch person reliability is given instead of the alpha reliability for the raw-data scores because raw-data scores are often inflated. Rasch person reliability does not count extreme scores and is usually a more conservative indicator of reliability. For example, the raw-data has a classical alpha reliability of .85. Although here the reliability difference of .04 is not a large difference, Rasch person reliability is usually a more generalizable estimate reliability. This is because the ordinal nature of the raw scores and the extreme scores for the raw data tend to inflate estimates compared to the interval nature of the Rasch measures.

Values from the Rasch analysis of scale functioning and structure are given in Table 8. As can be seen, the category with the fewest observations was 305, well above the guideline minimum of 10. The observed values distribution through the categories exhibits no abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed values are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.33, much less than the maximum criterion of 2.0. The threshold values all advance monotonically; however, the extreme thresholds do not meet the suggested value of 1. This suggests that the extreme categories might function more effectively with the extreme categories collapsed. However, a couple points should be noted. One, for scales with six categories the actual minimum value for categories to be decomposed into dichotomous variables is less than 1. Two, there is actually no strict requirement that scales need to be decomposable into dichotomous items (Linacre, 2002). In other words, the guidelines are suggested for an optimum level of information but in practice this might not hold. So if the amount of information for six categories returns the amount of information for five dichotomous items that would be optimal as mentioned in Chapter 3 (Linacre, 2002, 2006). However, if it returns the amount of information of four dichotomous items, it would be less than optimal but might be acceptable. In the case of thresholds advancing more than 5 logits, this indicates that information is too widely distributed for adequate measurement. To test if there is any improvement, the curiosity and exploration inventory was analyzed again with the first and second categories collapsed.

Table 8. *Rasch Rating Scale Functioning for Curiosity and Exploration Inventory*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	305	-.95	1.33	(none)
2 Disagree	547	-.64	1.07	-1.44
3 Slightly disagree	1305	-.17	.83	-1.22
4 Slightly agree	1562	.40	.86	-.04
5 Agree	954	.98	.87	1.15
6 Strongly agree	704	1.59	1.06	1.54

Re-Analysis of Curiosity and Exploration Inventory with Categories Collapsed

The effects of collapsing categories 1 and 2 for this scale are analyzed here. After collapsing, the fit statistics showed some small improvements as showed in Tables 9 and 7. The variance explained by the measures improved to 50.2%, a slight difference from 49.5%. The unexplained variance in the first contrast was similar with the same eigenvalue units of 1.7 and expressed as a percentage decreased from 8.8% to 8.6%. The collapsed categories had a better outfit mean-squared value of 1.19 in contrast with 1.33.

Table 9. *Rasch Rating Scale Functioning for Curiosity and Exploration Inventory (Categories 1 and 2 Collapsed)*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
2 (Strongly) disagree	852	-1.30	1.19	(none)
3 Slightly disagree	1305	-.70	.87	-1.43
4 Slightly agree	1562	-.06	.90	-.52
5 Agree	954	.60	.92	.75
6 Strongly agree	704	1.27	1.07	1.20

The Rasch item statistics for the curiosity and exploration inventory are presented in Table 10. Fit statistics ranged from .79 to 1.27, well within the targeted range of .5 to 1.5. The maximum misfit value increased slightly from 1.19 to 1.27. Item standard errors showed a slight increase for some items from .04 to .05 but these standard errors show fairly precise item measures. The Wright map in Figure 9 shows that items are well matched to the group.

Table 10. *Rasch Descriptive Statistics for Curiosity and Exploration Inventory (Categories 1 and 2 Collapsed)*

Item number	Measure (Logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
6	.90	.05	1.27	1.25	.55
3	.86	.05	1.05	1.04	.54
7	.62	.05	1.00	.98	.63
8	.28	.05	1.14	1.12	.66
1	.02	.05	.79	.81	.66
2	-.20	.05	.92	.92	.68
9	-.36	.05	.91	.90	.66
10	-.48	.05	1.09	1.19	.64
4	-.81	.05	.80	.80	.70
5	-.84	.05	1.04	1.04	.60

The statistics for this five-category scale were a mean of .33 logits with a standard deviation of .87. As seen in Table 11 skewness and kurtosis values were acceptable. The Rasch person reliability was .82, the Rasch person separation index was 2.10, the Rasch item reliability was .99, and the Rasch item separation was 12.11. Thus, there was a slight improvement from collapsing categories one and two, from Rasch person reliability .81 to .82 and person separation of 2.06 to 2.10.

```
-----------------------------------------------------------------
More curious students| More difficult to endorse items
    4                +
                     |
                  .  |
                     |
                     |
                     |
                     |
    3             .  +
                     |
                     |
                  .  |
                     |
                  .  |
                  .  |
    2                +
                  . T|
                 .#  |
                 .#  |
                 .#  |
                 ##  |T
                 ##  |
    1          .###  +
                .## S|  C3:  At best doing something challenging
                .##  |  C6:  Like doing frightening things
              .####  |S C7:  Always looking to challenge myself
          .########  |
              .####  |  C8:  Prefer excitingly unpredictable jobs
             .#####  |
    0          .###  +M C1:  Seek information in new situations
       .########### M|  C2:  Enjoys uncertainty of life
             .#####  |  C9:  Frequently challenge myself
            .######  |  C10: Embrace unknown people, events, places
            .######  |S
             .#####  |  C4:  Looking for new things/experiences
          .########  |  C5:  Challenges help grow and learn
   -1         .####  +
                 .# S|
                 ##  |T
                  .  |
                 .#  |
                 ##  |
                .##  |
   -2               T+
                  #  |
                     |
                  .  |
                     |
                  .  |
                     |
   -3                +
                  #  |
Less curious students|  Easier to endorse items
-----------------------------------------------------------------
```

Figure 9. Wright map of Curiosity and Exploration Inventory (categories 1 and 2 collapsed). Each “#” represents 5 persons. Each “.” represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

Table 11. *Descriptive Statistics for Modeled Positive Self Measures*

Variable	*k*	*M*	*SE*	95% *CI*	*SD*	*Skew*	*Kurtosis*	PR	PS
CEI	10	.33	.04	[.26, .41]	.87	.25	.66	.82	2.10
Flourishing	8	.27	.04	[.20, .34]	.87	.10	.76	.75	1.72
Hope (all)	8	.36	.05	[.26, .45]	1.13	.15	.48	.81	2.10
Hope/Agency	4	.44	.06	[.33, .56]	1.39	.30	1.00	.71	1.57
Hope/Paths	4	.44	.07	[.29, .59]	1.73	.13	.62	.77	1.81

Note. CEI = Curiosity and Exploration Inventory; SES = Standard Error of Skewness; SEK = Standard Error of Kurtosis; PR = Person Reliability; PS = Person Separation; SES = .11; SEK = .21.

Flourishing Scale

The Rasch analysis item statistics are presented in Table 12. Fit statistics range from .73 to 1.38, well within the targeted range of .5 to 1.5. Item standard errors of .05 show fairly precise item measures. The item-measure correlations show that the items are contributing variance toward the scale measure. The Wright map in Figure 10 shows that items are slightly easy to endorse as a group. The items and persons are suitably targeted. There are a few outliers at the extremes of the group but this is expected with large samples.

Table 12. *Rasch Descriptive Statistics for Flourishing Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
8	1.34	.05	1.01	1.01	.62
5	.25	.05	.80	.82	.67
6	.08	.05	.82	.84	.62
1	-.04	.05	.95	.95	.69
7	-.17	.05	1.38	1.36	.54
4	-.19	.05	.73	.73	.68
3	-.33	.05	1.50	1.14	.58
2	-.93	.05	1.14	1.16	.57

```
------------------------------------------------------------------
  Students higher     |  More difficult to endorse items
  in flourishing      |
    4             .+ |
                     |
                     |
                  .  |
                     |
    3                +
                  .  |
                     |
                  .  |
                  .  |
                  .  |
    2               T+
                  .  |
                 .#  |
                 .#  |
                 .#  |  Fl8: People respect me
           .####### S|T
    1         .####  +
             .#####  |
              #####  |S
              #####  |
            .###### M|
      .############  |  Fl5: Competent & capable; Fl6: Good life
    0      .#######  +M Fl1: Purposeful & meaningful life
             .#####  |  Fl4: Contribute to others; Fl7 Optimistic
              .####  |  Fl3: Engaged & interested daily activities
               .###  |
               .### S|S
              .####  |
   -1             #  +  Fl2: Social relationships supportive
                  .  |T
                .##  |
                  . T|
                  .  |
                  .  |
   -2             .  +
                     |
                  .  |
                     |
                     |
   -3                |
  Students lower     |  Easier to endorse items
  in flourishing     |
------------------------------------------------------------------
```

Figure 10. Wright map of Flourishing. Each "#" represents 6 persons. Each "." represents 1 to 5 persons. M = Mean, S = 1 SD, T = 2 SD.

The results of the unidimensionality analysis showed in Table 7 showed that the variance explained by the measures was 49.1%. This just misses the criterion of

50% suggesting some dimensionality in the data. As explained in the literature review section, this flourishing scale is an omnibus measure of positive constructs related to overall general positive psychological health. It was expected that very minor dimensions might exist but that they are not enough to harm measurement. The PCA of the Rasch residuals yielded a first contrast of 1.5 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 9.6%, below the 10.0% criterion. The empirical data suggest that even with some additional dimensionality in the data, it is not harming effective measurement.

Values from the Rasch analysis of scale functioning and structure are given in Table 13. As can be seen, the category with the fewest observations was 248, well above the guideline minimum count of 10. The observed counts distributed through the categories exhibits no abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.29, much less than the maximum criterion of 2.0. The threshold values all advance monotonically, however, the extreme thresholds do not meet the suggested value of 1.

The statistics for this scale were a mean of .27 logits with a standard deviation of .86 logits. As showed in Table 7, skewness and kurtosis values were acceptable. The Rasch person reliability was .75, the Rasch person separation index was 1.72, the Rasch item reliability was .99, and the Rasch item separation was 12.43.

Table 13. *Rasch Rating Scale Functioning for Flourishing*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	248	-1.28	1.29	(none)
2 Disagree	461	-.78	.92	-1.65
3 Slightly disagree	1007	-.14	.86	-1.19
4 Slightly agree	1418	.44	.94	-.18
5 Agree	709	.96	.87	1.36
6 Strongly agree	452	1.51	1.12	1.66

The reliability is only moderate on this scale but as explained previously, this is an omnibus measure of positive psychological functioning and differs from other measures in the sense that it is trying to account for perspectives that differ slightly. To take an example from language testing, a general measure of language proficiency might be composed of items that measure language through different skills such as listening and reading. Language proficiency can be considered as a single measure composed of differing language skills for the purpose of having a single overall measure. For other purposes, such as determining which language skill benefits from a particular teaching treatment, studying the subcomponents separately is needed. Another example could be made of reading skill, this might be measured through items that measure vocabulary knowledge, grammar knowledge, and discourse knowledge. For the purpose of determining reading skill this might be a single measure. For other purposes, it might be necessary to isolate the subcomponents into different types of knowledge. In other words, there is some multidimensionality in data of mental measurements. However, for the purposes of a single scale the question is, does the multidimensionality harm the single dimension being measured? In other words, is the single scale unidimensional enough (Diener, Inglehart, & Tay, 2013; Slocum-Gori & Zumbo, 2011; Slocum-Gori, Zumbo, Michalos, & Diener, 2009)? If items do not vary enough there is a danger of construct underrepresentation. If items vary too much then unidimensionality is threatened. As Humphreys (1962, p. 483) stated, "Controlled heterogeneity is the goal of test construction rather than maximum homogeneity in the statistical sense." This means that items need to be different enough for construct representation but not so similar that they are redundant. With very similar item-types, they will maximize homogeneity in the sense that they will correlate very highly together but they might also underrepresent the construct measured. For the purposes of this study, the reliability and unidimensionality of this measure are considered adequate.

One item was included for possible use in the flourishing scale that related to the Japanese term, *ikigai*, meaning purpose in life. The item had a high point-measure correlation of .72 with the rest of the flourishing measure. I decided not to use it to keep the flourishing scale intact as it was written to make it possible to compare with other uses of the scale. Including a Japanese specific term would have made the scale less generalizable so it was not included in any further analysis.

Hope Scale

As noted above, the hope construct consists of mental willpower or agentic thinking and waypower or pathways thinking for goals (Snyder, 1994). Agentic thinking refers to the agency or the mental energy and determination to reach goals. Pathways thinking refers to the mental ability to find ways to reach goals. These components usually co-vary (Lopez, 2013; Snyder, 1994, 2002). The analysis of the hope construct commences with an analysis of the eight item together, then separate analyses are conducted with the two components with four items each.

The Rasch item statistics for the whole scale are presented in Table 14. Fit statistics were well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 11 shows that items are slightly easy to endorse as a group. The items and persons are suitably targeted. There are a few outliers but this is expected in large samples.

Table 14. *Rasch Descriptive Statistics for Hope: Agentic and Pathways Thinking Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
7	.63	.05	1.38	1.39	.56
1	.34	.05	.95	.95	.67
4	.17	.05	.71	.73	.75
3	.16	.05	.95	.95	.73
6	-.23	.05	1.16	1.15	.65
5	-.28	.05	.92	.92	.74
8	-.28	.05	.89	.88	.72
2	-.51	.05	1.00	.98	.67

```
------------------------------------------------------------------
Students      .  +  More difficult to endorse items
higher in hope   |
                 |
                 |
              .  |
                 |
 4               +
                 |
              .  |
                 |
                 |
              .  |
 3               +
             .#  |
             .#  |
                T|
              #  |
            .##  |
2           .##  +
            .##  |
          #####  |
           .### S|
        .######  |
        .######  |
1        .#####  +
      ########   |
     .########   |T Hope7: Successful in life
    ##########   |
      .####### M|S Hope1: Many ways to get out of jam
     .########   |  Hope3: Many ways to get important things
0   .#########   +M Hope4: Find a way to solve problem
      ########   |  Hope6: Past prepared me for future
        .#####   |S Hope5: Pursue goals; Hope8: Meet goals
         .####   |  Hope2: Lots of ways around any problem
          .###   |T
       .###### S|
-1         ###   +
            .##  |
             ##  |
             .#  |
             .#  |
            .## T|
-2            .  +
              .  |
              #  |
                 |
              .  |
 Students     .  |
 lower in hope   +  Easier to endorse items
------------------------------------------------------------------
```

Figure 11. Wright Map of Hope: Agentic Thinking and Pathways Thinking Combined. Each “#” represents 4 persons. Each “.” Represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

The results of the unidimensionality analysis in Table 7 showed that the variance explained by the measures was 49.8%. This just misses the criterion of 50%, suggesting some dimensionality in the data although this might be due to lack of person separation. This suggests that the willpower and waypower components are closely but differently related to each other as theorized in the literature. The PCA of the Rasch residuals yielded a first contrast of 1.6, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 10.1%, just above the 10.0% criterion. The data suggest that combining the two components into a single scale might provide a hope measure that is unidimensional enough for some purposes but that for other purposes separating the two components might be warranted.

Values from the Rasch analysis of scale functioning and structure are given in Table 15. As can be seen, the category 1 had the fewest observations with a count of 159, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.27, much less than the maximum criterion of 2.0. The threshold values all advance monotonically, however, the extreme thresholds do not meet the suggested value of 1.0. The threshold values are acceptable, but it should be remembered that for other purposes, it might be desirable to collapse the categories.

Table 15. *Rasch Rating Scale Functioning for Hope (Combined Willpower and Waypower)*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	159	-1.40	1.27	(None)
2 Disagree	430	-.88	.99	-2.20
3 Slightly disagree	1096	-.22	.86	-1.44
4 Slightly agree	1517	.51	.88	-.16
5 Agree	721	1.20	.93	1.56
6 Strongly agree	372	1.88	1.14	2.23

The statistics for this scale were a mean of .36 logits with a standard deviation of 1.13. As seen in Table 11 skewness and kurtosis values were acceptable. The Rasch person reliability was .81, the Rasch person separation index was 2.10, the Rasch item reliability was .98, and the Rasch item separation was 6.79.

Although scale statistics were acceptable when the two hope components were analyzed together, they are theoretically distinct components. Therefore, in addition to analyzing them as a combined scale, the two components were analyzed separately as an agency or willpower component and as a pathways or waypower component.

Hope: Agentic Thinking Scale

The Rasch item statistics for the agentic thinking scale are presented in Table 16. Fit statistics range from .82 to 1.31, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 12 shows that items are easy to endorse as a group. The items are grouped together but persons distribution spans several logits. There are a few outliers but this is expected for large samples.

Table 16. *Rasch Descriptive Statistics for Hope: Agentic Thinking Component Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.74	.05	1.28	1.31	.65
2	-.21	.05	1.01	1.01	.74
3	-.26	.05	.82	.82	.80
4	-.27	.05	.86	.87	.76

The results of the unidimensionality analysis in Table 7 showed that the variance explained by the measures was 55.3%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of the Rasch residuals yielded a first contrast of 1.6, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 18.1%, above the 10.0% criterion. Agentic thinking as a subcomponent of hope is measured on a short scale of four items. Given that hope is a global self-construct and the shortness of the scale, the close

```
-----------------------------------------------------------------
Students higher  |  More difficult to endorse items
in hope: agency  |
4            .#  +
                 |
                 |
                 |
                 |
                 |
             .#  |
3               T+
                 |
            .##  |
              .  |
                 |
                 |
           .###  |
2                +
        .######  |
                S|
                 |
        #######  |
                 |
                 |
1     #########  +
                 |T
      .########  |  HA3: Successful in life
                 |
                M|S
   .###########  |
                 |
0   .##########  +M
                 |  HA1: Pursue goals  HA2: Past prepared me for future
              .  |  HA4: Meet goals
      .########  |S
                 |
          .####  |
                S|T
              .  +
          .####  |
                 |
            .##  |
              .  |
                 |
            ###  |
-2               +
             .# T|
                 |
                 |
              .  |
                 |
                 |
-3               +
              .  |
                 |
                 |
                 |
                 |
                 |
-4            .  +
Students lower in|  Easier to endorse items
hope: agency     |
-----------------------------------------------------------------
```

Figure 12. Wright Map of Hope: Agentic Thinking variable. Each “#” represents 6 persons. Each “.” represents 1 to 5 persons. M = mean, S = one standard deviation, T = two standard deviations from the mean.

grouping of the items and the additional percentage of unexplained variance in the first contrast is not a problem.

Values from the Rasch analysis of scale functioning and structure are given in Table 17. As can be seen, the category 1 had the fewest observations with a count of 81, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.43, less than the maximum criterion of 2.0. The threshold values all advance monotonically, with all thresholds near or above the suggested value of 1.0. This shows that although the items seem to exhibit a close measure grouping, the categories are sufficiently spread apart to provide adequate measurement.

Table 17. *Rasch Rating Scale Functioning for Hope: Agentic Thinking*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	81	-1.66	1.43	(None)
2 Disagree	229	-1.09	.98	-2.58
3 Slightly disagree	508	-.32	.92	-1.44
4 Slightly agree	743	.57	.85	-.21
5 Agree	381	1.45	.97	1.64
6 Strongly agree	204	2.25	1.07	2.59

The statistics for this scale were a mean of .44 logits with a standard deviation of 1.39. As seen in Table 11, skewness and kurtosis values were acceptable. The Rasch person reliability was .71, the Rasch person separation index was 1.57, the Rasch item reliability was .98, and the Rasch person separation index was 7.70.

Hope: Pathways Thinking Scale

The Rasch item statistics for the pathways thinking scale are presented in Table 18. Fit statistics range from .86 to 1.10, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 13 shows that items are easy to endorse as a group. The items are grouped together but persons distribution spans several logits. There are a few outliers but this is expected for large samples.

Table 18. *Rasch Descriptive Statistics for Hope: Pathways Component Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.42	.06	.94	.93	.78
2	.18	.06	.86	.86	.79
3	.17	.06	1.06	1.06	.80
4	-.76	.06	1.10	1.07	.75

The results of the unidimensionality analysis showed in Table 7 showed that the variance explained by the measures was 60.8%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of the Rasch residuals yielded a first contrast of 1.5, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 15.1%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher  |  More difficult to endorse items
in hope: pathways|
             .#  |
                 |
                 |
                 |
4                +
             .#  |
                T|
                 |
                 |
              #  |
3                +
                 |
            .##  |
                 |
         .###### |
                 |
2               S+
          .####  |
                 |
                 |
    .##########  |
                 |
1 .############  +
              .  |T
                 |
      .########  |S
                M|  HP1: Many ways to get out of jam
                 |  HP3: Many ways to get important things
0    .#########  +M HP4: Find a way to solve problem
                 |
                 |
      .########  |S
                 |
                 |T HP2: Lots of ways around any problem
-1       ######  +
              . S|
                 |
         .#####  |
                 |
            .##  |
-2               +
              .  |
            .##  |
                 |
            .##  |
                T|
-3               +
              .  |
                 |
                 |
              .  |
                 |
-4               +
                 |
                 |
                 |
              .  |
Students lower in|  Easier to endorse items
hope: pathways   |
------------------------------------------------------------------
```

Figure 13. Wright map of Hope: Pathways Thinking variable. Each “#” represents 6 persons. Each “.” represents 1 to 5 persons. M refers to mean, S to one standard deviation, and T refers to two standard deviations from the mean.

Waypower as a subcomponent of hope is measured on a short scale of four items. Given that hope is a global self-construct and the shortness of the scale, the close grouping of the items and the additional percentage of unexplained variance in the first contrast is not a problem.

Values from the Rasch analysis of scale functioning and structure are given in Table 19. As can be seen, the category 1 had the fewest observations with a count of 78, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.36, less than the maximum criterion of 2.0. The threshold values all advance monotonically with the all thresholds above the suggested value of 1.0. This shows that although the items seem to exhibit a close measure grouping, the categories are sufficiently spread apart to provide adequate measurement.

Table 19. *Rasch Rating Scale Functioning for Hope: Pathways Thinking*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	78	-2.61	1.17	(None)
2 Disagree	201	-1.64	.96	-3.12
3 Slightly disagree	588	-.48	.86	-2.07
4 Slightly agree	774	.76	.84	-.12
5 Agree	340	1.90	.93	2.08
6 Strongly agree	168	2.86	1.36	3.24

The statistics for this scale were a mean of .44 logits with a standard deviation of 1.73. As seen in Table 11, skewness and kurtosis values were acceptable. The Rasch person reliability was .77, the Rasch person separation index was 1.81, the Rasch item reliability was .98, and the Rasch item separation index was 7.34.

Modeled Positive L2 Self Scales

Interested-in-L2 Self Scale

The Rasch analysis item statistics are presented in Table 20. Fit statistics range from .70 to 1.69, outside the targeted range of .5 to 1.5. One item does not fit the model well and also shows a lower point-measure correlation. Other items show adequate fit to the model and the high item-measure correlations show that the items are contributing variance toward the scale measure. Item standard errors of .05 and .06 show fairly precise item measures. The person reliability is .87, the Rasch person separation is 2.59, the Rasch item reliability was .99, and the Rasch item separation was 9.38 showing that the scale might be acceptable but might be improved with the elimination of the misfitting item.

The misfitting item stated, "One English class passes like a moment." Upon reflection, this item might be interpreted to include other characteristics of an English class such as having friends in class or liking the teacher. This item was meant to tap into the idea of being interested in English makes a student engaged to the point where consciousness of time is lost. This interpretation would probably be

Table 20. *Rasch Descriptive Statistics for the Interested-in-L2 Scale Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.80	.05	.71	.72	.82
6	.61	.05	.96	.94	.81
2	.60	.05	1.54	1.69	.66
8	.06	.06	.83	.87	.77
3	-.25	.06	.70	.71	.82
4	-.47	.06	.91	.92	.78
5	-.55	.06	1.18	1.10	.77
7	-.78	.06	1.19	1.19	.70

endorsed by only the most interested students. However, students seem to be interpreting this item in different ways and it is only the third most difficult item to endorse. In addition to misfitting the Rasch model, this item at .60 logits is nearly the same as item 6 at .61 logits so information for this scale is already being provided for by another item and it is not needed. Considering the problems with this item and the lower point-measure correlation it was decided to remove this item from the scale. The items were then reanalyzed with the further use of this scale using seven items.

Table 21. *Descriptive Statistics for Positive L2 Self Scales*

Variable	*k*	*M*	*SE*	*95% CI*	*SD*	*Skew*	*Kurtosis*	P.R.	P.S.
Interested L2	8	1.18	.08	[1.03, 1.33]	1.76	0.33	1.09	.87	2.59
Interested L2 (item removed)	7	1.42	.09	[1.25, 1.59]	2.07	0.13	0.82	.87	2.64
L2 Passion	7	0.63	.07	[0.49, 0.78]	1.70	0.08	1.15	.87	2.62
L2 M.G.O.	8	0.46	.05	[0.35, 0.56]	1.27	-0.10	0.78	.84	2.29
L2 M.G.O. (item removed)	7	0.61	.06	[0.48, 0.74]	1.50	-0.04	1.07	.85	2.34

Note. M.G.O = Mastery Goal Orientation; P.R. = Person Reliability; P.S. = Person Separation; Standard Error of Skewness = .11; Standard Error of Kurtosis = .21.

The Rasch analysis item statistics for the seven-item scale are presented in Table 22. Fit statistics ranged from .78 to 1.34, well inside the targeted range of .5 to 1.5. Items showed adequate fit to the model and the item-measure correlations show that the items are contributing variance toward the scale measure. Item standard errors of .05 and .06 show fairly precise item measures. The Wright map in Figure 14 shows that items are easy to endorse as a group. The items and persons are suitably targeted with more information being located below the mean for this group. There are a few outliers at the extremes of the group but this is expected with large samples.

Table 22. *Rasch Descriptive Statistics for the Interested-in-L2 Scale Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	1.00	.06	.88	.90	.82
5	.79	.06	.99	.98	.83
7	.16	.06	.97	1.01	.77
2	-.19	.06	.79	.78	.83
3	-.43	.06	.91	.90	.81
4	-.53	.06	1.18	1.11	.80
6	-.79	.06	1.34	1.34	.71

The results of the unidimensionality analysis showed in Table 23 shows that the variance explained by the measure was 64.9%. This meets the criterion of 50%. The PCA of the Rasch residuals yielded a first contrast of 1.7 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 8.4%, below the 10.0% criterion.

```
-----------------------------------------------------------------
Students with higher |  More difficult to endorse items
interest in L2 self  |
    6        .#####  +
                     |
                     |
                     |
                ###  |
    5                +
                    T|
                     |
                .##  |
                     |
    4                +
              .####  |
                     |
               .###  |
            #######  |
    3               S+
              .####  |
          .########  |
                     |
            .######  |
    2   ###########  +
                     |
            .######  |
             .#####  |
         ########## M|T
    1       .######  +  I1: Look forward to English classes
                     |  I5: Very excited about English
           ########  |S
              #####  |
             .#####  |  I7: Learn in English class is meaningful
    0         #####  +M
                  .  |  I2: English lessons are enjoyable
               #### S|  I3: English-very interesting field of study
                .##  |S I4: English-very fascinating
                ###  |  I6: Learn in English lessons is important
   -1          ####  +
                 .#  |T
                ###  |
                 .#  |
                  .  |
   -2            .#  +
                  . T|
                  .  |
                  .  |
                  .  |
   -3                +
                  .  |
                     |
                  .  |
                  .  |
   -4             .  +
                     |
                     |
                     |
                  .  |
   -5             .  +
Students with lower  |  Easier to endorse items
interest in L2 self  |
-----------------------------------------------------------------
```

Figure 14. Wright map of Interested in L2 Self. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, and T = 2 SD.

Table 23. *Unidimensionality Analysis for Positive Self Variables*

Variable	Variance Explained by Measures %	Eigenvalue of the 1st contrast	Unexplained Variance in 1st Contrast %
Interested in L2 self	61.9	1.9	9.0
Interested in L2 self (item removed)	64.9	1.7	8.4
Passion for L2 learning	64.2	1.6	8.1
L2 mastery goal orientation	57.9	1.5	8.0
L2 mastery goal orientation (item removed)	61.6	1.5	8.2

Values from the Rasch analysis of scale functioning and structure are given in Table 24. As can be seen, the category with the fewest observations had a count of 121, well above the guideline minimum count of 10. The observed counts distributed through the categories exhibits no abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories.

Table 24. *Rasch Rating Scale Functioning for Interested-in-L2*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	121	-2.31	1.76	(None)
2 Disagree	204	-1.29	1.28	-2.65
3 Slightly disagree	533	-.34	.99	-1.75
4 Slightly agree	1168	.86	.85	-.46
5 Agree	974	2.13	.84	1.68
6 Strongly agree	762	3.43	1.01	3.18

The category with the poorest outfit value is an extreme category with a mean-square value of 1.76, less than the maximum criterion of 2.0. The threshold values all advance monotonically, however, the first category advances slightly less than the value of 1.

The statistics for this scale were a mean of 1.42 logits with a standard deviation of 2.07 logits. As showed in Table 21 skewness and kurtosis values acceptable. The Rasch person reliability was .87, the Rasch person separation index was 2.64, Rasch item reliability was .99, and the Rasch item separation index was 9.88.

Passion for L2 Learning Scale

The Rasch item statistics for the passion for L2 learning scale are presented in Table 25. Fit statistics range from .86 to 1.09, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 15 shows that items are easy to endorse as a group. The person distribution spans ten logits. There are a few outliers but this is expected for large samples.

Table 25. *Rasch Descriptive Statistics for Passion for L2 Learning Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
7	.86	.05	.97	1.07	.76
4	.81	.05	1.09	1.08	.79
5	.09	.05	1.10	1.13	.76
6	.01	.05	1.00	.97	.81
2	-.22	.05	.86	.90	.78
3	-.30	.06	.92	.94	.79
1	-1.25	.06	.97	.91	.75

The results of the unidimensionality analysis showed in Table 23 showed that the variance explained by the measures was 64.2%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.6 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 8.1%, below the 10.0% criterion.

Values from the Rasch analysis of scale functioning and structure are given in Table 26. As can be seen, category 1 had the fewest observations with a count of 197, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.51, less than the maximum criterion of 2.0. The threshold values advance monotonically with the thresholds near or above the suggested value of 1.0. This shows that although the items seem to exhibit a close measure grouping, the categories are sufficiently spread apart to provide adequate measurement.

```
------------------------------------------------------------------
 Students higher in  |  More difficult to endorse items
 L2 learning passion |
    5            .#  +
                     |
                 .#  |
                     |
                     |
                     |
    4           .##  +
                  .  |
                    T|
                .##  |
                     |
                .##  |
    3                +
               .###  |
                     |
              .####  |
                .##  |
                    S|
    2          .###  +
            .######  |
                     |
             .#####  |
         .#########  |T
            .######  |
    1                +  P7: Completely absorbed learning English
           .#######  |  P4: Like myself learning English
            .###### M|S
           .#######  |
                     |
        .##########  |  P5: English relates to other activities
    0     .########  +M P6: Passionate about learning English
            .######  |  P2: Learning allows English appreciation
             .#####  |  P3: English allows memorable experiences
               .###  |
                     |S
               .###  |
   -1         ##### S+
                .##  |
                .##  |T P1: English allows variety of experiences
                 ##  |
                 ##  |
                     |
   -2            .#  +
                 .#  |
                     |
                  # T|
                 .#  |
                     |
   -3             .  +
                     |
                  .  |
                     |
                     |
                  .  |
   -4                +
                     |
                     |
                     |
                     |
                     |
   -5             #  +
 Students lower in   | Easier to endorse items
 L2 learning passion |
------------------------------------------------------------------
```

Figure 15. Wright map for Passion for L2 Learning. Each "#" represents 4 persons. Each "." represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Table 26. *Rating Scale Functioning for Passion for L2 Learning*

Category	Count	Average measure	Outfit MNSQ	Andrich thesholds
1 Strongly disagree	197	-2.08	1.51	(None)
2 Disagree	380	-1.26	.93	-2.54
3 Slightly disagree	734	-.38	.96	-1.39
4 Slightly agree	1204	.67	.87	-.32
5 Agree	712	1.65	.91	1.64
6 Strongly agree	532	2.87	1.07	2.61

The statistics for this scale were a mean of .63 logits with a standard deviation of 1.70. As seen in Table 21 skewness and kurtosis values were acceptable. The Rasch person reliability was .87, the Rasch person separation index was 2.62, the Rasch item reliability was .99, and the Rasch item separation index was 11.92.

Mastery Goal Orientation Scale

The Rasch analysis item statistics are presented in Table 27. Fit statistics range from .76 to 1.63, outside the targeted range of .5 to 1.5. Item one does not fit the model well and also shows a lower point-measure correlation. Other items show adequate fit to the model and he high item-measure correlations show that the items are contributing variance toward the scale measure. Item standard errors of .05 and .06 show fairly precise item measures. The person reliability is .84 and the person separation is 2.29 showing that the scale might be acceptable but that it also might be improved with the elimination of the misfitting item.

Table 27. *Rasch Descriptive Statistics for the Mastery Goal Orientation Scale Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
5	1.09	.05	1.13	1.16	.67
1	.63	.05	1.53	1.63	.57
6	.51	.05	1.10	1.13	.63
7	-.07	.05	.80	.80	.79
2	-.11	.05	.87	.85	.76
3	-.23	.05	.74	.74	.80
8	-.50	.05	.77	.76	.80
4	-1.31	.06	1.01	.93	.72

The misfitting item stated, "My aim is to completely master the material presented in this class." Upon reflection, this item might be confusing in that master has a sense of completeness so that there is some redundancy in the wording, that is, completely "complete." This item was meant to tap into the idea of mastery goal orientation in English emphasizing the thoroughness of student learning. However, students seem to be interpreting this item in different ways. In addition to misfitting the Rasch model, this item at .63 logits has other items that are near enough so that information for this portion of the scale is already being provided for by another item and can be eliminated. Considering the problems with this item and the lower point-measure correlation it was decided to remove this item from the scale. The items were then reanalyzed with the further use of this scale using seven items.

The Rasch analysis item statistics for the seven-item scale are presented in Table 28. Fit statistics range from .78 to 1.34, well inside the targeted range of .5 to 1.5. Items show adequate fit to the model and the high item-measure correlations show that the items are contributing variance toward the scale measure. Item standard errors of .05 and .06 show fairly precise item measures. The Wright map in Figure 16 showed that items are easy to endorse as a group. There are a few outliers at the extremes of the group but this is expected with large samples.

Table 28. *Rasch Descriptive Statistics for the Mastery Goal Orientation Scale Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
4	1.33	.05	1.19	1.23	.70
5	.68	.05	1.24	1.34	.64
6	.02	.06	.90	.92	.78
1	-.03	.06	.97	.94	.76
2	-.15	.06	.81	.82	.80
7	-.46	.06	.80	.78	.81
3	-1.38	.06	1.05	.99	.73

The results of the unidimensionality analysis showed in Table 23 shows that the variance explained by the measures was 61.6%. This meets the criterion of 50%. The PCA of the Rasch residuals yielded a first contrast of 1.5 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 8.2%, below the 10.0% criterion.

```
------------------------------------------------------------------
Students stronger in |  More difficult to endorse items
mastery goal         |
orientation       #  |
                     |
                     |
                 .#  |
    4                +
                     |
                  .  |
                    T|
                     |
                 .#  |
    3                +
                .##  |
                .##  |
                     |
             .#####  |
             .#####  |
    2               S+
           .#######  |
            .######  |
                  .  |T
              #####  |  M4: Like learning difficult things in class
        ###########  |
    1       #######  +
                     |S
        .########## M|  M5: Like learning from my mistakes in class
         .#########  |
                  .  |
         .#########  |  M6: Focus on developing English skills
    0       .######  +M M1: Understand content of this course
            .######  |  M2: Goal is to learn much in class
                     |
            #######  |  M7: Like to study new topics in class
              #####  |
                ### S|S
   -1         .####  +
                     |
                 .#  |  M3: Want to learn much in class
                .##  |T
                 .#  |
                .##  |
   -2             .  +
                    T|
                 .#  |
                  .  |
                     |
                  .  |
   -3                +
                  .  |
                     |
                  #  |
                     |
   -4                +
                     |
                     |
                     |
                     |
                     |
Students weaker in  .|  Easier to endorse items
mastery goal         |
orientation          |
------------------------------------------------------------------
```

Figure 16. Wright map of Mastery Goal Orientation. Each "#" represents 4 persons. Each "." represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 29. As can be seen, the category with the fewest observations with a count of 181, was well above the guideline minimum value of 10. The observed counts distributed through the categories exhibits no abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.72, less than the maximum criterion of 2.0. The threshold values all advance monotonically, however, the first step does not advance by the criterion of 1. This suggests that this scale might function more effectively with the first two categories collapsed. The other categories are near or greater than 1, suggesting that the other categories are functioning effectively.

Table 29. *Rasch Rating Scale Functioning for Mastery Goal Orientation*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	181	-1.98	1.72	(None)
2 Disagree	306	-1.27	1.17	-2.43
3 Slightly disagree	796	-.40	.84	-1.73
4 Slightly agree	1308	.65	.89	-.30
5 Agree	718	1.71	.83	1.76
6 Strongly agree	449	2.84	1.01	2.70

The statistics for this scale were a mean of .61 logits with a standard deviation of 1.50. As seen in Table 21 skewness and kurtosis values were acceptable. The Rasch person reliability was .85, the Rasch person separation index was 2.35, the Rasch item reliability was .99, and the Rasch item separation was 13.66.

Modeled L2 Motivational Variables

L2 Speaking Self-Efficacy Scale

The Rasch item statistics for the L2 speaking self-efficacy measure are presented in Table 60. Fit statistics range from .71 to 1.30, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 17 shows that items are difficult to endorse as a group. The person distribution spans many logits. There are a few outliers but this is expected for large samples.

Table 30. *Rasch Descriptive Statistics for the L2 Speaking Self-Efficacy Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
9	1.77	.06	1.22	1.30	.58
8	.98	.06	1.29	1.26	.63
7	.92	.06	.83	.90	.73
5	-.13	.05	.72	.71	.81
6	-.17	.05	.96	.97	.77
1	-.23	.05	.83	.81	.79
4	-.27	.05	.82	.80	.78
3	-.97	.05	1.21	1.19	.74
2	-1.89	.05	1.12	1.09	.66

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
L2 speaking          |
self-efficacy     .  |
    5                +
                     |
                     |
                  .  |
                     |
    4                +
                     |
                  .  |
                  .  |
                     |
    3                +
                     |
                     |
                  .  |
                  .  |
    2             . T+T
                  #  |  SSE9: Give a speech
                     |
                  .  |
                  #  |
    1          .###  +S SSE7: Talk on phone; SSE8: Opinions
                .##  |
                 ## S|
              .####  |
              #####  |
    0        .#####  +M
        ###########  |  SSE1: Order meal; SSE4: Talk shopping;
            #######  |  SSE5: Talk school; SSE6: Talk about future
             .#####  |
              .#### M|
   -1        .#####  +S SSE3: Introduce myself at a party
        .##########  |
             ######  |
               .###  |
               .###  |  SSE2: Answer simple questions
   -2           .##  +T
                 ## S|
                ###  |
               .###  |
                  #  |
   -3           .##  +
                 .#  |
                     |
                 .# T|
                     |
   -4            .#  +
                     |
                 .#  |
                     |
                     |
   -5                +
                  .  |
                     |
                     |
                     |
   -6             #  +
Students lower in    |  Easier to endorse items
L2 speaking          |
self-efficacy        |
------------------------------------------------------------------
```
Figure 17. Wright map of L2 Speaking Self-Efficacy. Each “#” represents 5 persons. Each “.” represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

The results of the unidimensionality analysis showed in Table 31 shows that the variance explained by the measures was 65.1%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 5.9%, well below the 10.0% criterion.

Table 31. *Unidimensionality Analysis for L2 Motivational Variables*

Variable	Variance explained by measures %	1st contrast eigenvalue units	Unexplained variance in 1st contrast %
L2 speaking self-efficacy	65.1	1.5	5.9
L2 listening self-efficacy	60.7	1.9	8.1
L2 listening self-efficacy (categories collapsed)	60.7	1.8	8.0
L2 reading self-efficacy	60.2	1.5	8.3
L2 reading self-efficacy (categories collapsed)	60.5	1.5	8.3

Values from the Rasch analysis of scale functioning and structure are given in Table 32. As can be seen, category 1 had the fewest observations with a count of 167, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.79, less than the maximum criterion of 2.0. The threshold values all advance monotonically, though some advance less than the criterion of 1.

Table 32. *Rasch Rating Scale Functioning for L2 Speaking Self-Efficacy*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	825	-3.08	1.02	(none)
2 Disagree	980	-1.81	.80	-2.26
3 Slightly disagree	1292	-.83	.87	-1.55
4 Slightly agree	1205	.31	.92	-.24
5 Agree	375	1.14	1.16	1.81
6 Strongly agree	167	2.20	1.79	2.60

The statistics for this scale were a mean of -.90 logits with a standard deviation of 1.48. As seen in Table 33 skewness and kurtosis values were acceptable. The Rasch person reliability was .87, the Rasch person separation index was 2.64, the Rasch item reliability was 1.00, and the Rasch item separation was 17.73.

Table 33. *Descriptive Statistics of L2 Motivational Variables*

Variable	*k*	*M*	*SE*	95% *CI*	*SD*	*Skew*	*Kurtosis*	PR	PS
L2 SSE	9	-.90	.06	[-1.03, -0.78]	1.48	-.34	1.47	.87	2.64
L2 LSE	9	-.89	.06	[-1.01, -0.78]	1.40	-.38	1.34	.87	2.55
L2 RSE	7	-.84	.07	[-0.98, -0.70]	1.67	-.57	1.04	.86	2.46

Note. SSE = Speaking Self-Efficacy; LSE = Listening Self-Efficacy; RSE = Reading Self-Efficacy; PR = Person Reliability; PS = Person Separation; Standard Error of Skewness = .11; Standard Error of Kurtosis = .21.

L2 Listening Self-Efficacy Scale

The Rasch item statistics for the L2 listening self-efficacy measure are presented in Table 34. Fit statistics range from .80 to 1.13, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 18 shows that items are difficult to endorse as a group. The person distribution spans many logits. There are a few outliers but this is expected for large samples.

Table 34. *Rasch Descriptive Statistics for the L2 Listening Self-Efficacy Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
9	1.18	.06	.94	.97	.68
4	.74	.06	1.09	1.10	.68
5	.51	.05	.81	.80	.77
7	.37	.05	1.07	1.04	.71
8	.26	.05	.93	.92	.75
6	-.32	.05	1.07	1.13	.67
1	-.33	.05	.99	.97	.75
2	-.54	.05	.91	.89	.77
3	-1.87	.06	1.12	1.09	.66

The results of the unidimensionality analysis showed in Table 31 shows that the variance explained by the measures was 60.7%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.9 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 8.1%, well below the 10.0% criterion.

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
L2 listening         |
self-efficacy        |
    5             .  +
                     |
                     |
                     |
                     |
    4                +
                  .  |
                     |
                     |
                  .  |
    3             .  +
                  .  |
                  .  |
                  .  |
                  .  |
    2                +
                  . T|
                  .  |T
                 .#  |
                  #  |  LSE9: Details of movies
    1          .###  +
               ####  |S LSE4: Details of conversations
               .###  |  LSE5: Main ideas Movies
             ###### S|  LSE7: Main ideas of a speech
             .#####  |  LSE8: Main ideas of radio shows
    0         #####  +M
         .#########  |
      #############  |  LSE1: Main ideas conversation; LSE6: Main
           ########  |  ideas songs; LSE2: Airport announcements
            .###### M|S
   -1       #######  +
       ############  |
           .#######  |
             ######  |T
                .##  |  LSE3: Simple questions
   -2         .####  +
           .####### S|
               .###  |
                     |
                .##  |
   -3            .#  +
                ###  |
                     |
                 .# T|
                 .#  |
   -4                +
                     |
                 .#  |
                     |
                     |
   -5                +
                  .  |
                     |
                     |
                     |
   -6            .#  +
Students lower in    |  Easier to endorse items
L2 listening         |
self-efficacy        |
------------------------------------------------------------------
```

Figure 18. Wright map for L2 Listening Self-Efficacy. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 35. As can be seen, the category 1 had the fewest observations with a count of 135, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.84, less than the maximum criterion of 2.0. The threshold values all advance monotonically with category six advancing less than the criterion of 1. This suggests that this rating scale might function better with categories 5 and 6 collapsed.

Table 35. *Rasch Rating Scale Functioning for L2 Listening Self-Efficacy*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	730	-2.90	.91	(None)
2 Disagree	1045	-1.71	.84	-2.65
3 Slightly disagree	1366	-.80	.86	-1.52
4 Slightly agree	1220	.15	.98	-.28
5 Agree	345	1.08	1.10	1.81
6 Strongly agree	135	1.90	1.84	2.63

The statistics for this scale were a mean of -.89 logits with a standard deviation of 1.40. As seen in Table 33 skewness and kurtosis values were acceptable. The Rasch person reliability was .87, the person separation index was 2.55, the Rasch item reliability was 1.00, and the Rasch item separation was 15.09.

L2 Listening Self-Efficacy with Categories Collapsed

The Rasch item statistics for the L2 listening self-efficacy measure are presented in Table 36. Fit statistics range from .80 to 1.24, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 19 shows that items are slightly difficult to endorse as a group. The person distribution spans many logits. There are a few outliers but this is expected for large samples.

Table 36. *Rasch Descriptive Statistics for the L2 Listening Self-Efficacy Items (Categories 5 and 6 Collapsed)*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
9	1.25	.06	.94	.98	.69
4	.81	.06	1.08	1.10	.68
5	.56	.06	.81	.80	.76
7	.42	.06	1.00	.99	.72
8	.30	.06	.91	.90	.75
1	-.33	.06	1.00	.98	.74
6	-.35	.06	1.10	1.13	.67
2	-.56	.06	.91	.88	.76
3	-2.09	.07	1.24	1.13	.63

The results of the unidimensionality analysis showed in Table 31 shows that the variance explained by the measures was 60.7%. This surpasses the 50% criterion suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.8 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 8.0%, well below the 10.0% criterion.

```
-----------------------------------------------------------------
Students higher   .  +  More difficult to endorse items
In L2 listening      |
self-efficacy        |
                     |
                     |
                     |
    4             .  +
                     |
                     |
                  .  |
                     |
                  #  |
    3                +
                  #  |
                     |
                  # T|
                     |
                 .#  |
    2                +
                .##  |T
               .###  |
                .##  |
                  .  |
             .##### S|  LSE9: Details of movies
    1        .#####  +S
              #####  |  LSE4: Details of conversations
         ##########  |
            .######  |  LSE5: Main ideas Movies;      LSE 7: Main ideas of speech
            .######  |  LSE8: Main ideas of radio shows
                  .  |
    0      .#######  +M
            .###### M|
             ######  |  LSE1: Main ideas conversation LSE6: Main ideas songs
            .######  |  LSE2: Airport announcements
            .######  |
            #######  |
   -1        .#####  +S
               .###  |
              .####  |
                .## S|
              #####  |
               .###  |T
   -2                +
                .##  |  LSE3: Simple questions
                 .#  |
                  .  |
                .##  |
                     |
   -3            .# T+
                     |
                 .#  |
                     |
                     |
                 .#  |
   -4                +
                     |
                     |
                     |
                     |
                  .  |
                     |
   -5            .#  +
Students lower in L2 |  Easier to endorse items
listening self-efficacy
-----------------------------------------------------------------
```

Figure 19. Wright map for L2 Listening Self-Efficacy (categories collapsed). Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 37. As can be seen, the category 5 had the fewest observations with a count of 480, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.3, less than the maximum criterion of 2.0. The threshold values all advance monotonically more than the criterion of 1. This suggests that this rating scale functions better with category 6 collapsed.

Table 37. *Rasch Rating Scale Functioning for L2 Listening Self-Efficacy (with Categories 5 and 6 Collapsed)*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	730	-2.35	.92	(None)
2 Disagree	1045	-1.09	.87	-2.07
3 Slightly disagree	1366	-.11	.89	-.88
4 Slightly agree	1220	.94	1.02	.46
5 Agree	480	2.00	1.30	2.49

The statistics for this scale were a mean of -.19 logits with a standard deviation of 1.55. As seen in Table 33 skewness and kurtosis values were acceptable. The Rasch person reliability was .87, the Rasch person separation index was 2.58, the Rasch item reliability was 1.00, and the Rasch item separation index was 15.47.

L2 Reading Self-Efficacy Scale

The Rasch item statistics for the L2 reading self-efficacy measure are presented in Table 38. Fit statistics range from .86 to 1.17, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 20 shows that items are difficult to endorse as a group. The person distribution spans many logits. There are a few outliers but this is expected for large samples.

Table 38. *Rasch Descriptive Statistics for the L2 Reading Self-Efficacy Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
6	.86	.06	1.13	1.15	.71
5	.70	.06	1.01	1.00	.73
3	.01	.06	1.17	1.16	.72
4	-.19	.06	.88	.90	.80
7	-.19	.06	.88	.88	.79
2	-.53	.06	.87	.86	.80
1	-.66	.06	1.01	1.00	.78

The results of the unidimensionality analysis showed in Table 31 shows that the variance explained by the measures was 60.2%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 8.3%, well below the 10.0% criterion.

```
------------------------------------------------------------------
Students higher   .  +  More difficult to endorse items
in L2 reading        |
self-efficacy        |
                     |
                     |
                     |
    4             .  +
                     |
                     |
                  .  |
                     |
                     |
    3             .  +
                     |
                  #  |
                  .  |
                  # T|
                     |
    2                +
                 .#  |
                ###  |
                     |
                 .#  |
               .###  |
    1                +T
            .###### S|  RSE6: Novels
        .##########  |  RSE5: Newspaper articles
                     |S
             .#####  |
          .########  |
    0                +M RSE3: Traffic signs
          .########  |  RSE4: Travel pamphlet   RSE7: Web pages
         ##########  |
                  .  |S RSE2: Food menu
        .##########  |  RSE1: Main ideas of emails
              .#### M|
   -1      .#######  +T
                     |
            .######  |
            .######  |
             .#####  |
                     |
   -2          .###  +
            .######  |
                    S|
              .####  |
               .###  |
                  .  |
   -3            .#  +
                     |
               .###  |
                     |
                 .#  |
                    T|
   -4                +
                .##  |
                     |
                     |
                     |
                .##  |
   -5            ##  +
Students lower in L2 |  Easier to endorse items
reading self-efficacy|
------------------------------------------------------------------
```
Figure 20. Wright map of L2 Reading Self-Efficacy. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 39. As can be seen, the category 6 had the fewest observations with a count of 92, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.46, less than the maximum criterion of 2.0. The threshold values all advance monotonically but category six advances than the criterion of 1. This suggests that this rating scale might function better with categories five and six collapsed.

Table 39. *Rasch Rating Scale Functioning for L2 Reading Self-Efficacy*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	503	-3.08	1.12	(none)
2 Disagree	749	-1.83	.93	-2.93
3 Slightly disagree	1078	-.75	.82	-1.60
4 Slightly agree	1122	.26	.94	-.30
5 Agree	220	1.17	.98	2.28
6 Strongly agree	92	1.82	1.46	2.55

The statistics for this scale were a mean of -.84 logits with a standard deviation of 1.67. As seen in Table 33 skewness and kurtosis values were acceptable. The Rasch person reliability was .86, the Rasch person separation index was 2.46, the Rasch item reliability was .99, and the Rasch item separation was 9.09.

L2 Reading Self-Efficacy with Categories Collapsed

The Rasch item statistics for the L2 reading self-efficacy measure are presented in Table 40. Fit statistics range from .83 to 1.20, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The high item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 21 shows that items are slightly difficult to endorse as a group. The person distribution spans many logits. There are a few outliers but this is expected for large samples.

Table 40. *L2 Reading Self-Efficacy (Categories Collapsed)*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
6	.88	.06	1.13	1.15	.72
5	.73	.06	1.01	1.02	.74
3	.01	.06	1.18	1.20	.72
4	-.17	.06	.87	.87	.79
7	-.20	.06	.92	.91	.78
2	-.56	.06	.87	.83	.80
1	-.69	.06	1.00	.96	.78

The results of the unidimensionality analysis in Table 31 showed that the variance explained by the measures was 60.5%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 8.3%, below the 10.0% criterion.

Values from the Rasch analysis of scale functioning and structure are given in Table 41. As can be seen, the category 1 had the fewest observations with a count of 312, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.17, less than the maximum 2.0 criterion. The threshold values all advance monotonically more than the criterion of 1. This suggests that this rating scale functions better with categories 5 and 6 collapsed.

The statistics for this scale were a mean of -.15 logits with a standard deviation of 1.78. As seen in Table 33 skewness and kurtosis values were acceptable. The Rasch person reliability was .86, the person separation index was 2.47, the Rasch item reliability was .99, and the Rasch item separation was 8.96.

Table 41. *Rasch Rating Scale Functioning for L2 Reading Self-Efficacy (Categories Collapsed)*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	503	-2.48	1.12	(None)
2 Disagree	749	-1.21	.95	-2.32
3 Slightly disagree	1078	-.09	.83	-.97
4 Slightly agree	1122	1.01	.95	.39
5 (Strongly) agree	312	2.21	1.17	2.90

```
------------------------------------------------------------------
Students higher in L2|  More difficult to endorse items
reading self-efficacy|
                   # |
                     |
                     |
                     |
                     |
    4              # +
                     |
                     |
                     |
                   . |
                    T|
    3                +
                 ### |
                     |
                  .# |
                     |
                .### |
    2                +
                     |
              .##### |
                    S|
         .########## |
             .###### |T
    1                +
           .######## |  RSE6: Novels
                     |  RSE5: Newspaper articles
           .######## |S
         .########## |
                   . |
    0    .########## +M RSE3: Traffic signs
               .#### M|  RSE4: Travel pamphlet   RSE7: Web pages
                   . |
            .####### |S RSE2: Food menu
             .###### |  RSE1: Main ideas of emails
             .###### |
   -1                +
              .##### |T
                 .## |
             ####### |
                    S|
               .#### |
   -2           .### +
                   . |
                  .# |
                     |
                .### |
                     |
   -3             .# +
                     |
                    T|
                 .## |
                     |
                     |
   -4                +
                   . |
                 .## |
                     |
                     |
                     |
   -5             ## +
Students lower in L2 |  Easier to endorse items
reading self-efficacy|
------------------------------------------------------------------
```

Figure 21. Wright map for L2 Reading Self-Efficacy (categories collapsed). Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Chapter Summary

In this chapter Rasch analysis of items and scales for the main variables was done. These variables are components of latent constructs that form part of a structural model. For the positive self variables, Curiosity, Flourishing, and Hope were analyzed. The Curiosity variable did not meet a criterion outlined in the methods chapter so it was reanalyzed with the first and second categories collapsed. This improved Curiosity scale functioning. The Flourishing variable was found to function well. The Hope variable was found to function better by separately analyzing the two subscales as separate scales of Hope: Agency and Hope: Pathways. For the positive L2 self, Interested in L2 Self, Passion for L2 Learning, and L2 Mastery Goal Orientation were analyzed. An item in the Interested in L2 Self variable was found to perform poorly. The item was removed and the scale reanalyzed and found to function better. The Passion for L2 Learning variable was found to function well. An item in the Mastery Goal Orientation variable was found to function poorly so it was removed. The reanalyzed scale was found to function better. For the L2 motivational variables, L2 Speaking Self-Efficacy, L2 Listening Self-Efficacy, and L2 Reading Self-Efficacy were analyzed. The L2 Speaking Self-Efficacy variable functioned well. Both the L2 Listening Self-Efficacy and L2 Reading Self-Efficacy scales did not meet the criterion outlined in the methods chapter. In both scales categories five and six were collapsed and reanalyzed. After collapsing categories the scales functioned well. In the next chapter, Rasch analysis of items and scales for the peripheral variables used to establish validity was done.

CHAPTER 5

PRELIMINARY ANALYSIS:

PERIPHERAL INSTRUMENT VALIDITY EVIDENCE

This is the second chapter in a three chapter series of preliminary analyses. In the first stage (Chapter 4) the main instruments were examined to show that they fit the Rasch model. In this chapter Rasch analysis was also done for the peripheral instruments of this study. In the next stage (Chapter 6) analysis was focused on providing convergent and divergent relationship validity evidence among the main and peripheral instruments.

In this chapter the preliminary analyses of the peripheral instruments are described. Rasch analysis of the items were done and then analysis of the scale as a whole. Rasch analysis for this study was done with *Winsteps* software (Linacre, 2011). Item analysis included descriptions of fit statistics (.50 to 1.50), item difficulties, item standard errors, and item-person map as depicted in the Wright maps. Scale analysis included principle components analysis (PCA) of the Rasch residuals for descriptions of unidimensionality given by variance explained by the measures (50%), eigenvalues of unexplained variance in the first contrast (3.0), and percentage of unexplained variance in the first contrast (10%). Scale structure was examined through category observation counts and ascending orders of endorsability, and separation of thresholds. In addition, Rasch person reliability and separation, as well as traditional descriptive statistics, were given.

Self-Related Peripheral Variables for Validation

The variables analyzed in this chapter are the peripheral variables to the main variables of the overall study. The relationships among main variables and peripheral variables are analyzed in Chapter 6. In this chapter, they are examined in three sections from most general to most specific. This section is about the analysis of the self-related variables.

Self-Esteem

The Rasch item statistics for the self-esteem measure are presented in Table 42. Fit statistics range from .79 to 1.31, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 22 shows that items are well matched to the group. The person distribution spans eight logits. There are a few outliers but this is expected for large samples.

Table 42. *Descriptive Statistics for the Self-Esteem Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
2	.26	.05	.89	.89	.71
7	.26	.05	.79	.79	.75
5	.24	.05	1.10	1.10	.72
6	.10	.05	1.31	1.30	.69
4	-.39	.05	1.12	1.12	.66
3	-.03	.05	.99	.98	.71
1	-.45	.05	.79	.79	.75

```
------------------------------------------------------------------
  Students higher     |  More difficult to endorse items
  in self-esteem      |
    4             .  +
                     |
                     |
                  .  |
                     |
                  .  |
    3                +
                  .  |
                     |
                  .  |
                  # T|
                .##  |
    2                +
                 ##  |
               ####  |
                ###  |
              .####  |
               #### S|
    1             .  +
            .######  |
           .#######  |
         ##########  |T
         .#########  |S SE2: Good qualities   SE7: Useful person
                  .  |  SE5: Much proud of    SE6: Satisfied myself
    0    .######### M+M SE3: Life going good
      .############  |
          #########  |S SE4: Do things well as most other people
            .######  |T SE1: At least equal plane worth with others
             .#####  |
                  .  |
   -1       .######  +
              .####  |
               .### S|
              .####  |
                .##  |
                .##  |
   -2           .##  +
                .##  |
                  .  |
                 .# T|
                  #  |
                     |
   -3             .  +
                     |
                     |
                  .  |
                     |
                     |
   -4                +
                  .  |
                     |
                     |
                     |
                     |
   -5             .  +
  Students lower      |  Easier to endorse items
  in self-esteem      |
------------------------------------------------------------------
```

Figure 22. Wright map of Self-Esteem. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 51.7%. This surpasses the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.8 eigenvalue units, below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 12.2%, just above the 10.0% criterion. Although the percentage of unexplained variance in the first contrast is higher than desired, the proportion of explained variance to unexplained variance is satisfactory and the eigenvalue unit of unexplained variance does meet the criterion so that this measure can be considered to be essentially unidimensional.

Table 43. *Unidimensionality Analysis for Positive Self Related Variables*

Variable	Variance explained by measures %	Eigenvalue units in 1st contrast	Unexplained variance in 1st contrast %
Self-esteem	51.7	1.8	12.2
Satisfaction with life	58.2	1.6	13.6
Positive affect	48.6	1.7	11.2
Negative affect	49.3	1.7	10.9
Subjective happiness	61.3	1.6	8.9
Positive feeling	57.7	1.4	10.0
Negative feeling	48.2	1.5	13.1
Negative feeling (categories collapsed)	46.4	1.5	13.6
Positive social relationships	48.4	1.8	13.2
Grit (combined)	48.6	1.9	10.6
Grit: Perseverance	54.4	1.7	15.2
Grit: Passion	55.9	1.5	16.8
Hopelessness in achievement	48.7	1.5	12.8
Hopelessness in relationships	50.5	1.5	12.1

Values from the Rasch analysis of scale functioning and structure are given in Table 44. As can be seen, category 1 had the fewest observations with a count of

201, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.24, less than the maximum criterion of 2.0. The threshold values all advance monotonically but the extreme categories advance less than the criterion of 1.

Table 44. *Rasch Rating Scale Functioning for Self-Esteem*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	256	-1.94	1.08	(none)
2 Disagree	496	-1.17	.91	-2.24
3 Slightly disagree	1078	-.40	.89	-1.52
4 Slightly agree	1276	.35	.93	-.20
5 Agree	445	1.06	.97	1.71
6 Strongly agree	201	1.61	1.24	2.25

The statistics for this scale were a mean of -.07 logits with a standard deviation of 1.28. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .82, the Rasch person separation index was 2.11, the Rasch item reliability was .96, and the Rasch item separation was 5.15.

Table 45. *Descriptive Statistics for Self-Level Peripheral Variables*

	k	*M*	*SE*	95%*CI*	*SD*	Skew	Kurt.	*PR*	*PS*
Self-esteem	7	-0.07	.06	[-0.18, 0.04]	1.28	-.09	1.96	.82	2.11
Satisfaction with life	5	-0.20	.05	[-0.30, -0.10]	1.21	.17	1.68	.73	1.63
Positive affect	8	-0.16	.04	[-0.24, -0.08]	0.98	.56	1.90	.77	1.85
Negative affect	8	-0.77	.05	[-0.87, -0.66]	1.20	-.91	4.93	.77	1.85
Subjective happiness	7	0.85	.07	[0.71, 0.99]	1.66	.39	1.16	.86	2.46
Positive feeling	6	0.64	.06	[0.52, 0.77]	1.50	.12	1.42	.82	2.14
Negative feeling	6	-0.96	.06	[-1.07, -0.84]	1.37	-.99	1.89	.76	1.80
Positive social relationships	7	0.65	.04	[0.57, 0.73]	0.93	.65	2.97	.71	1.57
Grit	9	0.24	.04	[0.16, 0.32]	0.97	.64	3.46	.81	2.04
Grit-Perseverance	5	0.35	.05	[0.24, 0.45]	1.25	.48	1.42	.75	1.72
Grit-passion	4	0.17	.05	[0.07, 0.26]	1.14	-.11	2.28	.63	1.31
Hopelessness-achievement	6	-1.02	.06	[-1.13, -0.91]	1.30	-.85	2.01	.72	1.59
Hopelessness-relationships	6	-1.50	.06	[-1.62, -1.39]	1.34	-.78	0.71	.68	1.45

Note. Kurt. = Kurtosis; Standard Error of Skewness = .11; Standard Error of Kurtosis = .21.

Satisfaction with Life

The Rasch item statistics for the satisfaction with life measure are presented in Table 46. Fit statistics range from .56 to 1.68. One item falls outside the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 23 shows that items are difficult to endorse as a group. The person distribution spans eight logits. There are a few outliers but this is expected for large samples. The cause of the misfitting item was due to a small number of participants (4%). Given that the purpose of this measure in this study is to show the relationship to the main variables in this study a single misfitting item does not compromise this measure or study. It should be noted that it is known that the satisfaction with life scale contains a minor secondary dimension (Slocum-Gori, Zumbo, Michalos, & Diener, 2009). This is explored more in the unidimensionality analysis.

Table 46. *Rasch Descriptive Statistics for the Satisfaction with Life Items*

Item number	Measure (logits)	SE	Infit MNSQ	Outfit MNSQ	Point-measure correlation
5	.93	.05	1.68	1.68	.60
1	.63	.05	.92	.93	.71
2	-.26	.05	.69	.69	.77
3	-.31	.05	.56	.56	.84
4	-.98	.05	1.12	1.13	.66

The results of the unidimensionality analysis Table 43 showed that the variance explained by the measures was 58.2%. This surpasses the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.6 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 13.6%, above the 10.0% criterion. Although the percentage of unexplained variance in the first contrast is higher than desired, the proportion of explained variance to unexplained variance is satisfactory and the eigenvalue unit of unexplained variance does meet the criterion. Therefore, given the purpose of this measure for this study, even if a minor secondary dimension exists, this measure can be considered to be essentially unidimensional as suggested by the small percentage of participants misfitting the data (4%).

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
life satisfaction    |
    4             .  +
                  .  |
                     |
                     |
                     |
                     |
                     |
    3                +
                     |
                     |
                .##  |
                     |
                 .#  |
                    T|
    2                +
                  .  |
                     |
                 ##  |
                  .  |T
                ###  |
             .#####  |
    1               S+
              #####  |  LS5: Live life over, change almost nothing
                     |S
           .#######  |  LS1: Life close to ideal
                     |
          .########  |
         .#########  |
    0                +M
         .######### M|
                     |  LS2: Life conditions excellent
         .#########  |  LS3: Satisfied with my life
       .###########  |
                     |S
           .#######  |
   -1                +  LS4: So far gotten important things in life
            .######  |
                    S|
               ####  |T
               .###  |
                     |
                .##  |
   -2                +
                     |
                .##  |
                  . T|
                 .#  |
                     |
                     |
   -3            .#  +
                     |
                     |
                     |
                     |
                  .  |
                     |
   -4             .  +
Students lower in    |  Easier to endorse items
life satisfaction    |
------------------------------------------------------------------
```
Figure 23. Wright map of Satisfaction with Life. Each "#" represents 5 persons. Each "." represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 47. As can be seen, category six had the fewest observations with a count of 190, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.32, less than the maximum 2.0 criterion. The threshold values all advance monotonically some advance less than the 1.0 criterion.

Table 47. *Rasch Rating Scale Functioning for Satisfaction with Life*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	293	-1.96	1.09	(None)
2 Disagree	418	-1.24	.96	-1.95
3 Slightly disagree	714	-.49	.79	-1.34
4 Slightly agree	758	.32	.95	-.16
5 Agree	319	1.07	.94	1.48
6 Strongly agree	190	1.67	1.32	1.97

The statistics for this scale were a mean of -0.20 logits with a standard deviation of 1.21. As seen in Table 45, skewness and kurtosis values were acceptable. The Rasch person reliability was .73, the Rasch person separation index was 1.63, the Rasch item reliability was .99, and the Rasch item separation index was 12.95.

Positive Affect

The Rasch item statistics for the positive affect measure are presented in Table 48. Fit statistics range from .84 to 1.18, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 24 shows that items are slightly difficult to endorse as a group and were a tight grouping for the items. The person distribution spans eight logits. There are a few outliers more than expected for this large sample.

Table 48. *Rasch Descriptive Statistics for the Positive Affect Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
8	.47	.05	1.13	1.11	.70
7	.43	.05	.94	.92	.64
2	.32	.05	1.05	1.06	.64
3	.28	.05	1.18	1.16	.62
5	.19	.05	1.12	1.10	.63
4	-.14	.05	.84	.85	.68
1	-.65	.05	.88	.88	.67
6	-.89	.05	.90	.89	.65

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 48.6%. This is under the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.7 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 11.2%, above the 10.0% criterion.

```
----------------------------------------------------------------
Students higher in    |  More difficult to endorse items
positive affect       |
    5             .   +
                      |
                      |
                      |
                      |
                      |
    4             .   +
                      |
                      |
                      |
                      |
                      |
    3                 +
                      |
                  .   |
                      |
                  .   |
                  .   |
    2             .   +
                  #   |
                  # T|
                 .#   |
                 .#   |
                .##   |
    1        .#####   +T
               .### S|
               ####   |
        .########    |S Posaff7: Alert    Posaff8: Enthusiastic
            .#####   |  Posaff2: Proud    Posaff3: Strong
            .#####   |  Posaff5: Determined
    0      .######   +M
           .###### M|  Posaff4: Attentive
           .######   |
          .#######   |S
     .###########    |  Posaff1: Active
        .########    |  Posaff6: Excited
   -1      .######   +T
                    S|
               .###   |
               .###   |
               .###   |
                 .#   |
   -2             . T+
                      |
                  .   |
                  .   |
                      |
                  .   |
   -3                 +
Students lower in     |  Easier to endorse items
positive affect       |
----------------------------------------------------------------
```

Figure 24. Wright map of Positive Affect. Each “#” represents 5 persons. Each “.” represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 49. As can be seen, category six had the fewest observations with a count of 246, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.23, less than the maximum 2.0 criterion. The threshold values all advance monotonically but category six advances less than the 1.0 criterion.

Table 49. *Rasch Rating Scale Functioning for Positive Affect*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	275	-1.43	1.05	(None)
2 Disagree	958	-.92	1.00	-2.41
3 Slightly disagree	1364	-.34	.87	-.98
4 Slightly agree	961	.26	.89	.29
5 Agree	490	.87	.98	1.21
6 Strongly agree	246	1.45	1.23	1.89

The statistics for this scale were a mean of -0.16 logits with a standard deviation of 0.98. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .77, the Rasch person separation index was 1.85, the Rasch item reliability was .99, and the Rasch item separation was 11.42.

Negative Affect

The Rasch item statistics for the negative affect measure are presented in Table 50. Fit statistics range from .83 to 1.15, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 25 shows that items are difficult to endorse as a group and a tight grouping for the items. The person distribution spans nine logits. There are a few outliers more than expected for this large sample.

Table 50. *Rasch Descriptive Statistics for the Negative Affect Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
2	.82	.05	1.02	.95	.63
1	.64	.05	.88	.83	.68
3	.44	.05	.91	.91	.65
6	.14	.05	1.11	1.15	.60
7	-.19	.05	.95	.94	.64
8	-.20	.05	1.05	1.04	.65
5	-.79	.05	1.12	1.15	.60
4	-.85	.05	1.03	1.03	.64

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 49.3%. This is just under the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.7 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 10.9%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher    .  +  More difficult to endorse items
in negative affect    |
                      |
                      |
                      |
                      |
    3              .  +
                      |
                      |
                      |
                      |
                      |
    2              .  +
                   .  |
                      |
                   .  |
                   . T|
                   .  |T
    1             ##  +
                   #  |  Negaff2: Afraid
                  .#  |  Negaff1: Scared
                  .#  |S Negaff3: Upset
                 .## S|
                  .#  |  Negaff6: Jittery
    0          .####  +M
               .####  |  Negaff7: Ashamed   Negaff8: Irritable
                .###  |
       .############  |S
              .##### M|
              .#####  |  Negaff4: Nervous   Negaff5: Distressed
   -1        .######  +
               .####  |T
                      |
                 .##  |
                #### S|
                  .#  |
   -2            .##  +
                      |
                 .##  |
                   .  |
                      |
                   . T|
   -3                 +
                      |
                   .  |
                      |
                      |
                   .  |
   -4                 +
                      |
                      |
                   .  |
                      |
                      |
Students lower     #  +  Easier to endorse items
in negative affect    |
------------------------------------------------------------------
```

Figure 25. Wright map of Negative Affect. Each “#” represents 7 persons. Each “.” represents 1 to 6 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 51. As can be seen, category six had the fewest observations with a count of 134, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.44, less than the maximum criterion of 2.0. The threshold values all advance monotonically but the last categories advances less than the 1.0 criterion.

Table 51. *Rasch Rating Scale Functioning for Negative Affect*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	614	-2.18	.96	(None)
2 Disagree	1252	-1.17	.98	-2.41
3 Slightly disagree	1252	-.55	.90	-.87
4 Slightly agree	745	.06	.94	.25
5 Agree	297	.64	1.03	1.23
6 Strongly agree	134	1.07	1.44	1.81

The statistics for this scale were a mean of -0.77 logits with a standard deviation of 1.20. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .81, the Rasch person separation index was 2.04, the Rasch item reliability was .99, and the Rasch item separation was 11.42.

Subjective Happiness

The Rasch item statistics for the subjective happiness scale are presented in Table 52. Fit statistics range from .67 to 1.60. One item falls outside the targeted range of .5 to 1.5. Item 4 had an outfit value of 1.60. Outfit values between 1.5 and 2.0 are not contributing towards the scale but they might not degrade measurement (Linacre, 2011). Future uses of this scale should omit this item. When this item was removed and the scale reanalyzed, the person measure correlated at .99 with the original measure so it was retained for this study. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 26 shows that items are easy to endorse as a group. The person distribution spans nine logits. There are a few outliers but this is expected for large samples.

Table 52. *Rasch Descriptive Statistics for the Subjective Happiness Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
3	1.05	.05	1.17	1.18	.73
2	.33	.05	1.03	1.09	.73
7	.26	.05	.86	.85	.80
6	-.12	.06	.69	.68	.84
4	-.20	.06	1.60	1.60	.65
5	-.46	.06	.70	.67	.82
1	-.86	.06	.93	.93	.76

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 61.3%. This surpasses the criterion

```
-------------------------------------------------------------------
Students higher ###  +  More difficult to endorse items
in subjective        |
happiness      .###  |
                     |
                     |
                     |
    4           ###  +
                  . T|
                     |
             .#####  |
                     |
                     |
    3         .####  +
                .##  |
                     |
                ###  |
                  . S|
             .#####  |
    2    .#########  +
                     |
         .#########  |
        ###########  |
          #########  |
                  .  |T
    1    .#########  +  SH3: Very happy, enjoy life regardless
        .########## M|
       .###########  |
                     |S
      .############  |  SH2: Happy compared to peers  SH7: Really enjoying
      .############  |                                     life these days
    0       #######  +M
            #######  |  SH4: Happy compared to before SH6: Usually feel
                  .  |                                     extremely happy
            #######  |S SH5: Living a very happy life now
              .#### S|
           .#######  |  SH1: In general very happy person
   -1          .###  +
                .##  |T
               .###  |
                ###  |
                  .  |
                 .#  |
   -2             #  +
                  . T|
                 .#  |
                     |
                  .  |
                  .  |
   -3             .  +
                     |
                  .  |
                     |
                  .  |
                     |
   -4             .  +
Students lower in    |  Easier to endorse items
subjective happiness |
-------------------------------------------------------------------
```

Figure 26. Wright map of the Subjective Happiness. Each "#" represents 3 persons. Each "." represents 1 to 2 persons. M = Mean, S = 1 SD, T = 2 SD.

of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.6 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 8.9%, below the 10.0% criterion.

Values from the Rasch analysis of scale functioning and structure are given in Table 53. As can be seen, category 1 had the fewest observations with a count of 140, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.47, less than the maximum criterion of 2.0. The threshold values all advance monotonically but the first threshold advances less than the 1.0 criterion.

Table 53. *Rasch Rating Scale Functioning for Subjective Happiness*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	140	-1.88	1.47	(none)
2 Disagree	287	-1.05	1.12	-2.36
3 Slightly disagree	729	-.32	.92	-1.59
4 Slightly agree	1218	.66	.95	-.28
5 Agree	804	1.66	.83	1.55
6 Strongly agree	578	2.94	1.03	2.67

The statistics for this scale were a mean of .85 logits with a standard deviation of 1.66. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .86 and the person separation index was 2.46, the Rasch item reliability was .99, and the Rasch item separation was 9.76.

Positive Feeling

The Rasch item statistics for the positive feeling measure are presented in Table 54. Fit statistics range from .67 to 1.37, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 27 shows that items are slightly easy to endorse as a group and a tight grouping for the items. The person distribution spans ten logits. There are a few outliers as expected for this large sample.

Table 54. *Rasch Descriptive Statistics for the SPANE: Positive Feeling Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.74	.06	1.33	1.37	.66
6	.48	.06	.96	.95	.74
5	.07	.06	1.23	1.22	.70
2	-.03	.06	.67	.68	.80
4	-.51	.06	.96	.94	.79
3	-.75	.06	.79	.79	.77

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 57.7%. This is over the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.4 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 10.0, just at the 10.0% criterion.

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
positive emotion     |
    5             #  +
                     |
                  .  |
                     |
                     |
                     |
    4                +
                .##  |
                     |
                    T|
                  #  |
                     |
    3          .###  +
                  .  |
               ####  |
                     |
              .####  |
             .#####  |
    2               S+
             .#####  |
            #######  |
                  .  |
         ##########  |
         ##########  |
    1                +T
         ##########  |
                    M|  Posfeel1: Positive
        .##########  |S Posfeel6: Contented
          #########  |
                  .  |
    0      .#######  +M Posfeel2: Good     Posfeel5: Joyful
            #######  |
                  .  |
            #######  |S Posfeel4: Happy
                     |
           ######## S|  Posfeel3: Pleasant
   -1      .#######  +T
                  .  |
             .#####  |
                     |
                 .#  |
                .##  |
   -2                +
                  . T|
                  .  |
                     |
                     |
                  .  |
   -3             .  +
                     |
                     |
                     |
                     |
                     |
   -4                +
                     |
                     |
                     |
                  .  |
                     |
   -5             .  +
Students lower in    |  Easier to endorse items
positive feeling     |
------------------------------------------------------------------
```

Figure 27. Wright map of Positive Feeling. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD,T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 55. As can be seen, category one had the fewest observations with a count of 82, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and with more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is category two with a mean-square value of 1.14, less than the maximum criterion of 2.0. The threshold values all advance monotonically near the 1.0 criterion.

Table 55. *Rasch Rating Scale Functioning for Positive Feeling*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	82	-2.05	1.10	(None)
2 Disagree	293	-1.01	1.14	-2.89
3 Slightly disagree	868	-.22	.98	-.78
4 Slightly agree	1012	.68	.88	.85
5 Agree	624	1.72	.91	1.68
6 Strongly agree	344	2.71	1.04	2.83

The statistics for this scale were a mean of .64 logits with a standard deviation of 1.50. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .82 and the person separation index was 2.14, the Rasch item reliability was .99, and the Rasch item separation was 8.85.

Negative Feeling

The Rasch item statistics for the negative feeling measure are presented in Table 56. Fit statistics range from .81 to 1.29, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 28 shows that items are very difficult to endorse as a group and a tight grouping for the items. The person distribution spans eight logits. There are more outliers than expected for this large sample.

Table 56. *Rasch Descriptive Statistics for the SPANE: Negative Feeling Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
3	.22	.05	.81	.85	.70
2	.16	.05	.92	.95	.68
1	.02	.05	1.00	.98	.67
6	-.07	.05	1.11	1.13	.68
5	-.10	.05	1.20	1.29	.61
4	-.24	.05	.91	.89	.71

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 48.2%. This is just under the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 13.1%, above the 10.0% criterion.

```
-------------------------------------------------------------------
Students higher in    |  More difficult to endorse items
negative feeling      |
    3                 +
                  .   |
                      |
                      |
                  .   |
                      |
                  .   |
    2             .   +
                      |
                  .   |
                    T|
                  .   |
                  .   |
                 .#   |
    1             .   +
                 ##   |
                 ##   |
               ####   |
               ####   |
             .##### S|T Negfeel3: Unpleasant
           ########   |S Negfeel2: Bad
    0                 +M Negfeel1: Negative   Negfeel6: Angry
         .#########   |S Negfeel5: Afraid
        .##########   |T Negfeel4: Sad
                      |
        .##########   |
        .##########   |
                  . M|
   -1  .###########   +
                  .   |
        .##########   |
                  .   |
          .########   |
                      |
            #######   |
   -2               S+
             .#####   |
                  .   |
              .####   |
                  .   |
                  .   |
               .###   |
   -3                 +
                      |
                    T|
                ###   |
                      |
                      |
                      |
   -4                 +
                      |
                ###   |
                      |
                      |
                      |
                      |
   -5           ###   +
Students lower in     |  Easier to endorse items
negative feeling      |
-------------------------------------------------------------------
```

Figure 28. Wright map of Negative Feeling. Each “#” represents 4 persons. Each “.” represents 1 to 2 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 57. As can be seen, category six had the fewest observations with a count of 71, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.31, less than the maximum criterion of 2.0. The threshold values all advance monotonically but the last categories advances less than the 1.0 criterion.

Table 57. *Rasch Rating Scale Functioning for Negative Feeling*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	551	-2.34	1.07	(None)
2 Disagree	1020	-1.27	.87	-2.59
3 Slightly disagree	968	-.52	.87	-.82
4 Slightly agree	492	.00	1.06	.39
5 Agree	150	.50	1.11	1.41
6 Strongly agree	71	1.14	1.31	1.60

The statistics for this scale were a mean of -.96 logits with a standard deviation of 1.37. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .76, the Rasch person separation index was 1.80, the Rasch item reliability was .88, and the Rasch item separation was 2.71.

Negative Feeling with Categories Collapsed

The Rasch item statistics for the negative feeling with categories five and six collapsed are presented in Table 58. Fit statistics range from .82 to 1.30, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 29 shows that items are very difficult to endorse as a group and a tight grouping for the items. The person distribution spans eight logits. There are more outliers than expected for this large sample.

Table 58. *Rasch Descriptive Statistics for the SPANE: Negative Feeling Items (Categories Collapsed)*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
3	.22	.06	.82	.86	.71
2	.16	.06	.95	.97	.69
1	.03	.05	.98	.97	.68
6	-.06	.05	1.09	1.11	.69
5	-.09	.05	1.20	1.30	.62
4	-.26	.05	.93	.91	.72

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 46.4%. This is under the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 13.6%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher in    |  More difficult to endorse items
negative feeling      |
    4             .   +
                      |
                      |
                      |
                  .   |
                      |
                      |
    3                 +
                      |
                  .   |
                      |
                      |
                      |
                      |
    2            .#  T+
                      |
                .##   |
                      |
                 .#   |
                      |
               .###   |
    1         #####   +
                     S|
             .#####   |
          .########   |
         .#########   |
                      |T Negfeel3: Unpleasant
       .###########   |S Negfeel2: Bad
    0    ##########   +M Negfeel1: Negative    Negfeel6: Angry
                  .   |S Negfeel5: Afraid
        .##########   |T Negfeel4: Sad
                  .  M|
       .###########   |
                  .   |
       .###########   |
   -1             .   +
          .########   |
                      |
            #######   |
                     S|
             .#####   |
                  .   |
   -2         #####   +
                  .   |
                  .   |
               .###   |
                      |
                      |
                     T|
   -3           ###   +
                      |
                      |
                      |
                      |
                      |
                ###   |
   -4           ###   +
Students lower in     |  Easier to endorse items
negative feeling      |
------------------------------------------------------------------
```

Figure 29. Wright map of Negative Feeling (categories collapsed). Each "#" represents 4 persons. Each "." represents 1 to 2 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 59. As can be seen, category five had the fewest observations with a count of 221, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.21, less than the maximum criterion of 2.0. The threshold values all advance monotonically but the last threshold advances less than the 1.0 criterion.

Table 59. *Rasch Rating Scale Functioning for Negative Feeling*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	511	-1.92	1.07	(None)
2 Disagree	1020	-.83	.89	-2.16
3 Slightly disagree	968	-.07	.87	-.38
4 Slightly agree	492	.49	1.08	.86
5 (Strongly) agree	221	1.09	1.21	1.68

The statistics for this scale were a mean of -.49 logits with a standard deviation of 1.44. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .77, the Rasch person separation index was 1.82, the Rasch item reliability was .77, and Rasch item separation was 1.82.

Positive Social Relationships

The Rasch item statistics for the positive social relationships measure are presented in Table 60. Fit statistics range from .92 to 1.20, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 30 shows that items are easy to endorse as a group. The person distribution spans six logits. There are a few outliers but this is expected for large samples.

Table 60. *Rasch Descriptive Statistics for the Positive Social Relationships Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
6	.79	.04	1.01	1.03	.57
7	.59	.05	.95	.95	.62
3	.55	.05	.98	.98	.59
5	-.05	.05	1.20	1.19	.57
4	-.47	.05	.92	.92	.67
1	-.49	.05	.95	.92	.67
2	-.92	.05	1.05	.99	.61

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 48.4%. This misses the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.8 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 13.2%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
positive social      |
relationships        |
    4             #  +
                  .  |
                     |
                     |
                     |
                     |
                  .  |
    3                +
                     |
                  .  |
                     |
                .##  |
                    T|
                .##  |
    2                +
               ####  |
             .#####  |
            .######  |
                    S|
            .######  |T
         .#########  |
    1 .############  +
          #########  |  PSR6: Will contribute to better society
       ############  |
                  . M|S PSR3: Seen as compassionate   PSR 7: Can
          .########  |  interact with different groups or cultures
       .###########  |
           .#######  |
    0     .########  +M PSR5: Believe people are kind
              #####  |
                .## S|
               ####  |  PSR1: Have people to talk to  PSR4: Experienced
               .###  |S warm and trusting relationships
               .###  |
                 ##  |  PSR2: Trust my friends
   -1            .#  +
                  . T|
                  #  |T
                 .#  |
                     |
                  .  |
                  .  |
   -2                +
                  .  |
Students lower in    |  Easier to endorse items
positive social      |
relationships        |
------------------------------------------------------------------
```

Figure 30. Wright map of the Positive Social Relationships. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 61. As can be seen, category 1 had the fewest observations with a count of 134, well above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.42, less than the maximum criterion of 2.0. The threshold values all advance monotonically and some advance less than the 1.0 criterion.

Table 61. *Rasch Rating Scale Functioning for Positive Social Relationships*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	134	-.85	1.42	(None)
2 Disagree	291	-.42	1.20	-1.56
3 Slightly disagree	714	-.06	.88	-1.15
4 Slightly agree	1214	.51	.89	-.24
5 Agree	761	1.14	.84	1.31
6 Strongly agree	637	1.81	.97	1.64

The statistics for this scale were a mean of .65 logits with a standard deviation of .93. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .71, the person separation index was 1.57, the Rasch item reliability was .99, and the Rasch item separation was 12.71.

Grit

As mentioned in Chapter 2, the grit construct consists of trait-level components of perseverance and passion for long-term goals (Duckworth et al., 2007; Duckworth & Quinn, 2009). Perseverance for long-term goals means maintaining continuity in terms of effort over years of time. Passion for long-term goals means maintaining continuity in terms of interest over years of time. Grit is maintaining stamina toward long-term challenging goals. In this analysis, grit is examined with these two components combined and then examined as two separate components.

The Rasch item statistics for the grit measure are presented in Table 62. Fit statistics range from .66 to 1.51, just outside the targeted range of .5 to 1.5.

Table 62. *Rasch Descriptive Statistics for the Grit Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
2	.80	.05	1.10	1.13	.49
3	.65	.05	.97	.96	.64
5	.44	.05	1.04	1.03	.69
8	.12	.05	.94	.93	.71
7	-.13	.05	.79	.78	.68
6	-.19	.05	1.11	1.12	.62
1	-.39	.05	.66	.66	.74
4	-.60	.05	1.51	1.47	.59
9	-.69	.05	.86	.87	.64

Item standard errors show fairly precise item measures. The item-measure correlations show that items are not contributing variance toward the scale as in the other measures. The Wright map in Figure 31 shows that items are easy to endorse as a group. The person distribution spans 6.5 logits. There are a few outliers but this is expected for large samples.

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 48.6%. This misses the criterion of 50% suggesting some dimensionality in the data. This suggests that the perseverance and passion components are closely but differently related to each other as theorized in the literature. The PCA of the Rasch residuals yielded a first contrast of 1.9, below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 10.6%, above the 10.0% criterion. The data suggest that combining the two components into a single scale might provide a grit measure that is unidimensional enough for some purposes but that for other purposes separating the two components might be warranted.

Values from the Rasch analysis of scale functioning and structure are given in Table 63. As can be seen, category 1 had the fewest observations with a count of 275, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.26, much less than the maximum criterion of 2.0. The threshold values all advance monotonically, however,

```
-------------------------------------------------------------------
Students higher in    |  More difficult to endorse items
Grit                  |
    4              .  +
                      |
                      |
                      |
                      |
                      |
                      |
    3              .  +
                      |
                   .  |
                      |
                   .  |
                      |
                   .  |
    2             .# T+
                   .  |
                  .#  |
                  .#  |
                 .##  |
                 ###  |
                .### S|
    1          .####  +T
                .###  |  Grit2: Not distracted when working project
          .#########  |  Grit3: Focus on projects that take months
               .####  |S
            .#######  |  Grit5: Achieved a goal that took years
              ######  |
            .####### M|  Grit8: I am a hard worker
    0   ############  +M
                .###  |  Grit6: Overcome setbacks to complete challenge
              .#####  |  Grit7: Finish what I begin
              .#####  |  Grit1: Pursue goal until finished
               .####  |S Grit4: Maintain interest in topic for years
               ##### S|  Grit9: Generally I am a hard worker
                 .##  |
   -1            .##  +T
                   #  |
                  .#  |
                   .  |
                  .# T|
                   .  |
                   .  |
   -2              .  +
                   .  |
                      |
                   .  |
                      |
                   .  |
                      |
   -3                 +
Students lower in     |  Easier to endorse items
Grit                  |
-------------------------------------------------------------------
```

Figure 31. Wright map of Grit. Each “#” represents 5 persons. Each “.” represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

the extreme thresholds do not meet the suggested value of 1.0. For some purposes, the threshold values might be acceptable, but it should be remembered that for other purposes, it might be desirable to collapse the extreme categories.

Table 63. *Rasch Rating Scale Functioning for Grit*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	275	-1.18	1.26	(None)
2 Disagree	575	-.72	.95	-1.74
3 Slightly disagree	1248	-.22	.88	-1.19
4 Slightly agree	1453	.41	.86	-.04
5 Agree	785	.95	.97	1.26
6 Strongly agree	497	1.47	1.09	1.71

The statistics for this scale were a mean of .24 logits with a standard deviation of .97. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .81, the Rasch person separation index was 2.04, the Rasch item reliability was .99, and the Rasch item separation was 10.53.

Grit: Perseverance Component

The Rasch item statistics for the perseverance scale are presented in Table 64. Fit statistics range from .82 to 1.19, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 32 shows that items are slightly easy to endorse as a group. The items are grouped together but persons distribution spans seven logits. There are a few outliers but this is expected for large samples.

Table 64. *Rasch Descriptive Statistics for the Grit: Perseverance Component Items*

Item number	Measure (logits)	SE	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.62	.05	1.19	1.19	.70
4	.25	.05	.90	.89	.77
3	-.05	.05	.82	.83	.73
2	-.11	.05	1.18	1.17	.67
5	-.71	.05	.86	.88	.71

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 54.4%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of the Rasch residuals yielded a first contrast of 1.7, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 15.2%, above the 10.0% criterion. Perseverance as a subcomponent of grit is measured on a short scale of five items. The percentage of unexplained variance suggests that there is an additional dimension in the data. However, with an eigenvalue of 1.7, the additional dimension does not rise to the value of two items and given that it is well below the criterion of 3.0, the additional dimension is not a problem.

Values from the Rasch analysis of scale functioning and structure are given in Table 65. As can be seen, the category 1 had the fewest observations with a count of 136, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
perseverance grit    |
    4             #  +
                  .  |
                     |
                     |
                     |
                     |
                  #  |
    3                +
                     |
                    T|
                 ##  |
                     |
               .###  |
    2         .####  +
                     |
                     |
             .#####  |
              .#### S|
                     |
          .########  |
    1                +
            .######  |T
                     |
        .##########  |  PersGrit1: Achieved a goal that took years
                     |S
          .######## M|  PersGrit4: I am a hard worker
                  .  |
    0   .##########  +M PersGrit3: Finish what I begin
         .#########  |  PersGrit2: Overcome setbacks to complete
                  .  |
            .######  |S
                     |
             .#####  |  PersGrit5: Generally I am a hard worker
            .###### S|T
   -1                +
               ####  |
                     |
                 .#  |
                 .#  |
                     |
                 .#  |
   -2               T+
                 .#  |
                     |
                     |
                 .#  |
                     |
                     |
   -3             .  +
Students lower in    |  Easier to endorse items
perseverance grit    |
------------------------------------------------------------------
```

Figure 32. Wright map of Grit: Perseverance component. Each "#" represents 5 persons. Each "." represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.14, less than the maximum criterion of 2.0. The threshold values all advance monotonically with the thresholds below the suggested value of 1.0. This shows that although the items seem to exhibit a close measure grouping, the categories are sufficiently spread apart to provide adequate measurement.

Table 65. *Rasch Rating Scale Functioning for Grit: Perseverance*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	136	-1.67	1.11	(None)
2 Disagree	287	-.90	1.01	-2.02
3 Slightly disagree	644	-.29	.84	-1.37
4 Slightly agree	858	.47	.91	-.19
5 Agree	482	1.21	1.00	1.37
6 Strongly agree	279	1.92	1.14	2.21

The statistics for this scale were a mean of .35 logits with a standard deviation of 1.25. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .75, the person separation index was 1.72, the Rasch item reliability was .99, and the Rasch item separation was 8.47.

Grit: Passion Component

The Rasch item statistics for the passion scale are presented in Table 66. Fit statistics range from .84 to 1.24, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 33 shows that items are difficult to endorse as a group. The items are grouped together but persons distribution spans eight logits. There are a few outliers but this is expected for large samples.

Table 66. *Rasch Descriptive Statistics for the Grit: Passion Component Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
2	0.77	0.05	.98	1.04	.61
3	0.60	0.05	.89	.88	.72
1	-0.57	0.05	.84	.86	.69
4	-0.80	0.05	1.24	1.22	.71

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 55.9%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of the Rasch residuals yielded a first contrast of 1.5, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 16.8%, above the 10.0% criterion. The percentage of unexplained variance suggests that there is an additional dimension in the data. However, with an eigenvalue of 1.5, the additional dimension does not rise to the value of two items and given that it is well below the 3.0 criterion, the additional dimension is not a problem.

Values from the Rasch analysis of scale functioning and structure are given in Table 67. As can be seen, category 1 had the fewest observations with a count of

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
passion grit         |
    4             .  +
                     |
                     |
                  .  |
                     |
                     |
                     |
    3                +
                     |
                  .  |
                     |
                     |
                  # T|
                     |
    2                +
                .##  |
                     |
                .##  |
                     |T
             .##### S|
                     |
    1       .######  +
                  .  |
           ########  |S PasnGrit2: Not distracted working project
                     |  PasnGrit3: Focus projects that take months
       .###########  |
                     |
                  . M|
    0      ########  +M
                     |
          #########  |
                     |
            .######  |  PasnGrit1: Pursue goal until finished
                     |S
              .#### S|  PasnGrit4: Maintain interest in topic years
   -1                +
              .####  |
                  .  |
                     |T
                 .#  |
                     |
                     |
   -2            .# T+
                  .  |
                     |
                 .#  |
                     |
                     |
                  .  |
   -3                +
                     |
                     |
                     |
                     |
                  .  |
                     |
   -4             .  +
Students lower in    | Easier to endorse items
passion grit         |
------------------------------------------------------------------
```

Figure 33. Wright map of Grit: Passion Component. Each "#" represents 7 persons. Each "." represents 1 to 6 persons. M = Mean, S = 1 SD, T = 2 SD.

139, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.33, less than the maximum criterion of 2.0. The threshold values all advance monotonically with some of the thresholds below the suggested value of 1.0. This shows that although the items seem to exhibit a close measure grouping, the categories are sufficiently spread apart to provide adequate measurement.

Table 67. *Rasch Rating Scale Functioning for the Grit: Passion Items*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	139	-1.61	1.33	(None)
2 Disagree	288	-1.05	.88	-2.13
3 Slightly disagree	604	-.29	.91	-1.32
4 Slightly agree	595	.50	.85	.14
5 Agree	303	1.15	.98	1.47
6 Strongly agree	218	1.82	1.11	1.84

The statistics for this scale were a mean of .17 logits with a standard deviation of 1.14. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .63, the person separation index was 1.31, the Rasch item reliability was .99, and the Rasch item separation was 13.77.

Hopelessness in Achievement

The Rasch item statistics for the hopelessness in achievement measure are presented in Table 68. Fit statistics range from .72 to 1.41, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 34 shows that items are very difficult to endorse as a group. The person distribution spans eight logits. There are a few outliers but this is expected for large samples.

Table 68. *Rasch Descriptive Statistics for the Hopelessness in Achievement Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.80	.06	1.23	1.23	.59
2	.18	.05	.84	.84	.71
3	.01	.05	.74	.72	.74
4	-.06	.05	.85	.85	.70
6	-.43	.05	1.35	1.41	.60
5	-.51	.05	1.01	1.04	.69

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 48.7%. This is less than the criterion of 50% suggesting some multidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 12.8%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
hopelessness in      |
achievement          |
    4             .  +
                     |
                     |
                     |
                     |
                     |
                     |
    3                +
                     |
                     |
                     |
                  .  |
                     |
                     |
    2                +
                     |
                  .  |
                     |
                     |
                  . T|
                 .#  |
    1             #  +
                 .#  |T HLA1: Work won't be successful
                .##  |
                  .  |
               .###  |S
              .####  |
            .###### S|  HLA2: Can't be satisfied with my work
    0     #########  +M HLA3: Won't be satisfied future work;
                  .  |  HLA4: Never achieve work goal
        .##########  |
       ############  |S HLA6: Upset about future
                  .  |  HLA5: Work will not be like what I want
       .###########  |
           ######## M|T
   -1       .######  +
                     |
       .###########  |
                     |
         .#########  |
            .######  |
                     |
   -2        .##### S+
                     |
              #####  |
                  .  |
                     |
                ###  |
                     |
   -3             .  +
               .### T|
                     |
                     |
                     |
                     |
              .####  |
   -4          .###  +
Students lower in    |  Easier to endorse items
hopelessness in      |
achievement          |
------------------------------------------------------------------
```

Figure 34. Wright map of Hopelessness in Achievement. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 69. As can be seen, category 6 had the fewest observations with a count of 69, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.52, less than the maximum criterion of 2.0. The threshold values all advance monotonically some advance less than the 1.0 criterion. This suggests that two categories are somewhat underutilized.

Table 69. *Rasch Rating Scale Functioning for Hopelessness in Achievement*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	663	-2.30	.98	(None)
2 Disagree	797	-1.32	.86	-2.06
3 Slightly disagree	996	-.56	.80	-1.16
4 Slightly agree	533	.00	.92	.28
5 Agree	159	.25	1.48	1.36
6 Strongly agree	69	.57	1.52	1.57

The statistics for this scale were a mean of -1.02 logits with a standard deviation of 1.30. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .72, the Rasch person separation index was 1.59, the Rasch item reliability was .98, and the Rasch item separation was 8.03.

Hopelessness in Achievement with Categories Collapsed

The Rasch item statistics for the hopelessness in achievement measure with category six collapsed are presented in Table 70. Fit statistics range from .73 to 1.36, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 35 shows that items are difficult to endorse as a group. The person distribution spans eight logits. There are a few outliers but this is expected for large samples.

Table 70. *Rasch Descriptive Statistics for the Hopelessness in Achievement with Categories 5 and 6 Collapsed*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
1	.84	.06	1.19	1.21	.60
2	.18	.05	.85	.84	.72
3	.01	.05	.74	.73	.74
4	-.06	.05	.83	.83	.71
6	-.42	.05	1.32	1.36	.61
5	-.55	.05	1.07	1.06	.70

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 49.2%. This is less than the criterion of 50% suggesting some multidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 12.5%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher   .  +  More difficult to endorse items
in hopelessness      |
achievement          |
                     |
                  .  |
                     |
                     |
    3                +
                     |
                     |
                     |
                     |
                  .  |
                     |
    2                +
                 .# T|
                     |
                .##  |
                     |
                ###  |
                 .#  |
    1                +
              #####  |T HLA1: Work won’t be successful
            .###### S|
                  .  |
         .#########  |S
        .##########  |
                  .  |  HLA2: Can’t be satisfied with my work
    0  .###########  +M HLA3: Won’t be satisfied future work
      .############  |  HLA4: Never achieve work goal
                  .  |
           ######## M|S HLA6: Upset about future
            #######  |  HLA5: Work will not be like what I want
                     |
      .############  |T
   -1                +
          #########  |
            .######  |
                  .  |
             .##### S|
                     |
             .#####  |
   -2             .  +
                     |
                ###  |
                     |
                  .  |
               .### T|
                     |
   -3                +
                     |
                     |
              .####  |
                     |
                     |
                     |
   -4          .###  +
Students lower in    |  Easier to endorse items
hopelessness in      |
achievement          |
------------------------------------------------------------------
```

Figure 35. Wright map of the Hopelessness in Achievement (categories collapsed). Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 71. As can be seen, category 1 had the fewest observations with a count of 228, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.50, less than the maximum criterion of 2.0. The threshold values all advance monotonically some advance less than the 1.0 criterion.

Table 71. *Rasch Rating Scale Functioning for Hopelessness in Achievement (Categories 5 and 6 Collapsed)*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	663	-1.92	.97	(None)
2 Disagree	797	-.90	.87	-1.66
3 Slightly disagree	996	-.11	.80	-.73
4 Slightly agree	533	.51	.95	.76
5 (Strongly) agree	228	.80	1.50	1.63

The statistics for this scale were a mean of -0.57 logits with a standard deviation of 1.35. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .75, the Rasch person separation index was 1.74, the Rasch item reliability was .98, and the Rasch item separation was 8.05.

Hopelessness in Relationships

The Rasch item statistics for the hopelessness in relationships measure are presented in Table 72. Fit statistics range from .64 to 1.43, within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 36 shows that items are very difficult to endorse as a group. The person distribution spans many logits. There are no outliers.

Table 72. *Rasch Descriptive Statistics for the Hopelessness in Relationships Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
4	.70	.06	1.43	1.29	.57
5	.57	.06	.69	.64	.70
2	.24	.05	.92	.92	.68
3	-.44	.05	1.17	1.14	.65
1	-.49	.05	.95	.97	.67
6	-.59	.05	.99	1.03	.69

The results of the unidimensionality analysis showed in Table 43 showed that the variance explained by the measures was 49.2%. This is less than the criterion of 50% suggesting some multidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 13.0%, above the 10.0% criterion.

```
------------------------------------------------------------------
Students higher in   +  More difficult to endorse items
hopelessness in      |
relationships        |
                 .   |
                     |
                     |
                 .   |
                 .   |
    1           .#   +T
                 . T|
                .#   |  HLR4: Lost best relationship; not fixed in future
                 .   |  HLR5: Won't have better relationships in future
                 .   |S
                 .   |
              ####   |  HLR2: Not have nice relationships now or in future
              .###   |
    0         .###   +M
                 .   |
           .###### S|
           .######   |
                     |S HLR1: Not satisfied with relationships;
          .#######   |  HLR3: Future relationships unsatisfactory;
       .##########   |  HLR6: Relationships with friends won't become what
                     |        I want
   -1       ########   +T
                     |
       .########## M|
                 .   |
      .###########   |
                     |
      .###########   |
                 .   |
   -2     .#########   +
                     |
            .#####   |
                   S|
                     |
        .#########   |
                 .   |
                     |
   -3       ########   +
                     |
                     |
                   T|
                     |
                     |
              .###   |
                     |
   -4       ########   +
Students lower in    |  Easier to endorse items
hopelessness in      |
relationships        |
------------------------------------------------------------------
```

Figure 36. Wright map of Hopelessness in Relationships. Each "#" represents 4 persons. Each "." represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 73. As can be seen, category 6 had the fewest observations with a count of 48, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 2.13, higher than the maximum criterion of 2.0. The threshold values all advance monotonically some advance less than the 1.0 criterion. This suggests that this rating scale might function better with categories 5 and 6 collapsed.

Table 73. *Rasch Rating Scale Functioning for Hopelessness of Interpersonal Relations*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	1051	-2.38	.95	(None)
2 Disagree	826	-1.48	.82	-1.85
3 Slightly disagree	847	-.72	.70	-1.18
4 Slightly agree	345	-.21	1.13	.39
5 Agree	108	.24	1.31	1.23
6 Strongly agree	48	.33	2.13	1.42

The statistics for this scale were a mean of -1.50 logits with a standard deviation of 1.34. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .68, the person separation index was 1.45, the Rasch item reliability was .99, and the Rasch item separation was 9.32.

Hopelessness in Relationships with Categories Collapsed

The Rasch item statistics for the hopelessness in relationships measure did not meet the criterion with categories five and six collapsed. The scale was collapsed further with categories four, five, and six collapsed and presented in Table 74. Fit statistics range from .65 to 1.29, within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 37 shows that items are difficult to endorse as a group. The person distribution spans seven logits. There were no outliers.

Table 74. *Rasch Descriptive Statistics for the Hopelessness in Relationships Items (Categories Collapsed)*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
4	.85	.06	1.29	1.22	.62
5	.64	.06	.69	.65	.74
2	.29	.06	.88	.87	.72
3	-.50	.06	1.13	1.15	.68
1	-.59	.06	.96	.94	.70
6	-.70	.06	1.14	1.10	.69

The results of the unidimensionality analysis in Table 43 showed that the variance explained by the measures was 50.5%. This met the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 12.1%, above the 10.0% criterion.

```
-----------------------------------------------------------------
Students higher in    |  More difficult to endorse items
hopelessness in       |
relationships         |
    4           .   +
                    |
                    |
                    |
                    |
                    |
                    |
               .#   |
    3               +
                    |
                    |
                    |
                    |
                    |
               .#   |
                    |
                    |
    2              T+
             ###    |
                .   |
                    |
           .####    |
                    |T
                    |
    1     .#####   +
                   S|  HLR4: Lost best relationship; won't be fixed in future
         .######    |
                    |S HLR5: Won't have better relationships in future
          ######    |
                    |
        ########    |  HLR2: Don't have nice relationships now or in future
                    |
    0     ######   +M
                    |
      .#########    |
                . M |
       .########    |  HLR3: Future relationships unsatisfactory
                    |S HLR1: Not satisfied with relationships
                    |  HLR6: Relationships with friends won't become what I want
     .##########    |
   -1           .   +
       .########    |
                    |T
           .####    |
                    |
                   S|
       .########    |
                .   |
   -2               +
         .######    |
                    |
                    |
                    |
                    |
                   T|
                    |
   -3  #########    +
Students lower in   |  Easier to endorse items
hopelessness in     |
relationships       |
-----------------------------------------------------------------
```

Figure 37. Wright map of Hopelessness in Relationships (categories collapsed). Each "#" represents 5 persons. Each "." represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

Values from the Rasch analysis of scale functioning and structure are given in Table 75. As can be seen, category 4 had the fewest observations with a count of 501, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.39, lower than the maximum criterion of 2.0. The threshold values all advance monotonically but the first step advances less than the 1.0 criterion.

Table 75. *Rasch Rating Scale Functioning for Hopelessness of Interpersonal Relations*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	1051	-1.63	.95	(None)
2 Disagree	826	-.58	.87	-1.05
3 Slightly disagree	847	.35	.74	-.23
4 Slightly (to strongly) agree	501	1.04	1.39	1.28

The statistics for this scale were a mean of -.55 logits with a standard deviation of 1.54. As seen in Table 45 skewness and kurtosis values were acceptable. The Rasch person reliability was .73, the Rasch person separation index was 1.66, the Rasch item reliability was .99, the Rasch item separation was 9.69.

L2 Domain Related Peripheral Variables for Validation

The variables analyzed here are peripheral variables to the main variables of the overall study. The relationships among these variables are analyzed in the next chapter. This section is about the analysis of the L2 domain related variables.

Ideal L2 Self

The Rasch item statistics for the ideal L2 self measure are presented in Table 76. Fit statistics range from .69 to 1.26, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 38 shows that items are easy to endorse as a group. The person distribution spans eight logits. There are a few outliers but this is expected for large samples.

Table 76. *Rasch Descriptive Statistics for the Ideal L2 Self Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
5	.50	.05	.93	.92	.81
3	.42	.05	.98	1.01	.80
2	.39	.05	.70	.69	.85
1	.11	.05	1.22	1.20	.76
4	-.08	.05	.89	.95	.79
6	-1.35	.06	1.26	1.20	.65

The results of the unidimensionality analysis showed in Table 83 showed that the variance explained by the measures was 65.4%. This surpasses the criterion

```
-----------------------------------------------------------------
Students higher in   |  More difficult to endorse items
ideal L2 self        |
    4   .##########  +
           .#######  |
                     |
                     |
                    T|
                     |
              .####  |
    3                +
                     |
               ####  |
                     |
                     |
            .######  |
                     |
    2      ######## S+
                     |
          .########  |
                     |
           .#######  |
          #########  |T
                     |
    1    .#########  +
            .######  |
          .########  |
                    M|S IS5: Imagine speaking English with friends
          .########  |  IS2: Future career imagine use of English;
          #########  |  IS3: Often imagine myself speak English
          .########  |  IS1: Future things to do require English
    0             .  +M
           .#######  |  IS4: Dream to use English effectively in future
            .######  |
        .##########  |
            #######  |S
            .######  |
              .#### S|
   -1             .  +
               .###  |
                 .#  |T IS6: When I think about future, important that I use
               .###  |       English
                     |
               .###  |
                     |
   -2            .#  +
                  .  |
                 .# T|
                     |
                 .#  |
                     |
                     |
   -3             .  +
                     |
                     |
                     |
                     |
                  .  |
                     |
   -4             #  +
Students lower in    |  Easier to endorse items
Ideal L2 Self        |
-----------------------------------------------------------------
```

Figure 38. Wright map of Ideal L2 Self. Each "#" represents 3 persons. Each "." represents 1 to 2 persons. M = Mean, S = 1 SD, T = 2 SD.

of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 2.0 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 11.4%, over the 10.0% criterion. Values from the Rasch analysis of scale functioning and structure are given in Table 77. As can be seen, category 1 had the fewest observations with a count of 236, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations but relatively similar counts in the higher categories. The observed counts take a slight dip in category five or get a slight bump in category six. There might be a slight tendency toward an extreme category on this scale. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.14, much less than the maximum criterion of 2.0. The threshold values all advance monotonically with category one advancing less than the 1.0 criterion.

Table 77. *Rasch Rating Scale Functioning for Ideal L2 Self*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	236	-1.70	1.12	(None)
2 Disagree	301	-.92	1.03	-1.60
3 Slightly disagree	570	-.31	.87	-1.24
4 Slightly agree	734	.46	.93	-.18
5 Agree	628	1.41	.92	1.03
6 Strongly agree	756	2.44	1.14	1.99

The statistics for this scale were a mean of .82 logits with a standard deviation of 1.80. As seen in Table 84 skewness and kurtosis values were acceptable. The Rasch person reliability was .84, the Rasch person separation index was 2.26, the Rasch item reliability was .99, and the Rasch item separation was 11.90.

Prosociality Goals

The Rasch item statistics for the prosociality goals measure are presented in Table 78. Fit statistics range from .69 to 1.36, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 39 shows that items are easy to endorse as a group. The person distribution spans eight logits. There are a few outliers but this is expected for large samples.

Table 78. *Rasch Descriptive Statistics for the Prosociality Goals Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
5	1.49	.05	1.27	1.36	.50
4	.76	.05	1.26	1.28	.64
7	.38	.05	.70	.69	.74
3	-.01	.05	1.11	1.12	.65
6	-.12	.05	.76	.75	.73
1	-1.19	.06	.93	.91	.62
2	-1.31	.06	.91	.87	.64

The results of the unidimensionality analysis showed in Table 83 showed that the variance explained by the measures was 56.8%. This surpasses the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
prosociality goals   |
    5             .  +
                     |
                  .  |
                     |
                     |
                     |
    4                +
                  .  |
                     |
                     |
                  #  |
                     |
    3            .#  +
                    T|
                .##  |
                     |
              .####  |
              .####  |
    2                +
           .####### S|T
           .#######  |
                  .  |  PSG5: Feel sense of social belonging in college
           ########  |
         .#########  |
    1   .##########  +S
                  .  |  PSG4: Can express disagreement with classmates
       .########### M|
      .############  |
       ############  |  PSG7: Feel understood by classmates
          .########  |
    0      .#######  +M PSG3: Cooperate with classmates to learn
             .#####  |  PSG6: Feel accepted by classmates
              ##### S|
                ###  |
               .###  |
                     |
   -1            .#  +S
                  .  |  PSG1: Help school friends with problem
                 .# T|  PSG2: Cheer up school friends with problem
                  #  |
                  .  |
                  .  |T
   -2             .  +
                     |
                  .  |
                  .  |
                     |
                  .  |
   -3                +
Students lower in    |  Easier to endorse items
Prosociality goals   |
------------------------------------------------------------------
```

Figure 39. Wright map of the Prosociality Goals. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

contrast of 1.8 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 11.0%, just above the 10.0% criterion. Although the percentage of unexplained variance in the first contrast is higher than desired, the proportion of explained variance to unexplained variance is satisfactory and the eigenvalue unit of unexplained variance does meet the criterion so that this measure can be considered to be essentially unidimensional.

Values from the Rasch analysis of scale functioning and structure are given in Table 79. As can be seen, category 1 had the fewest observations with a count of 139, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.60, less than the maximum criterion of 2.0. The threshold values all advance monotonically but some advance less than the 1.0 criterion. The extreme categories advance the least.

Table 79. *Rasch Rating Scale Functioning for Prosociality Goals*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	139	-1.32	1.60	(None)
2 Disagree	275	-.97	.96	-1.96
3 Slightly disagree	630	-.26	.90	-1.37
4 Slightly agree	1313	.57	.92	-.52
5 Agree	843	1.57	.80	1.49
6 Strongly agree	561	2.39	1.06	2.36

The statistics for this scale were a mean of 0.76 logits with a standard deviation of 1.08. As seen in Table 84 skewness and kurtosis values were acceptable. The Rasch person reliability was .73, the Rasch person separation index was 1.64, the Rasch item reliability was 1.00, and the Rasch item separation was 17.00.

Prosociality Goals with Categories Collapsed

The Rasch item statistics for the prosociality goals measure with categories one and two collapsed are presented in Table 80. Fit statistics range from .68 to 1.39, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 40 shows that items are easy to endorse as a group. The person distribution spans ten logits. There are a few outliers but this is expected for large samples.

Table 80. *Rasch Descriptive Statistics for the Prosociality Goals Items (Category Collapsed)*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
5	1.68	.06	1.35	1.39	.47
4	.78	.05	1.28	1.27	.62
7	.39	.05	.68	.69	.73
3	-.04	.05	1.10	1.11	.65
6	-.14	.06	.76	.75	.74
1	-1.26	.06	.93	.92	.65
2	-1.41	.06	.86	.84	.67

The results of the unidimensionality analysis in Table 83 showed that the variance explained by the measures was 57.6%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.8 eigenvalue units, well below the

```
----------------------------------------------------------------
Students higher in  |  More difficult to endorse items
prosociality goal   |
    5            .  +
                    |
                    |
                    |
                 .  |
                    |
    4               +
                    |
                    |
                 .  |
                    |
                    |
    3            #  +
                    |
                    |
                .#  |
                   T|
               .##  |
    2        .####  +T
                    |
             .####  |  SG5: Feel sense of social belonging in college
                    |
         .######## S|
          ########  |
    1            .  +S
         #########  |  SG4: Can express disagreement with classmates
         .########  |
       .##########  |
                 .  |  SG7: Feel understood by classmates
      .########### M|
    0 ############  +M SG3: Cooperate with classmates to learn
                 .  |  SG6: Feel accepted by classmates
      ############  |
         .########  |
          ########  |
                   S|
   -1       .#####  +S
            .#####  |
               .##  |  SG1: Help school friends with problem;
                    |  SG2: Cheer up school friends with problem
               ###  |
                 .  |
   -2           .# T+T
                .#  |
                    |
                 .  |
                    |
                    |
   -3            .  +
                    |
                    |
                    |
                    |
                    |
   -4               +
                 .  |
                    |
                    |
                    |
                    |
Students lower in . +  Easier to endorse items
Prosociality goals  |
----------------------------------------------------------------
```

Figure 40. Wright map of the Prosociality Goals (categories collapsed). Each "#" represents 4 persons. Each "." represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

3.0 criterion. The percentage of unexplained variance in the first contrast was 10.8%, just above the 10.0% criterion. Although the percentage of unexplained variance in the first contrast is higher than desired, the proportion of explained variance to unexplained variance is satisfactory and the eigenvalue unit of unexplained variance does meet the criterion so that this measure can be considered to be essentially unidimensional.

Values from the Rasch analysis of scale functioning and structure are given in Table 81. As can be seen, category 2 had the fewest observations with a count of 414, well about the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.15, less than the maximum criterion of 2.0. The threshold values all advance monotonically but some advance less than the 1.0 criterion.

Table 81. *Rasch Rating Scale Functioning for Prosociality Goals*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
2 (Strongly) disagree	414	-1.73	1.15	(None)
3 Slightly disagree	630	-.92	0.94	-1.79
4 Slightly agree	1313	.01	0.96	-1.13
5 Agree	843	1.09	0.84	.98
6 Strongly agree	561	2.00	1.06	1.93

The statistics for this scale were a mean of .20 logits with a standard deviation of 1.20. As seen in Table 84 skewness and kurtosis values were acceptable. The Rasch person reliability was .76, the Rasch person separation index was 1.78, the Rasch item reliability was 1.00, and the Rasch item separation was 17.11.

Math Self-Concept

The Rasch item statistics for math self-concept measure are presented in Table 82. Fit statistics range from .56 to 1.38, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 43 shows that items are easy to endorse as a group. The person distribution spans ten logits. There are a few outliers but this is expected for large samples.

Table 82. *Descriptive Statistics for the Math Self-Concept Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
3	.56	.07	1.10	1.34	.83
2	.20	.06	.66	.67	.88
6	.20	.06	.94	.98	.86
1	.16	.06	.60	.56	.89
5	-.02	.06	.90	.96	.87
7	-.05	.06	1.32	1.38	.83
8	-.25	.06	1.21	1.31	.85
4	-.80	.06	1.18	1.15	.87

Table 83. *Unidimensionality Analysis for L2 Self Related Variables*

Variable	Variance explained by measures %	Eigenvalue units in 1st contrast	Unexplained variance in 1st contrast %
Ideal L2 self	65.4	2.0	11.4
Prosociality goals	56.8	1.8	11.0
Math self-concept	73.5	1.6	5.1
Math self-concept (categories collapsed)	71.0	1.6	5.7

Table 84. *Descriptive Statistics of L2 Domain Related Variables*

	k	*M*	*SE*	*95%CI*	*SD*	Skew	Kurt.	*PR*	*PS*
Ideal L2 self	6	0.82	.08	[0.67, 0.98]	1.80	.49	0.61	.84	2.26
Prosociality goals	7	0.76	.05	[0.67, 0.85]	1.08	.34	2.08	.73	1.64
Prosociality goals (col. category)	7	0.20	.05	[0.10, 0.30]	1.20	-.14	3.07	.76	1.78
Math SC	8	-2.02	.11	[-2.24, -1.80]	2.58	.34	-0.20	.85	2.38
Math SC (col. category)	8	-1.38	.12	[-1.61, -1.15]	1.44	.42	-0.28	.85	2.40

Note. Col. category = Collapsed categories; Kurt. = Kurtosis; Standard Error of Skewness = .11; Standard Error of Kurtosis = .21.

The results of the unidimensionality analysis showed in Table 83 showed that the variance explained by the measures was 73.5%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.6 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 5.1%, well under the 10.0% criterion.

Values from the Rasch analysis of scale functioning and structure are given in Table 85. As can be seen, category six had the fewest observations with a count of 219, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations but relatively

```
-----------------------------------------------------------------
Students higher   .  +  More difficult to endorse items
in math self-concept |
                  .  |
                     |
                     |
                     |
    4                +
                  .  |
                     |
                     |
                  .  |
                  .  |
    3             .  +
                     |
                  . T|
                  .  |
                  .  |
                     |
    2             .  +
                  .  |
                  .  |
                  .  |
                  #  |
                  .  |
    1             .  +
                  .  |
                  . S|T
                 .#  |  MSC3: Math is easy for me
                 .#  |S
                 .#  |  MSC1: Very good at math   MSC2: Always been good at math
    0             .  +M MSC 6: Math best subject
                  #  |  MSC5: Learn math quickly  MSC7: Enjoy learning math
                  #  |S MSC8: Good understanding of math
                .##  |
                 .#  |T
                 .#  |  MSC4: Get good grades in math
   -1            .#  +
                 .# M|
                 .#  |
                 .#  |
                 .#  |
                 .#  |
   -2           .##  +
                 .#  |
                 .#  |
                     |
                .##  |
                 .#  |
   -3                +
                .## S|
                     |
                     |
               .###  |
                     |
   -4                +
                     |
               .###  |
                     |
                     |
                     |
   -5  .###########  +
Students lower in    |  Easier to endorse items
math self-concept    |
-----------------------------------------------------------------
```

Figure 41. Wright map for Math Self-Concept. Each "#" represents 9 persons. Each "." represents 1 to 9 persons. M = Mean, S = 1 SD, T = 2 SD.

higher counts in the lower categories. The observed counts decrease from categories one to six. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 2.17, higher than the maximum criterion of 2.0. The threshold values all advance monotonically with category one advancing near the 1.0 criterion.

Table 85. *Rasch Rating Scale Functioning for Math Self-Concept*

Category	Count	Average measure	Outfit MNSQ	Step structure
1 Strongly disagree	1808	-3.18	.98	(None)
2 Disagree	754	-1.92	.79	-2.29
3 Slightly disagree	665	-.91	.88	-1.32
4 Slightly agree	561	.18	1.08	-.27
5 Agree	283	1.38	1.19	1.37
6 Strongly agree	219	2.51	2.17	2.50

The statistics for this scale were a mean of -2.02 logits with a standard deviation of 2.58. As seen in Table 84 skewness and kurtosis values were acceptable. The Rasch person reliability was .85, the Rasch person separation index was 2.38, the Rasch item reliability was .97, and the Rasch item separation was 5.59.

Math Self-Concept with Categories Collapsed

The Rasch item statistics for math self-concept measure with categories five and six collapsed are presented in Table 86. Fit statistics range from .58 to 1.43, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 42 shows that items are easy to endorse as a group. The person distribution spans nine logits. There were no outliers.

Table 86. *Math Self-Concept with Categories Collapsed*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
3	.57	.07	1.18	1.43	.84
6	.25	.07	.92	.97	.86
2	.21	.07	.68	.68	.89
1	.19	.07	.64	.58	.89
5	-.04	.07	.90	1.00	.87
7	-.04	.07	1.28	1.35	.84
8	-.29	.07	1.17	1.23	.85
4	-.85	.07	1.19	1.14	.86

The results of the unidimensionality analysis in Table 83 showed that the variance explained by the measures was 71.0%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.6 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 5.7%, well under the 10.0% criterion.

Values from the Rasch analysis of scale functioning and structure are given in Table 87. As can be seen, category five had the fewest observations with a count of 502, well above the minimum criterion of 10. The observed counts are distributed.

```
------------------------------------------------------------------
Students higher  .#  +  More difficult to endorse items
in math self-concept |
                     |
                     |
                     |
                 .#  |
    4                +
                     |
                     |
                     |
                  #  |
                    T|
    3                +
                  .  |
                     |
                 .#  |
                     |
                  .  |
    2                +
                  .  |
                  .  |
                     |
                 .#  |
                 .# S|
    1            .#  +
                  .  |T
                     |
                  .  |  MSC3: Math is easy for me
                  #  |S MSC6: Math best subject
                 .#  |  MSC1: Very good at math  MSC2: Always been good
    0            .#  +M at math   MSC5: Learn math quickly;
                 .#  |  MSC7: Enjoy learning math
               .###  |S MSC8: Good understanding of math
                 .#  |
                 .# M|
                 .#  |T MSC4: Get good grades in math
   -1            .#  +
                 .#  |
                  .  |
                 ##  |
                 .#  |
                 .#  |
   -2            .#  +
                     |
                .##  |
                     |
                .## S|
                     |
   -3                +
               .###  |
                     |
                     |
                     |
               .###  |
   -4  .###########  +
Students lower in    |  Easier to endorse items
math self-concept    |
------------------------------------------------------------------
```

Figure 42. Wright map of Math Self-Concept (categories collapsed). Each “#” represents 9 persons. Each “.” represents 1 to 9 persons. M = Mean, S = 1 SD, T = 2 SD.

through the categories with no exhibitions of abnormal fluctuations but relatively higher counts in the lower categories. The observed counts decrease from categories one to five. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.81, lower than the maximum criterion of 2.0. The threshold values all advance monotonically with categories one advancing at the 1.0 criterion or more.

Table 87. *Rasch Rating Scale Functioning for Math Self-Concept*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	1808	-2.65	0.96	(None)
2 Disagree	754	-1.35	0.80	-1.74
3 Slightly disagree	665	-.29	0.90	-0.74
4 Slightly agree	561	.95	1.11	0.40
5 (Strongly) agree	502	2.36	1.81	2.09

The statistics for this scale were a mean of -1.38 logits with a standard deviation of 2.73. As seen in Table 84, a skewness value .42 and a kurtosis value of -.28 were acceptable. The Rasch person reliability was .85, the Rasch person separation index was 2.40, the Rasch item reliability was .97, and the Rasch item separation was 5.55.

L2 Motivation Related Peripheral Variables for Validation

The variables analyzed here are peripheral variables to the main variables of the overall study. The relationships among these variables are analyzed in the next chapter. This section is about the analysis of the L2 motivation-related variables.

Intended Learning Effort

The Rasch item statistics for the intended learning effort measure are presented in Table 88. Fit statistics range from .64 to 1.40, within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 43 shows that items are slightly easy to endorse as a group. The person distribution spans nine logits. There are a few outliers but this is expected for large samples.

Table 88. *Rasch Descriptive Statistics for the Intended Learning Effort Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
5	1.04	.05	1.02	1.07	.60
6	.51	.05	.98	.97	.71
8	.21	.05	1.12	1.11	.75
4	.02	.05	1.40	1.39	.64
1	-.03	.05	.64	.65	.78
7	-.03	.05	.74	.75	.73
3	-.40	.05	.91	.90	.78
2	-1.32	.05	1.16	1.06	.70

```
-------------------------------------------------------------------
Students higher in    |  More difficult to endorse items
intended learning     |
effort                |
    4             .  +
                      |
                      |
                      |
                  .   |
                      |
    3                +
                  .   |
                      |
                  #   |
                  . T|
                  #   |
    2             .  +
                      |
               .###   |
             .#####   |
               ### S|T
               .###   |
    1         .####  +  ILE5: Doing best to learn English
             .#####   |
             .#####   |S
              ####    |  ILE6: If possible would watch English TV
       .###########   |
            .###### M|  ILE8: If no school English try elsewhere
    0        ######  +M ILE1: Work hard at learning English
           .#######   |  ILE4: Listen English song on radio
            #######   |  ILE7: Makes great effort to learn English
             .#####   |  ILE3: Would take English course in future
                .##   |S
                .##   |
   -1          .### S+
             .#####   |
                 .#   |T ILE2: Extremely important to learn English
                  #   |
                 .#   |
                 .#   |
   -2                +
                 .# T|
                  .   |
                      |
                  .   |
                  .   |
   -3                +
                  .   |
                  .   |
                      |
                      |
                      |
   -4             .  +
                      |
                      |
                      |
                      |
                      |
   -5             .  +
Students lower in     |  Easier to endorse items
intended learning     |
effort                |
-------------------------------------------------------------------
```

Figure 43. Wright map of Intended Learning Effort. Each "#" represents 5 persons. Each "." represents 1 to 4 persons. M = Mean, S = 1 SD, T = 2 SD.

The results of the unidimensionality analysis showed in Table 89 showed that the variance explained by the measures was 58.4%. This surpasses the criterion of 50% suggesting unidimensionality. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.8 eigenvalue units, well below the 3.0 criterion. The percentage of unexplained variance in the first contrast was 9.3%, below the 10.0% criterion.

Table 89. *Unidimensionality Analysis for L2 Motivation Related Variables*

Variable	Variance explained by measures %	Eigenvalue units in 1st contrast	Unexplained variance in 1st contrast %
Intended learning effort	58.4	1.8	9.3
Persistent effort at L2 learning	61.1	1.5	9.9

Values from the Rasch analysis of scale functioning and structure are given in Table 90. As can be seen, category 1 had the fewest observations with a count of 403, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increase monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.35, less than the maximum criterion of 2.0. The threshold values all advance monotonically but the extreme categories advance less than the 1.0 criterion.

Table 90. *Rasch Rating Scale Functioning for Intended Learning Effort*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	403	-1.68	1.35	(None)
2 Disagree	568	-.98	0.98	-1.75
3 Slightly disagree	1017	-.35	0.82	-1.16
4 Slightly agree	1197	.43	0.82	-.09
5 Agree	625	1.05	0.98	1.35
6 Strongly agree	483	1.81	1.03	1.65

The statistics for this scale were a mean of .08 logits with a standard deviation of 1.20. As seen in Table 91 skewness and kurtosis values were acceptable. The Rasch person reliability was .85, the Rasch person separation index was 2.34, the Rasch item reliability was .99, and the Rasch item separation was 12.99.

Table 91. *Rasch Descriptive Statistics of L2 Motivation Related Variables*

	k	*M*	*SE*	*95%CI*	*SD*	Skew	Kurt.	*PR*	*PS*
Intended learning effort	8	.08	.05	[-.02, .18]	1.20	-.71	2.36	.85	2.34
Persistent effort at L2 learning	6	-.15	.06	[-.27, -.03]	1.43	-.03	1.67	.81	2.04

Note. Standard Error of Skewness = .11; Standard Error of Kurtosis = .21; Kurt. = Kurtosis

Persistent Effort at L2 Learning

The Rasch item statistics for the persistent effort at L2 learning measure are presented in Table 92. Fit statistics range from .70 to 1.25, well within the targeted range of .5 to 1.5. Item standard errors show fairly precise item measures. The item-measure correlations show that items are contributing variance toward the scale. The Wright map in Figure 44 shows that items are difficult to endorse as a group. The person distribution spans eleven logits. There are a few outliers but this is expected for large samples.

Table 92. *Rasch Descriptive Statistics for Persistent Effort at L2 Learning Items*

Item number	Measure (logits)	*SE*	Infit MNSQ	Outfit MNSQ	Point-measure correlation
3	1.79	.06	1.20	1.20	.67
6	.04	.05	1.05	1.05	.77
1	-.26	.05	.96	.95	.75
4	-.32	.05	.70	.70	.78
5	-.62	.05	.82	.81	.77
2	-.63	.05	1.24	1.25	.65

The results of the unidimensionality analysis in Table 89 showed that the variance explained by the measures was 61.1%. This surpasses the criterion of 50%. The PCA of Rasch residuals yielded unexplained variance in the first contrast of 1.5 eigenvalue units, well below the criterion of 3.0. The percentage of unexplained variance in the first contrast was 9.9%, just below the 10.0% criterion.

Values from the Rasch analysis of scale functioning and structure are given in Table 93. As can be seen, category 6 had the fewest observations with a count of 164, above the minimum criterion of 10. The observed counts are distributed through the categories with no exhibitions of abnormal fluctuations and have more observations near the middle than the extremes. This suggests that the observed counts are appropriately distributed. Average measure values increased monotonically across all categories. The category with the poorest outfit value is an extreme category with a mean-square value of 1.38, less than the maximum

```
------------------------------------------------------------------
Students higher in   |  More difficult to endorse items
persistent effort at |
L2 learning          |
    6             .  +
                     |
                     |
                     |
                     |
    5                +
                     |
                     |
                     |
                     |
    4             .  +
                     |
                     |
                  .  |
                     |
    3             .  +
                     |
                  . T|
                 .#  |
                     |
    2            ##  +
               .###  |  PE3: Study by myself hours each week
              #####  |T
                     |
           .####### S|
    1      ########  +
         ##########  |S
                  .  |
         ##########  |
        .##########  |
    0 .############  +M PE6: Try to make English study a habit
                  . M|  PE1: Keep trying to understand if problems
         .#########  |  PE4: Spend more time if trouble learning
            .######  |  PE2: Keep up with schoolwork   PE5: Spend extra time if
            .######  |S                                     activity difficult
   -1             .  +
         .#########  |
           .#######  |
             .##### S|T
              #####  |
   -2                +
                .##  |
               ####  |
                 ##  |
                    T|
   -3             .  +
                     |
                  #  |
                     |
                  .  |
   -4                +
                     |
                     |
                  #  |
                     |
   -5             .  +
Students lower in    |  Easier to endorse items
persistent effort at |
L2 learning          |
------------------------------------------------------------------
```

Figure 44. Wright map of Persistent Effort at L2 Learning. Each “#” represents 4 persons. Each “.” represents 1 to 3 persons. M = Mean, S = 1 SD, T = 2 SD.

criterion of 2.0. The threshold values all advance monotonically but the extreme categories advance slightly than the 1.0 criterion. For the purposes of this study, this rating scale is functioning adequately.

Table 93. *Rasch Rating Scale Functioning for Persistent Effort at L2 Learning*

Category	Count	Average measure	Outfit MNSQ	Andrich thresholds
1 Strongly disagree	280	-2.69	1.27	(None)
2 Disagree	438	-1.63	0.89	-2.60
3 Slightly disagree	847	-.54	0.86	-1.71
4 Slightly agree	1041	.50	0.90	-.18
5 Agree	422	1.42	0.85	1.79
6 Strongly agree	164	1.87	1.38	2.70

The statistics for this scale were a mean of -.15 logits with a standard deviation of 1.43. As seen in Table 91, skewness and kurtosis values were acceptable. The Rasch person reliability was .81, the Rasch person separation index was 2.04, the Rasch item reliability was 1.00, and the Rasch item separation was 14.48.

Chapter Summary

In this chapter Rasch analysis of items and scales for the peripheral variables was done. The variables in Table 94 are used to provide external validity evidence for the modeled main variables. For the positive self variables, Self-Esteem, Satisfaction with Life, Positive Affect, Negative Affect, Subjective Happiness, Positive Feeling, Negative Feeling, Positive Social Relationships, Grit, Hopelessness in Achievement, and Hopelessness in Relationships were analyzed.

The Negative Feeling variable did not meet a criterion outlined in the methods chapter so it was reanalyzed with the fifth and sixth categories collapsed. This improved the Rasch person reliability from .76 to .77. The Subjective Happiness variable was found to have one misfitting item. The person measures correlation of the variable with and without the misfitting item was .99. This was similar enough to be retained for this study but future studies should use the scale with the misfitting item removed. The Hopelessness in Achievement variable did not meet the criterion and was reanalyzed with categories five and six collapsed. The measure again did not meet a criterion and the categories were collapsed further with category four. This improved the Rasch person reliability from .68 to .73. For the positive L2 self-level variables, Ideal L2 Self, Prosociality Goals, and Math Self-Concept were analyzed. The Prosociality Goals scale functioned better with categories one and two collapsed. The Math Self-Concept scale functioned better with categories five and six collapsed with Rasch person separation improved from 2.38 to 2.40. For the L2 motivational variables, the Intended Learning Effort scale and the Persistent Effort at L2 Learning scale were analyzed. They were both found to function effectively. In the next chapter, convergent and divergent validity evidence is provided.

Table 94. *Constructs Used in this Study for Validity Evidence*

Construct level	Positive relationships (Convergent validity)	Negative or no relationships (Divergent validity)
Positive self	Self-esteem, Satisfaction in life, Positive affect, Subjective happiness, Positive feeling, Positive social relationships, Grit	Negative affect, negative feeling, hopelessness in achievement, hopelessness in relationships
Positive L2 self	Prosociality goals, Ideal L2 self	Math self-concept
L2 motivational constructs	Intended learning effort, Persistent effort at L2 learning	

CHAPTER 6

PRELIMINARY ANALYSIS:

CONVERGENT AND DIVERGENT VALIDITY EVIDENCE

Internal validity evidence was presented in Chapter 4 for the modeled variables and in Chapter 5 for the peripheral variables by examining psychometric characteristics of items and scales. The preliminary analysis in this chapter is to provide external validity evidence for the modeled variables. The modeled variables are situated in a nomological network with other related but different variables (Cronbach & Meehl, 1955; Loevinger, 1957). Strong positive relationships shown by high positive correlations are taken as convergent validity evidence while inverse or negative relationships shown by negative correlations are taken as divergent validity evidence. Inverse correlations also show that instrument items are not being responded to without discrimination. In other words, participants are not responding to sets of items in blocks or other fixed patterns but are considering the items individually.

Relationships Among General Self Variables

The participant measures from the Rasch analysis were calculated, screened, and cleaned according to the processes described in Chapters 4 and 5. Pearson product-moment correlations were calculated and displayed in Table 91.

Table 95. *Correlation Matrix of Positive Self and Related Variables*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. CEI	1														
2. Flr	.66	1													
3. HP	.67	.65	1												
4. HA	.64	.74	.65	1											
5. SE	.57	.78	.64	.64	1										
6. SWL	.45	.72	.50	.67	.72	1									
7. PA	.45	.50	.40	.43	.49	.38	1								
8. NA	-.02	-.12	-.12	-.10	-.14	-.18	.23	1							
9. SH	.52	.73	.51	.62	.71	.77	.41	-.17	1						
10. PF	.42	.50	.37	.46	.47	.46	.74	.07	.55	1					
11. NF	-.07	-.14	-.14	-.12	-.12	-.19	.21	.76	-.20	.06	1				
12. PSR	.62	.75	.58	.66	.65	.62	.50	-.09	.67	.55	-.16	1			
13. GPe	.57	.58	.57	.69	.50	.45	.37	.03	.37	.32	-.02	.56	1		
14. GPa	.58	.58	.58	.67	.48	.45	.32	-.07	.37	.29	-.09	.51	.69	1	
15. HAc	-.34	-.43	-.36	-.44	-.37	-.37	-.24	.22	-.38	-.26	.24	-.40	-.30	-.30	1
16. HRe	-.27	-.39	-.29	-.32	-.32	-.32	-.25	.23	-.43	-.31	.27	-.49	-.22	-.16	.56

Note. n = 539. CEI = Curiosity and Exploration Inventory; Flr = Flourishing; HP = Hope: Pathways; HA = Hope: Agency; SE = Self-Esteem; SWL = Satisfaction with Life; PA = Positive Affect; NA = Negative Affect; SH = Subjective Happiness; PF = Positive Feeling; NF = Negative Feeling; PSR = Positive Social Relationships; GPe = Grit: Perseverance; GPa = Grit: Passion; HAc = Hopelessness: Achievement; HRe = Hopelessness: Relationships; p < .01 all correlations except 1, 3, 10, 12, 13, 14 with Negative Affect and 1, 10, 13, 14 with Negative Feeling

The variables were all correlated significantly ($p < .01$) except for some correlations with the Negative Affect and Negative Feeling variables.

As shown in Table 95, the modeled variables (numbers 1-4) curiosity and exploration inventory, flourishing, and the two hope subcomponents have moderate to strong correlations with other general positive psychology variables. For example, the curiosity variable shows correlations over .50 with self-esteem, subjective happiness, positive social relationships, perseverance component of grit, and the passion component of grit. The flourishing variable likewise shows correlations of .50 or over with self-esteem, satisfaction with life, positive affect, positive feeling, subjective happiness, positive social relationships, perseverance component of grit, and the passion component of grit. The hope variables also show correlations over .50 with self-esteem, satisfaction with life, subjective happiness, positive social relationships, perseverance component of grit, and the passion component of grit. The modeled variables also have no or small correlations with variables associated with negative psychological constructs. For example, the curiosity variable has negative correlations with hopelessness in achievements and hopelessness in relationships and non-significant relationships with negative affect and negative feelings.

The purpose of the correlation matrix is to set the modeled variables in a nomological network. This situates the modeled variables by showing relationships with other positive psychology variables and provides external convergent evidence. The modeled variables also show varying degrees of relationship among other modeled variables and with peripheral variables and this shows external discriminant validity evidence. The modeled variables also show no relationships or negative relationships with negative affect and feeling in addition to components of hopelessness, thus showing that there is also external divergent evidence. The correlation matrix also provides some support for the analyses in Chapters 4 and 5. For example, although hope and grit can be used as a complete scale for some purposes, breaking them into the theorized components is supported by the correlations shown. The agency component of hope correlates .65 with the pathways thinking component of hope but the agency component has a stronger relationship with the perseverance component of grit with a correlation of .69. This indicates that the hope components are measuring along different dimensions of hope and supports using the components separately.

Relationships Among L2 Domain Variables

The correlation matrix in Table 96 shows the modeled positive L2 self variables of interest in L2 self, passion for L2 learning, and mastery goal orientation have moderate to strong correlations with other L2 domain variables and no or weak correlation with the math self-concept variable. The modeled variables have strong positive correlations among modeled variables and moderate to strong correlations with other related academic domain variables. The modeled variables of interest and passion and the peripheral variable of ideal L2 self did not have significant correlations with the math self-concept variable. Mastery goal orientation and prosociality goals had weak significant correlations with math self-concept.

Table 96. *Correlation Matrix of Positive L2 Self and Related Variables*

Variables	1	2	3	4	5	6
1. Interest in L2 self	—					
2. Passion for L2	.88	—				
3. Mastery goal orientation	.82	.80	—			
4. Ideal L2 self	.75	.80	.66	—		
5. Prosociality goals	.55	.55	.53	.48	—	
6. Math self-concept	-.02	.04	.12	.04	.16	—

Note. $n = 539$. $p < .01$ all correlations except 1, 2, and 4 with Math Self-Concept

Relationship Among Motivational Variables

Modeled motivational self-efficacy measures correlated strongly with each other and had moderate to strong correlations with other motivational variables as shown in Table 97. The variables of self-efficacy had stronger correlations with each and the effort variables correlated higher with each other.

Table 97. *Correlation Matrix of L2 Motivation Variables*

Variables	1	2	3	4	5
1. Listening self-Efficacy	—				
2. Reading Self-Efficacy	.84	—			
3. Speaking Self-Efficacy	.83	.82	—		
4. Intended Learning Effort	.64	.62	.66	—	
5. Persistent Effort	.66	.66	.70	.80	—

Note. $n = 539$. $p < .01$ all correlations.

Chapter Summary

This chapter provided validity evidence beyond the item and scale analysis done in Chapter 4 by showing relationships at the three different levels used in this study. The main variables to be modeled showed expected positive relationships to related peripheral variables and expected negative relationships to contrasting variables. This provides evidence that students were discriminating among items and scales and shows that scales are measuring different but related constructs.

CHAPTER 7

RESULTS

In this chapter the results of the confirmatory factor analysis (CFA) and the structural equation model (SEM) are presented to answer the research questions given in Chapter 2. The first three research questions asked if composite constructs of positive self-concept, positive L2 self, and L2 motivation could be constructed. Research question 4 asked about the structural relationships among these constructs. Research question 5 asked about the relationships among these constructs with the substitution of L2 proficiency for L2 motivation.

To review briefly, in chapter four raw scale scores were transformed into interval measures given in Rasch logit units. The interval measures are used to perform the CFA to answer research questions 1 and 2. The interval measures and TOEIC scores are used to create a CFA answering research question 3. Structural modeling is conducted to answer research questions 4 and 5.

Confirmatory Factor Analysis

The first confirmatory factor analysis, based on the interval measures derived from the Rasch analysis was performed using AMOS (Arbuckle, 2007). The model is presented in Figure 45 where ovals represent latent variables and rectangles represent measured variables. Lines with arrows on the ends represent relationships among variables and absence of a line represents a lack of relationship. A three factor model of positive self-concept, positive L2 self, and L2 motivation was constructed. Measured variables of curiosity, flourishing, and hope served as indicators of positive self-concept. Measured variables of interested-in-L2 self, passion for L2 learning, and mastery L2 goal orientation served as indicators of positive L2 self. Measured variables of reading self-efficacy, speaking self-efficacy, and listening self-efficacy served as indicators of L2 motivation.

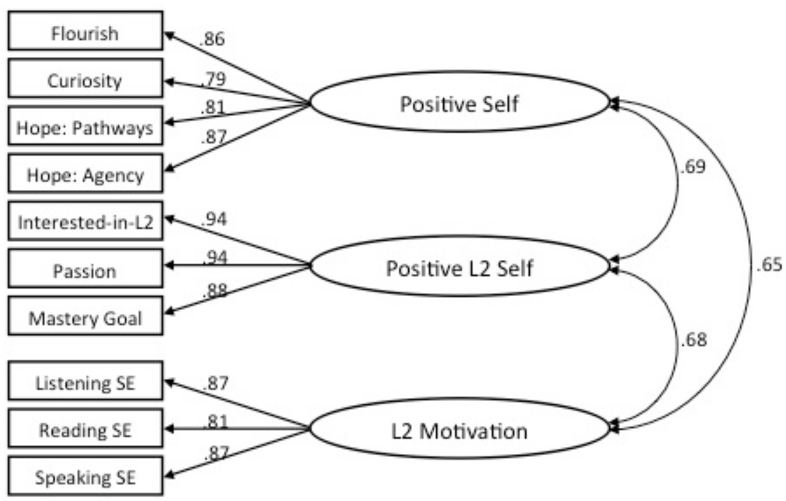

Figure 45. Confirmatory factor analysis with L2 Motivation. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; SE = Self-Efficacy.

The results of the CFA indicated the latent factors and measurement indicators showed adequate fit (X^2 = 121.32, p < .001, CFI = .98, RMSEA = .074 CIs [.060, .088], AIC =167.23). The chi-squared statistic suggested that the data are not fitting the model, but this statistic is sensitive to sample size and is often not met (Byrne, 2010). The CFI indicates good fit but the RMSEA indicates adequate to mediocre fit. RMSEA tends to become smaller with larger *N*-sizes and more variables; with few latent variables and not large *N*-sizes it is not uncommon for RMSEA to be larger (Kenny, Kaniskan, & McCoach, 2014; Kenny & McCoach, 2003). Although one fit index showed adequate to mediocre fit, these results are not poor and are deemed to be acceptable to support a model of three latent variables of positive self at three levels of generality/specificity. First, at a global self-concept level a composite latent variable positive self was supported. It was composed of the three measured variables curiosity, flourishing, and hope. At this global level, the positive self is referenced only by the self, not any particular domain, or specific activity. Second, a positive L2 self at a mid-level or domain level which is a composite latent variable of three measured variables of interested-in-L2 self, passion for L2 learning, and mastery goal orientation was supported. The positive L2 self lacks the generality of a global self but is more specifically referenced to an L2 domain. Third, a positive L2 self at highly specific level that because it is proximal to learning is considered a type of L2 motivation was supported. The positive self at this highly specific level references self-efficacy in three subdomains of L2 reading, L2 listening, and L2 speaking in reference to specific language tasks.

The second confirmatory factor analysis based on the interval measures derived from the Rasch analysis and TOEIC Bridge scores. The measures of positive self-concept and positive L2 self were derived from self-report scales. The TOEIC Bridge scores are objective scores of L2 proficiency. Using objective L2 proficiency measures ensures that a model is not created solely due to a method effect of a single type of measure. As another validating step in demonstrating the beneficial effects of positive self-concept and positive L2 self, a variable of L2 proficiency, unrelated to the development of the self-report measures was included.

The second model, including the factor of L2 proficiency, is presented in Figure 46. This model was identical to the first except that in place of L2 motivation, L2 proficiency was substituted. Sub-scores on TOEIC Bridge were used as measured variables of reading proficiency and listening proficiency as indicators of L2 proficiency. The two factors of positive self and positive L2 self were hypothesized to covary with each other. The two factors of positive L2 self and L2 proficiency were hypothesized to covary with each other.

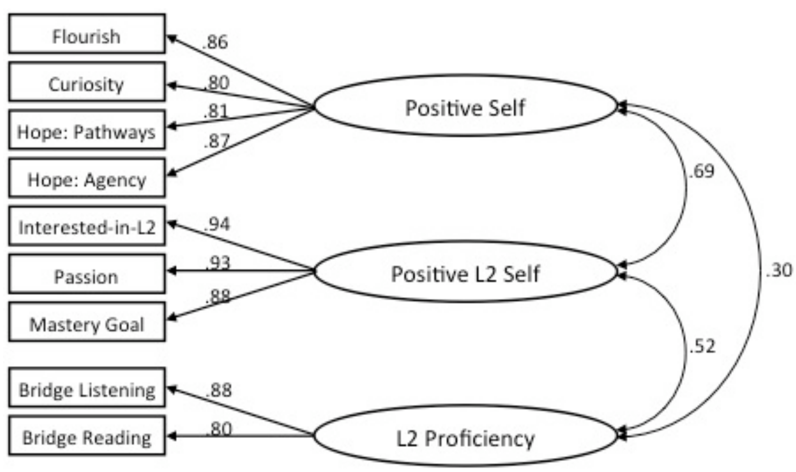

Figure 46. Confirmatory factor analysis with L2 Proficiency (TOEIC Bridge). Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; Bridge = TOEIC Bridge.

The results of the CFA indicated that a positive self-construct latent variable and a positive L2 self latent variable fit a model with L2 proficiency thus providing further support of research questions 1 and 2. The model with these three latent variables showed that it failed the chi-squared test but other goodness-of-fit indexes suggest adequate fit ($X^2 = 76.77$, $p < .001$, CFI = .98, RMSEA = .065, CIs [.049, .082], AIC = 136.77). These CFA results support a model that included latent variables of positive self-concept, positive L2 self with variables of L2 proficiency.

The third confirmatory factor analysis was based on the interval measures derived from the Rasch analysis and TOEIC scores. This was identical to the previous model except TOEIC scores were substituted for the TOEIC Bridge scores. Roughly half the participants had taken the TOEIC Bridge as a proficiency measure and half had taken the TOEIC as a proficiency measure. To provide additional support for research question 1, 2, and 5, it should not matter what particular L2 proficiency measure is used. In other words, if relationships exist among the modeled variables then substituting different but similar measures provides additional supporting measures.

The model with L2 proficiency is presented in Figure 47. A three factor model of positive self-concept, positive L2 self, and L2 proficiency was hypothesized. Measured variables of curiosity, flourishing, and hope served as indicators of positive self-concept. Measured variables of interested-in-L2 self, passion for L2 learning, and mastery L2 goal orientation served as indicators of positive L2 self. Sub-scores on TOEIC were used as measured variables of reading proficiency and listening proficiency, together indicating L2 proficiency. The two factors of positive self and positive L2 self were hypothesized to covary with each other. The two factors of positive L2 self and L2 proficiency were hypothesized to covary with each other.

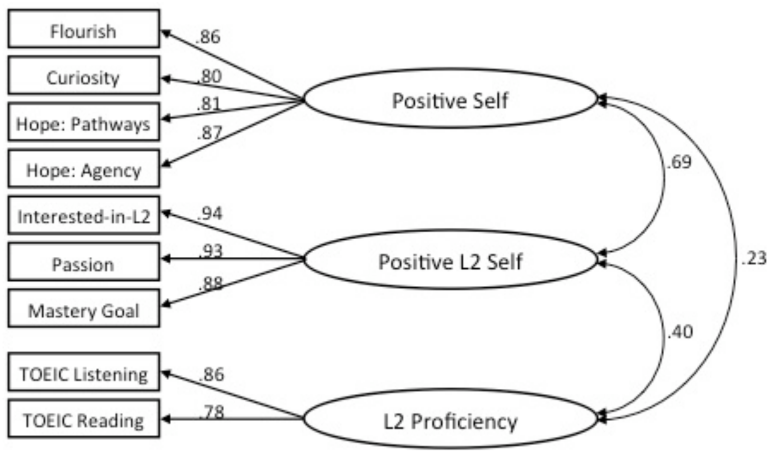

Figure 47. Confirmatory factor analysis with L2 Proficiency (TOEIC). Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation.

The results of the third CFA indicated that a positive self-construct latent variable and a positive L2 self latent variable fit a model with a different measure of L2 proficiency. The model with these three latent variables showed that it failed the chi-squared test but other goodness-of-fit indexes suggest adequate fit (X^2 = 80.66, $p < .001$, CFI = .98, RMSEA = .068, CIs [.052, .084], AIC = 140.66). These CFA results support a model that included latent variables of positive self-concept, positive L2 self with variables of L2 proficiency as measured by the TOEIC. This demonstrated the model worked whether the L2 proficiency measure is TOEIC or TOEIC Bridge.

Equivalent Measurement Models

Kline (2011) suggested that equivalent or near equivalent models should be considered. Two models are tested here. First, a single factor model showed in Figure 48 was tested. Fit indexes show a poor fit to the one factor model (X^2 = 1474.28, p < .001, CFI = .70, RMSEA = .28, CIs [.27, .30], AIC = 1514.28).

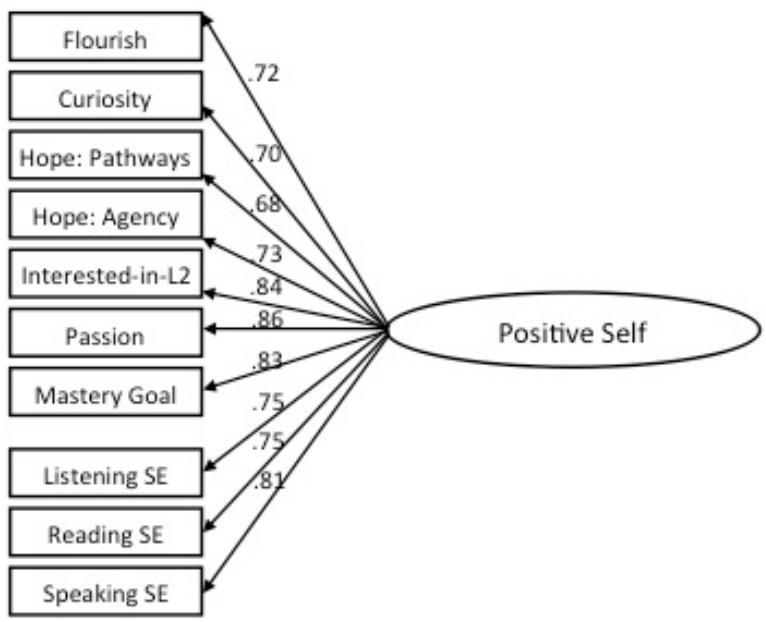

Figure 48. Structural equation model with Single Factor. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; SE = Self-Efficacy.

A second model with two factors showed in Figure 49 is tested next. Fit indexes again show a poor fit to the two factor model (X^2 = 778.94, p < .001, CFI = .84, RMSEA = .21, CIs [.19, .22], AIC = 820.94).

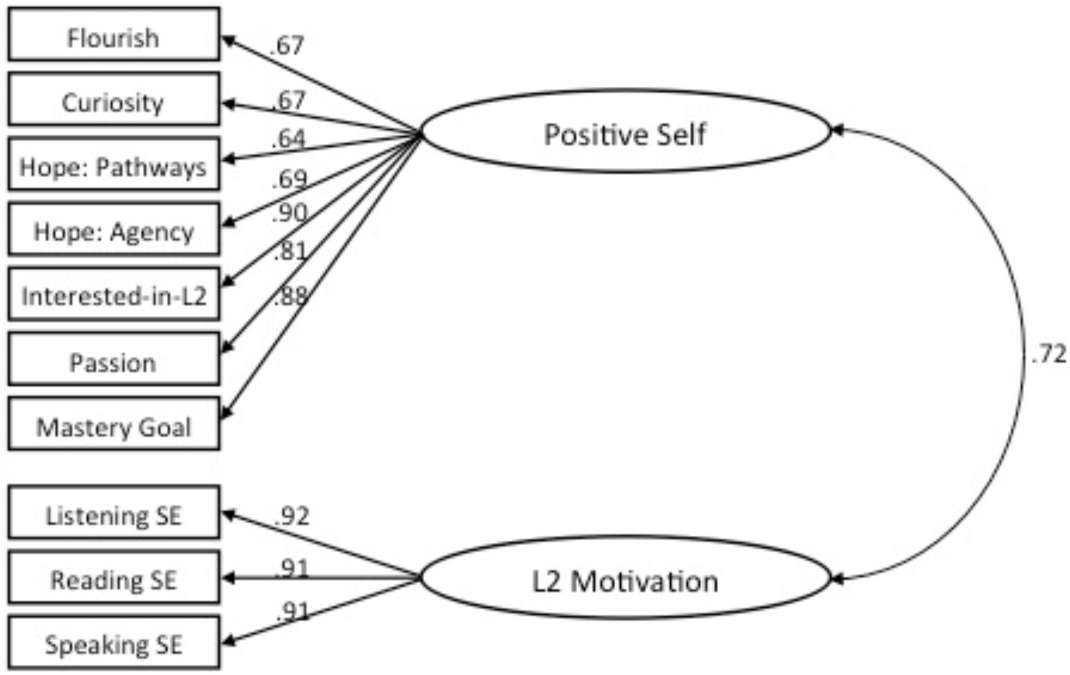

Figure 49. Structural equation model with Two Factors. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; SE = Self-Efficacy.

The three CFAs answered the first three research questions which asked can composite constructs of positive self, positive L2 self, and L2 motivation be constructed. Also, CFAs two and three support the further investigation of research question 5, which looked at the structural relationships to L2 proficiency. As shown in fit indexes summarized in Table 98, latent constructs in the last three CFAs were constructed with reasonable fit to the models proposed while alternative models in CFAs one and two showed poor fit.

Table 98. *Summary of Models and Fit Indexes*

Model	X^2 p < .001	CFI	RMSEA [90% CIs]	AIC
1 Factor	1474.28	.70	.28 [.27, .30]	1514.28
2 Factors	778.94	.84	.21 [.19, .22]	820.94
3 Factors (Motivation)	121.32	.98	.07 [.06, .09]	167.23
3 Factors (L2 Prof. a)	76.77	.98	.07 [.05, .08]	136.77
3 Factors (L2 Prof. b)	80.66	.98	.07 [.05, .08]	140.66

Note. Prof. a = TOEIC Bridge; Prof. b = TOEIC.

Structural Equation Modeling

Structural equation modeling with maximum likelihood estimation (MLE) answered RQs four and five. By assessing the fit of the hypothesized structural model for the relationships among three latent constructs, it confirmed the structural relationships among positive self-concept, positive L2 self, and L2 motivation (RQ4) and the structural relationships among positive self-concept, positive L2 self, and L2 proficiency (RQ5). This model, based on hierarchical constructs of positive self, proposed that a most stable general positive self-concept would lead to a less stable positive L2 self, which in turn would lead to least stable L2 motivation. The relationships in this three-level hierarchy would be stronger than that between the two levels of positive self and L2 motivation. The results of the CFA supported the model with adequate fit (X^2 = 121.32, p < .001, CFI = .98, RMSEA = .074 CIs [.061, .088], AIC =167.32). The structural component showed that L2 motivation was influenced by positive L2 self and positive L2 self in turn was influenced by positive self-concept with stronger paths than between positive self and L2 motivation. Figure 50 illustrates the structural model with loadings of the measured variables, regression paths. All parameters are significant, p < .05.

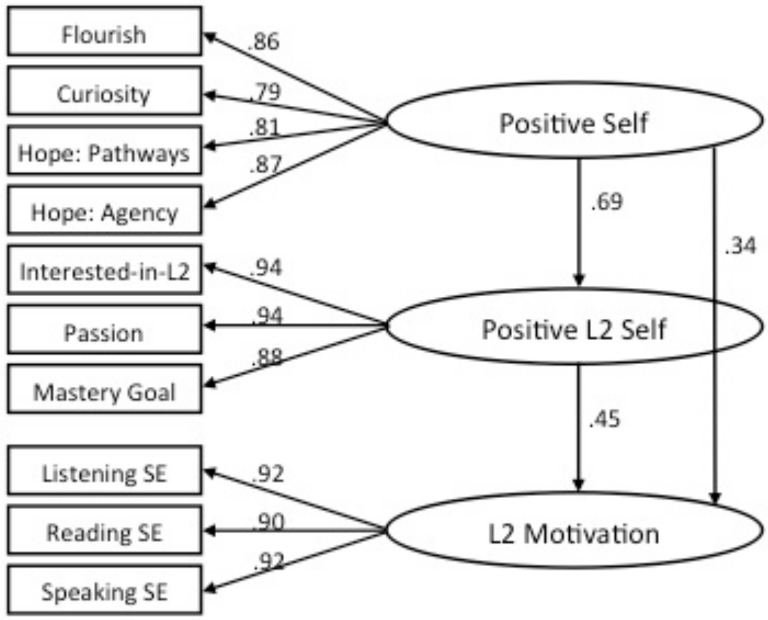

Figure 50. Structural Equation Model with Three Factors. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; SE = Self-Efficacy.

An alternative structural model was considered where only the hierarchy in the model was tested, that is, with no relationship between positive self and L2 motivation. Failure to test alternative models might lead to a confirmation bias (Kline, 2011, p. 14). Figure 51 illustrates the alternative model with positive L2 self regressed to positive L2 self regressed to L2 motivation. The results of the alternative model showed poor model fit compared to the original model (X^2 = 163.98, p < .001, CFI = .97, RMSEA = .088 CIs [.075, .101], AIC = 207.98).

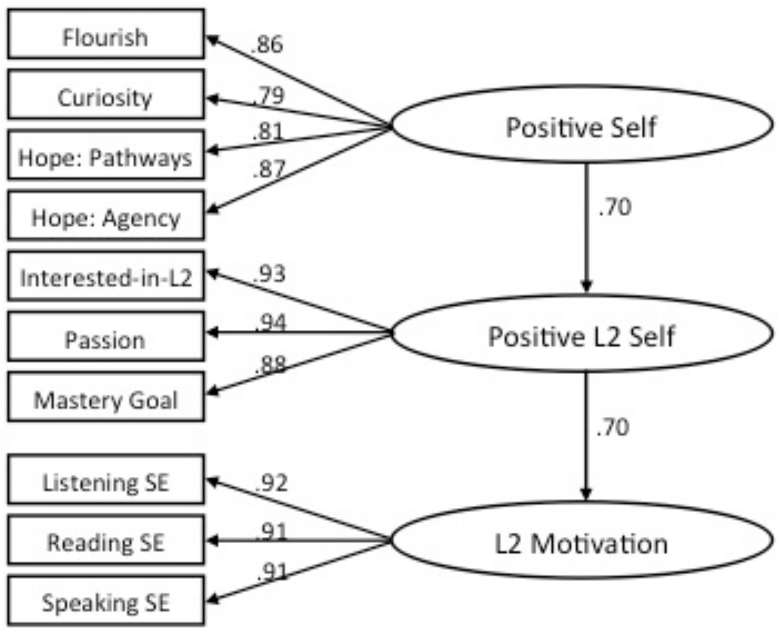

Figure 51. Structural equation alternative model with three factors. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; SE = Self-Efficacy.

To extend the findings of the CFA that positive self-concept and positive L2 self showed a relationship with L2 proficiency as measured by TOEIC Bridge scores a structural model with MLE was performed to assess the fit of the hypothesized structural model for the relationships among the three latent constructs. This model, based on hierarchical constructs of positive self, proposed that a more stable general positive self-concept would lead to a relatively stable positive L2 self which in turn would lead to L2 proficiency. The results of the CFA confirmed that the indicators would load on the latent variables with good fit to the model. The structural component showed that L2 proficiency regressed on positive L2 self and positive L2 self in turn regressed on positive self-concept. Figure 52 shows the structural model with loadings of the measured variables and regression paths (X^2 = 77.87, p < .001, CFI = .98, RMSEA = .064 CIs [.048, .080], AIC =167.23). All parameters are significant, p < .05.

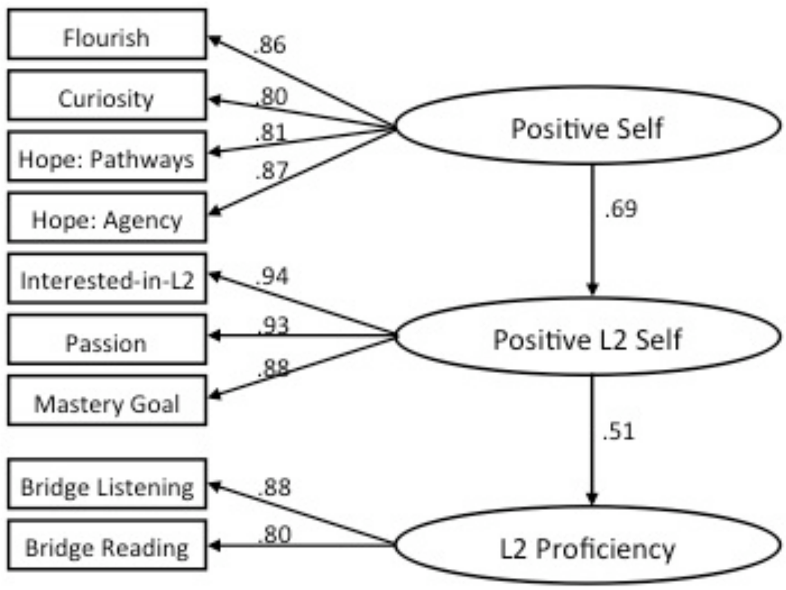

Figure 52. Structural equation model with Proficiency. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; Bridge = TOEIC Bridge.

An alternative structural model was considered where positive self-construct would lead to both positive L2 self and L2 proficiency. Figure 53 illustrates the alternative model with both positive L2 self and L2 motivation both regressed on positive self-concept. The additional parameter was not significant. The results of the alternative model showed similar but slightly poorer model fit with previous model (X^2 = 76.77, p < .001, CFI = .98, RMSEA = .065 CIs [.049, .082], AIC =136.77).

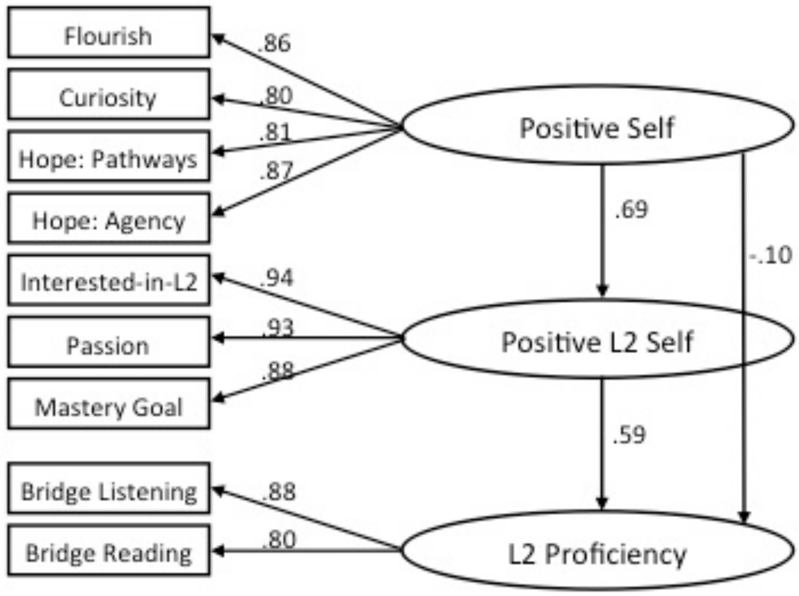

Figure 53. Alternative structural equation model with L2 Proficiency. Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; Bridge = TOEIC Bridge.

To extend the findings of the CFA that positive self-concept and positive L2 self showed a relationship with L2 proficiency also as measured by TOEIC scores, a structural model with MLE was performed to assess the fit of the hypothesized structural model for the relationships among the three latent constructs. This model, based on hierarchical constructs of positive self proposed that a more stable general positive self-concept would lead to a relatively stable positive L2 self which in turn would lead to L2 proficiency. The results of the CFA confirmed that the indicators would load on the latent variables with good fit to the model. The structural component showed that L2 proficiency regressed on positive L2 self and positive L2 self in turn regressed on positive self-concept. Figure 54 illustrates the structural model with loadings of the measured variables and regression paths (X^2 = 81.45, *p*

< .001, CFI = .98, RMSEA = .066 CIs [.051, .083], AIC =139.45). All parameters are significant, $p < .05$.

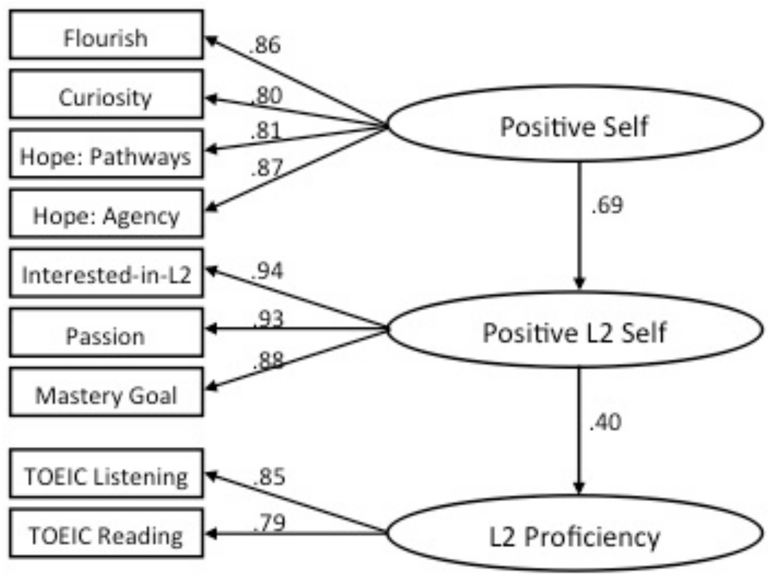

Figure 54. Structural equation model with Proficiency (TOEIC). Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation.

An alternative structural model was then considered, where positive self-construct would lead to both positive L2 self and L2 proficiency as measured by the TOEIC. Figure 55 illustrates the alternative model with both positive L2 self and L2 motivation both regressed on positive self-concept. The additional parameter was not significant. The results of the alternative model showed poorer model fit with the previous model ($X^2 = 80.66$, $p < .001$, CFI = .98, RMSEA = .068 CIs [.052, .084], AIC =140.84). The additional path from the positive self to L2 proficiency caused slightly worse model fit but it did not exceed the criterion value of ΔCFI $\leq .01$ so this showed that the two models are not statistically distinguishable.

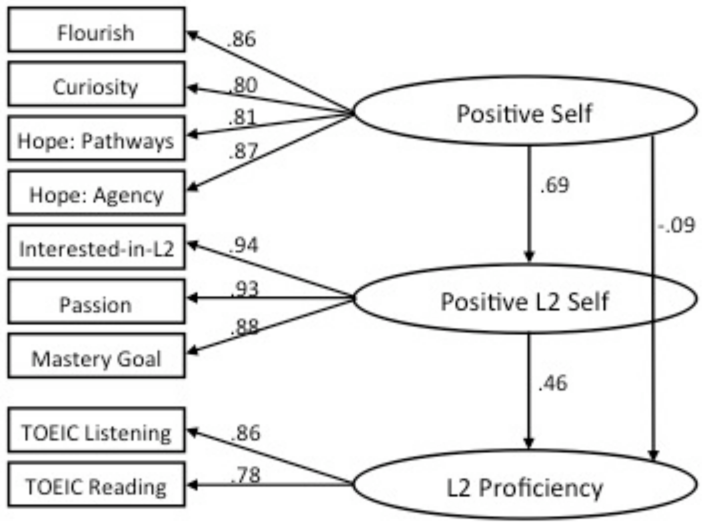

Figure 55. Alternative structural equation model with Proficiency (TOEIC). Passion = Passion for L2 Learning; Mastery Goal = Mastery Goal Orientation; Bridge = TOEIC.

Cross-Validation and Invariance Testing

Although the CFA and the SEM results confirmed the expected relationships, there were two different proficiency groups in this study. One group had taken the TOEIC Bridge (n = 221) and one had taken the TOEIC (n = 275). Theoretically, because both tests measure second language proficiency, a cross-validation study can show equivalence of causal structures. Cross-validation provides some evidence that the model is more generalizable than for a single study.

Cross-validation involves testing the components of the measurement model and structural model for invariance (also called equivalence by Byrne, 2010) across groups. There are different levels of invariance (Cheung & Rensvold, 2002;

Vandenberg & Lance, 2000). These levels or steps can be thought of as a taxonomy of invariance (Little & Slegers, 2005), which is represented in Table 99.

Table 99. *Model Comparisons of Fit Values for Proficiency Groups*

Model	CFI	RMSEA	CIs	AIC
Unconstrained	.976	.051	[.039, .063]	200.95
Measurement weights	.978	.046	[.035, .057]	188.95
Structural weights	.979	.045	[.033, .056]	184.95
Structural covariances	.979	.044	[.033, .055]	182.95
Structural residuals	.980	.042	[.031, .053]	178.95
Measurement residuals	.983	.036	[.025, .047]	160.95

Measurement variables, paths, and factorial structure are tested to see if they replicate in different groups. Configural invariance is in the measurement model step where measured variable and latent constructs have the same patterns of loadings. This is the most basic form of invariance, requiring only the same pattern of estimates among variables, not equal coefficients. The criterion for configural invariance was met by the ΔCFI $\leq .01$ (Cheung & Rensvold, 2002). Weak factorial invariance also called metric invariance or pattern invariance is an additional check to test that relative factor loadings across groups are invariant. This is tested by constraining the factor loadings of each manifest variable to be equal across groups. Comparing the ΔCFI showed less than the criterion value of .01. Strong factorial invariance, also called scalar invariance, adds the structural covariances or factorial invariance. The means of the indicators are constrained to be equal across groups. The criterion for strong factorial invariance was met by the criterion ΔCFI $\leq .01$. Finally, strict factorial invariance means that the structure and measurement residuals are invariant. Byrne (2010) noted that the last step is rarely met in practice. The residual variances were equated and met the criterion of $\Delta CFI \leq .01$ indicating strict factorial invariance. These tests showed that the model is completely invariant across the two groups.

Chapter Summary

The first part of this chapter answered the first three research questions relating to the possibility of constructing composite variables of positive self-concept, positive L2 self, L2 motivation, and L2 proficiency. CFA was first done for the three composite variables of positive self-concept, positive L2 self, and L2 motivation. The results showed that the latent factors and the measurement indicators adequately fit the model. Two additional CFAs were done, substituting proficiency measures for L2 motivation. The results again showed that the latent factors and the measurement indicators adequately fit the three factor models. Two additional alternative models were tested. The first, a one factor model with all indicators loading on a single latent variable, had poor fit to the model. The second, a two factor model with the indicators loading on two latent variables, also had poor model fit and did not meet the criterion value with a large ΔCFI value. The two models with proficiency showed no ΔCFI difference and met the criterion of $\Delta CFI \leq .01$. The SEM results answered the remaining two research questions about the structural relationships among the positive self-concept, positive L2 self, and L2 motivation and L2 proficiency. The first SEM fit a model that had a stronger path from positive L2 self to L2 motivation than positive self-concept to L2 motivation. The second SEM model showed poorer fit to a model with one less path from positive self-concept to L2 motivation. The third, fourth, fifth, and sixth models fit with paths from positive self-concept to positive L2 self to L2 proficiency, and fit with a model that had an additional path from positive self-concept to L2 proficiency, although the additional path was not significant.

Cross-validation or invariance testing was done for the two different types of proficiency measures. Changes in CFI were examined for comparative tests of group invariance. CFI fit indexes were used for comparing increasingly constrained models. The tests of measurement invariance indicated that all factor loadings, structural paths, residuals of the measured variables and factors were equivalent across the two proficiency groups and were well fitting.

CHAPTER 8

DISCUSSION

In this chapter the multiple research questions are presented individually and then the overarching aim of how constructs from positive psychology can be integrated with constructs from second language learning motivation are summarized and discussed. Finally, this chapter ends with some implications from this study.

Research Question 1: Positive Self-concept

Research question 1 asked if a composite construct of positive self-concept could be constructed. Constructs and scales of curiosity, hope, and flourishing were taken from the field of positive psychology. The individual scales underwent Rasch analysis and item level and scale level data were presented for psychometric validation. External validation evidence was presented through convergent and divergent validity with other scales that have a relationship with positive self-concept. Finally, a confirmatory factor analysis was done to show that a composite construct of positive self-concept could be constructed.

The Curiosity and Exploration Inventory II (CEI II) was adapted to match the response format of the other measures in this study. The Rasch analysis indicated overall good item fit and scale values. However, step threshold for the beginning response categories of the scale did not meet the suggested guideline so category 1 was collapsed with category 2. When this was done variance explained by the measures increased from 49.5% to 50.2% and unexplained variance in the first contrast decreased from 8.8% to 8.6%. This is a slight increase and for some purposes might not matter, but it was done here because of the importance as part of the model of positive self-concept. Although minor dimensions exist because of the multidimensional nature of curiosity and exploration, these dimensions are similar enough to contribute to a single measure and this was how it was used. This is in line with the recommendation by the creators of the measure (Kashdan, Gallagher et al., 2009) to use the CEI II as a single measure. Person reliability also slightly increased from collapsing the categories from .81 to .82. The curiosity measure showed a moderate to strong positive relationship to other measures used in positive psychology. For example, it correlated above .60 with flourishing, hope, and positive social relationships. Curiosity also showed a negative relationship to achievement hopelessness (-.34) and relationship hopelessness (-.27).

The flourishing measure was adapted to match the response format of the other measures in this study. The Rasch analysis indicated overall good item fit and scale values. However, the variance explained by the measures was a little low at 49.1%. The eigenvalue units in the first contrast was only 1.5 and the percentage of unexplained variance in the first contrast was 9.6% so the unidimensionality was considered acceptable. In other words, any secondary dimension in the data was considered low enough to not affect good measurement. The person reliability was also a little low at .75. The flourishing measure was meant to be an omnibus measure of positive flourishing so some minor dimensions and lower reliability should be expected and this is acceptable for the purpose of this study. The flourishing measure showed a strong positive relationship to other measures used in positive psychology. For example, it correlated above .70 with agency hope, self-esteem, satisfaction with life, subjective happiness, and positive social relationships. Flourishing also showed a negative relationship to achievement hopelessness (-.43) and relationship hopelessness (-.39).

The hope measure was adapted to match the response format of the other measures in this study. The Rasch analysis indicated overall good item fit and scale values. The person reliability was .81. However, the variance explained by the measures was a little low at 49.8%. This value and the fact that the hope construct is composed of two theoretically distinct components of mental willpower or agentic thinking and waypower or pathways thinking for goals (Lopez, 2013; Snyder, 1994, 2002) led to reanalysis of the scale as separate agency and pathways thinking components.

The agency-thinking component of hope on reanalysis showed good item fit and an increase in variance explained by the measures to 55.3%. The person reliability decreased to a low .71. Also, the unexplained variance in the first contrast increased to 18.1%. The eigenvalue of 1.6 suggested that the percentage value was due to the shortness of the scale with only four items. Given that the agency component of hope as used in this study is a global trait-like self-construct and the few number of items, these poorer than expected values were considered acceptable.

The pathways thinking component of hope on reanalysis showed good item fit an increase in variance explained by the measures to 60.8%. As with the other component, the person reliability decreased to .77. Also, the unexplained variance in the first contrast increased to 15.1%. However, again as with the other hope component, the low eigenvalue of 1.5 suggested that the percentage value was due to the shortness of the scale with only four items. Given that the pathways component of hope as used in this study is a global trait-like self-construct and the few number of items, these poorer than expected values were considered acceptable.

Both components of hope measures showed a moderate to strong positive relationship to other measures used in positive psychology. For example, they correlated above .60 with curiosity, flourishing, self-esteem, and each other. The agency component of hope also correlated above .60 with satisfaction with life, subjective happiness, positive social relationships and both perseverance and passion components of grit. The two components of hope also showed a negative relationship to achievement hopelessness and relationship hopelessness.

The results of the preliminary analyses showed that the measurement scales had good psychometric properties at both the item and scale levels. The preliminary analyses of the relationships among peripheral variables related to positive psychology provided validity evidence external to the main positive psychology variables (Loevinger, 1957). Negative relationships were found for variables that indicate a dysphoric self. The positive and negative relationships provide evidence that the positive self-variables are situated in a nomological network of positive psychology constructs. In other words, even though all constructed measures were self-reports, evidence was provided that measures were found to be distinctive and positively or negatively related to similar or dissimilar constructs.

In sum, the results of the confirmatory factor analysis indicated that a composite positive self-concept could indeed be constructed from the components of curiosity, flourishing, and the two components of agentic hope and pathways hope with adequate fit to the model. This was possible when either L2 proficiency or L2 motivation was included as part of the overall model.

Research Question 2: Positive L2 Self

Research question 2 asked if a composite construct of positive L2 self could be constructed. Constructs and scales of interest in L2 self and passion for L2 learning were constructed based on theories of positive psychology. A construct and scale of mastery goal orientation was adapted from the field of achievement goal theory. The individual scales underwent Rasch analysis and item level and scale level data were presented for psychometric validation. External validation evidence was presented through convergent and divergent validity with other scales that have a relationship with positive L2 self. Finally, a confirmatory factor analysis was done to show that a composite construct of positive L2 self could be constructed.

The Rasch analysis of the interested in L2 self scale indicated that one of the eight items misfit both the infit and outfit mean-squares outside the targeted range of .5 to 1.5. In addition, this item had a much lower point-measure correlation so this item was dropped from the scale. On reanalysis, the remaining seven items showed good item fit statistics and scale characteristics. The variance explained by the measures was 64.9%. The person reliability was .87.

The interested in L2 self scale showed strong correlations of .80 or over to passion for L2 learning and mastery goal orientation. It also showed a .75 correlation with another L2 domain variable, the ideal L2 self. It correlated at .55 with the prosociality goals scale and showed a nonsignificant relationship with the math self-concept scale. The math self-concept scale was not expected to have a relationship with the interest in L2 self scale. However, if students were responding to the survey in what they might have seen as a socially desirable way, that is, to please the survey giver, or in response sets that failed to distinguish among item types, then a positive relationship would be found. The fact that there is no relationship suggests that students were distinguishing between item types and were not exhibiting a strong social desirability bias for the sake of appearances.

The Rasch analysis of the passion for L2 learning scale showed good item fit and scale characteristics. The variance explained by the measures was 64.2% and the person reliability was .87. The passion for L2 learning scale showed a .80

correlation with mastery goal orientation and ideal L2 self. It showed a moderate correlation with prosociality goals at .55 and a non-significant relationship with math self-concept.

The Rasch analysis of the mastery goal orientation scale showed that one item did not fit the model, falling outside the targeted range of .5 to 1.5 for both the infit and outfit mean-square. Furthermore, it had a lower point-measure correlation so this item was dropped from the study. Upon reanalysis, the remaining seven items all fit well within the targeted guidelines. The scale characteristics were also good. The variance explained by the measures was 61.6% and the person reliability was .85. The mastery goal orientation scale showed strong correlations of .80 or over to passion for L2 learning and mastery goal orientation. It also showed a .66 correlation with ideal L2 self and .53 correlation with prosociality goals. The mastery goal orientation had a weak relationship to math self-concept with a correlation of .12.

The results of the preliminary analyses showed that the measurement scales as components of a positive L2 self had good psychometric properties at both the item and scale levels. The preliminary analyses of the peripheral variables provided validity evidence external to the main variables of relationships to other variables shown to be related to a positive academic self or identity. A weak relationship was found between mastery orientation and math self-concept. The positive and negative relationships provide evidence that the positive L2 self variables are situated in a nomological network of positive L2 domain constructs.

The results of the confirmatory factor analysis indicated that a composite positive L2 self could be constructed from the components of interested in L2 self, passion for L2 learning, and master L2 goal orientation with adequate fit to the model. This was possible when either L2 proficiency or L2 motivation was included as part of the overall model.

Research Question 3: L2 Motivation

Research question 3 asked if a composite construct of L2 motivation could be constructed. Constructs and scales of L2 speaking self-efficacy, L2 listening self-efficacy, and L2 reading were created based on previous research in second language studies, educational psychology, and positive psychology. The individual scales underwent Rasch analysis, and item level and scale level data were analyzed for psychometric internal validation. External validation evidence was presented through convergent and divergent validity with other scales that have a relationship with L2 motivation. Finally, a confirmatory factor analysis was done to show that a composite construct of L2 motivation could be constructed.

The Rasch analysis of the L2 speaking self-efficacy scale showed good item fit and scale characteristics. The variance explained by the measures was 65.1% and the person reliability was .87. The L2 speaking self-efficacy scale showed a .82 correlation with L2 reading self-efficacy and .83 with L2 listening self-efficacy. It showed a moderate correlation with intended learning effort at .66 and a .70 correlation with persistent effort.

The Rasch analysis of the L2 listening self-efficacy scale showed generally good item fit and scale characteristics. The variance explained by the measures was 60.7% and the person reliability was .87. However, the lower than expected step advancement and relatively low count for category 6 suggested that this category be collapsed with category 5. It should be noted that it is not uncommon for extreme categories to be less endorsed. The categories at the ends of the response options are designed to be more difficult to endorse to prevent floor or ceiling type effects. With three or four categories, collapsing might constrain measurement. The scales in this study all used six categories and collapsing categories, if all other indicators of quality are still high, does not cause a problem. After collapsing the last two categories, the variance explained by the measures was again 60.7% and the person reliability was again .87 although the Andrich thresholds advanced better for more separation and the outfit mean square dropped considerably for improved scale stability. The L2 listening self-efficacy scale showed a .84 correlation with L2 reading self-efficacy and .83 with L2 speaking self-efficacy. It showed a moderate correlation with intended learning effort at .64 and a .66 correlation with persistent effort.

The Rasch analysis of the L2 reading self-efficacy scale showed generally good item fit and scale characteristics. The variance explained by the measures was 60.2% and the person reliability was .86. However, similar to listening self-efficacy, the lower than expected step advancement and relatively low count for category six suggested that this category be collapsed with category five. After collapsing the last two categories, although the person reliability did not improve, the variance explained by the measures improved to 60.5% and the Andrich thresholds advanced better for more separation and the outfit mean square dropped considerably for improved scale stability. The L2 reading self-efficacy scale showed a .84 correlation with L2 listening self-efficacy and .82 with L2 speaking self-efficacy. It showed a moderate correlation with intended learning effort at .66 and a .70 correlation with persistent effort.

The results of the confirmatory factor analysis indicated that a composite construct of L2 motivation could be constructed from the components of L2 speaking self-efficacy, L2 listening self-efficacy, and L2 reading self-efficacy with adequate fit to the model. Although L2 motivation in this study is based on self-efficacy, it is not a general measure of L2 self-efficacy. This study measured with scales based on specific L2 skills and items based on specific tasks. It is possible to imagine someone with very high language aptitude endorsing a L2 learning general item or scale that measured the ability to learn another language in general, which might be called L2 language learning self-efficacy, but that is not the approach taken here.

Research Question 4: Structural Relationships with Motivation

Research question 4 asked to what extent does positive self-concept, affect L2 motivation with positive L2 self as a moderator? To review briefly, as discussed in the literature review, positive self-concept is a general view of the total self that is relatively stable and trait-like. This is not to say that it is not changeable, just that change might take time and interventions or treatments might need some degree of frequency and perseverance over time. Also, like other stable aspects of the self, it does not mean that an individual always behaves in a deterministic and mechanical manner. Thus, positive self-construct refers to general tendencies, not a fixed pattern of absolutes.

A positive L2 self is a domain level self where the domain is second language learning. It is less general than positive self-concept but as a motivational construct distant from language learning behavior. A positive L2 self requires some degree of developing expertise in order to delineate the domain of second language learning and this also gives it a degree of stability. In other words, a novice with almost no experience could have a positive language learning experience but until enough of these experiences are strung together they are experienced as isolated incidents or situations. With enough repeated experience, competence develops and the type of experience is recognized as belonging to a particular domain. Again, as with positive self-concept, this does not mean that it is unchangeable or that it determines behavior in an absolute fashion.

L2 motivation, in contrast to positive self-concept and positive L2 self, is proximal rather than distant to language learning activity and achievement. The measured variables for the latent construct are specific to language skills and the items to specific activities. The variables reference the self but a highly specific self that is competent to complete language tasks.

The structural relationship is from the general to the specific, just as general psychological traits lead to a relative tendency of an exhibition of psychological states. The positive L2 self is in effect a moderating variable between the most general positive self and a highly specific motivational self. As the moderating variable, or middle-level variable, the structural relationship should be stronger between it and the variables it moderates. In addition, because the three levels of specificity are all self-referenced variables, positive self and L2 motivational self should show a relationship. The structural equation model showed that these were indeed the relationships found. The positive self-concept to positive L2 self path showed a standardized beta weight of .69 and the positive L2 self to L2 motivation path showed a standardized beta weight of .45. The positive self-concept path to L2 motivation showed a weaker standardized beta weight of .34.

Research Question 5: Structural Relationships with L2 Proficiency

Research question 5 asked to what extent does positive self-concept affect L2 proficiency with positive L2 self as a moderator? As with L2 motivation, a positive self-concept is more distal and a positive L2 self is a mediating variable that bridges self and L2 so the directionality should be similar to that in the model answering research question 4. Unlike the motivational measures of self-efficacy, L2 proficiency does not directly reference the self and also is not a self-report measure. One possible problem with creating a model with only subjective self-reports is that there might be a method effect where there is common variance due to a singular method of collecting data. The use of an objective outcome measure, L2 proficiency, showed that the model fit is not solely due to method effects. The directional relationships should again be between the positive self-concept and positive L2 self and then the positive L2 self and L2 proficiency, but not between positive self-concept and L2 proficiency, because the latter is more distant or removed from self-concept, especially because it lacks reference to the self. The structural equation model showed that these were indeed the relationships. The positive self-concept to positive L2 self path showed a standardized beta weight of .69 and the positive L2 self to L2 proficiency path showed a standardized beta weight of .51.

Further validation evidence for the model was provided by the cross-validation study. There were two different groups in this study: one group had taken the TOEIC Bridge and one had taken the TOEIC. By testing the measurement model and structural model for invariance or equivalence across groups, additional evidence for the generalizability of the model was provided.

General Discussion

Two aspects of this study in particular merit further discussion. They relate to the self-focused nature of this study and the role of goals.

It is important to note that the constructs and variables at different levels of generality are concerned with the self, and this concern raises questions. For example, does this imply some sort of selfishness or self-focused individuality at the expense of others? In addition, as noted in the literature review, perfectionist selves, ideal selves, some forms of self-esteem, and passionate selves might be maladaptive. In this study, care was taken to include variables showed to be adaptive in previous studies. So, for example, harmonious L2 passion was measured rather than obsessive L2 passion. Harmonious L2 passion is characterized by being in harmony with other life domains and values including positive social relationships with others while obsessive L2 passion would be a more isolated passion that conflicts with and excludes other life domains and values.

In addition to the care in choosing socially adaptive variables, part of the validation process included showing relationships with peripheral variables that measured positive sociality. This was done at the general global level with the positive social relationships measure and prosociality goals measure at the language learning domain level. Correlations were moderate to high with the modeled variables. Although concepts of self run though this model, it should be understood as an adaptive self that is socially conscious and is positively related to social relationships.

Another important aspect the study is the implicit role of goals. At different levels of specificity goals are related to some variables. Goals are related to the construct of hope, which consists of two subcomponents of agentic thinking (that one has the agency to pursue goals) and pathways thinking (that one can find a path even in the face of problems toward a goal).

The mastery goal orientation construct relates beliefs one has about the purpose of pursuing goals, as the name implies. This goal orientation is concerned with an absolute mastering of material to be learned and increasing competence. In achievement goal theory, mastery goal orientation is often contrasted with performance goal orientation that is based on displaying competence of relative normative standards.

Goals also feature in the self-efficacy variables, which refer to competence beliefs about specific activities. In a sense, they are competence beliefs about goal-like tasks. Self-efficacy targets can be so proximal that they do not seem like goals but this proximalness also gives them predictive strength as motivational variables.

Theoretical Implications

This study has implications for applying positive psychology to L2 learners in a Japanese context at three levels of specificity. The first part of this study established that aspects of the self for Japanese students can be reliably and precisely measured on single dimensions at all three levels. Composite constructs of positive self-concept, positive L2 self, and L2 motivation can be constructed. The second part of the study suggests that at the global positive self level, relatively new constructs from positive psychology have direct relationships to a positive L2 self and L2 motivation, and, therefore, confirm the promise of applying positive psychology to motivation and learning.

The implications of the two structural models are that a global positive self-concept does not have a direct relationship to L2 proficiency but it does have a direct relationship to a positive L2 self and L2 motivation. This was expected because there are many pathways to a positive self-concept. For example, a person might excel at a musical skill, a sport, or an academic domain that feeds into the person's general sense of self. However, a teacher could try to develop connections among positive self-concept and positive L2 self and L2 motivation.

Pedagogical Implications

A question might arise regarding whether it is always desirable to have a positive self-concept, positive L2 self, and L2 motivation. My answer is that a positive self-concept is desirable for every individual but there are many paths, for example, through art, music, or a particular academic domain, with L2 learning being one of them. A positive L2 self is desirable for anyone learning a language but the content and strength differs for different people. For example, someone who wants to learn an L2 for travel might have a different positive L2 self than someone who wants to learn an L2 for work purposes, which in turn differs from that of someone who eventually wants to teach the L2. Second language motivation on the other hand should be high for anyone learning a second language because of the direct connection between the two. Not everyone needs to be in the process of learning an L2, but if they are it is desirable to have a high level of motivation.

In the language learning literature there have been calls to integrate self-esteem (Rubio, 2007) and happiness (Helgesen, 2006), into the language classroom. More specific implications are that in addition students could be made aware of other constructs that contribute to a global positive self, such as: curiosity (Kashdan, 2009), flourishing (Diener et al., 2010; Seligman, 2011), and hope (Lopez, 2013; Snyder, 1994, 2000a, 2000b). In addition to the main variables, a number of peripheral variables were adapted or constructed and found to have good measurement properties with the sample in this study. Although the peripheral variables were used to provide external validation evidence, future studies could explore the relationship of one or more of these variables and some aspect of language learning or motivation. For example, self-control or grit (Duckworth, Peterson, et al., 2007; Baumeister & Tierney, 2011) showed relationships in other learning domains and it might be fruitful to explore if there is a relationship to the language learning domain beyond that shown here. It might be that the global grit construct has a stronger relationship to the measure of persistent effort for learning an L2 (in effect grit for L2) and this might have a stronger relationship to language learning. In other words, future studies could follow the specificity pattern in this study to test additional models.

There have been many studies in the last decade of the ideal L2 self and as pointed out in the literature review, an ideal L2 self or ought-to L2 self might be problematic in the Japanese context. One problem is that as in self-discrepancy theory, the ought-to self has been linked in psychological studies to depression, anxiety, distress and mental disorders (Cornette, Strauman, Abramson, & Busch, 2009; Higgins, Klein, & Strauman, 1985; Strauman, 1989; Strauman & Higgins, 1998). Ideals feature prominently in perfectionism, which can have an adaptive type or maladaptive type (Blatt, 1995; Hamachek, 1978; Hewitt & Flett, 1991; Nugent, 2000). For example, maladaptive perfectionists have high ideals that are often difficult to meet leading to dysphoria. Maladaptive perfectionism is related to being afraid of making mistakes, hesitations in making decisions and acting on them, and a host of pathological and mental health problems. In the Japanese educational context where students are often represented as being reticent, afraid of making mistakes, or generally demotivated, then it might be that for some students the discrepancy between ideal and actual self is so great that it does not lead to motivation but withdrawal. In discrepancy theory, for some people the ought-to self causes anxiety and ideal self causes depression (Higgins, 1987, 2012). In Japan, the phenomenon of *hikikomori* or self-seclusion due to anxiety, is recognized as a growing problem (Furlong, 2008; Teo, 2010). Suicide and attempted suicide rates, due in part to depression, in Japan are among the highest in the world and have remained a problem in recent years (Hidaka, Operario, Takenaka, Omori, Ichikawa, & Shirasaka, 2008; Nakao & Takeuchi, 2006). Teachers need to be aware that self-discrepancies can in some situations be maladaptive and they need to be careful when advising students.

Positive psychology is not about trying to be happy all the time or about putting a positive “spin” on negative experiences (Kashdan & Biswas-Diener, 2014; Peterson, 2006; Peterson & Seligman, 2004). Being “positive” when actually feeling negative is something like lying to oneself and creates unnecessary negative internal tensions. This can happen with imagined futures or present selves.

A problem with anticipating future events and conditions is that people tend to misjudge or be wrong about the positive effect on well-being. People tend to be poor affective forecasters because they are prone to cognitive biases that cause them to inaccurately predict outcomes (Gilbert, 2006; Lyubomirsky, 2013; Wilson & Gilbert, 2005). For example, people are prone to overestimate how happy they would feel if they win a prize. This is called the impact bias and it affects the intensity and duration of estimated emotion.

Research has found that variables in positive psychology outperform variables such as intelligence or aptitude (Duckworth & Seligman, 2005, 2006; Seligman, 2011). This suggest that there is support for Friedman (2007) advocating that in the future education should stress curiosity and passion as mattering more than intelligence.

As previously noted, positive psychology is not about wishful thinking or self-deception but is concerned with being authentic with your true self (Harter, 2002; Peterson, 2006; Schlegel, Hicks, Arndt, & King, 2009). People are susceptible to many biases and self-enhancing strategies that cause self-distortions (Leary, 2004). As Harter (2012) has noted, “authenticity of the self can be compromised by the tendencies to inflate, becloud, and distort the real inner self" (p. 329). Self-reflection and self-understanding can help to uncover the real self. The authentic true self has a relationship to well-being and meaningfulness in life (Schlegel & Hicks, 2011). As Seligman (2011) has stated, "well-being is a combination of feeling good as well as actually having meaning, good relationships, and accomplishment" (p. 25).

As in positive psychology, an authentic positive L2 self is based on similar elements of authenticity. One element has to do with discovering oneself through self-understanding and self-acceptance (Ryff & Singer, 2008). Another element is self-realization (Ryan & Deci, 2001) or what the humanistic psychologists called actualizing tendency or self-actualization (Maslow, 1968, 1970; Rogers, 1951, 1961). Identifying and using strengths is another element that can be realized in many different constellations (Linley, 2008; Proctor, Maltby, & Linley, 2009; Peterson & Seligman, 2004; Weber, Wagner, & Ruch, in press). Curiosity, hope, and flourishing are strengths for learning but they can be combined with other strengths such as kindness, fairness, or humor for example, depending on the individual. In addition there are the many specific elements that make up one's values, purposes, goals, interests, and passions (e.g., Emmons, 2003; Kashdan, 2009; Silvia, 2006). Alignment of these various elements leads to a more authentic self and higher levels of well-being (Sheldon, 2002, 2004, 2014).

Developing a sense of agency, competence, learning, and enjoyment is an important part of education and unites the variables that are a part of the presented model. These variables have a thread of learning or acquiring novelity/complexity running through them. Underlying self-efficacy is a personal sense of agency and competence. The same is true of hope, although at a more general trait-level in my model. Agency, competence, learning, and enjoyment underlie domain-specific interest, passion for learning, L2 mastery goal orientation, curiosity, and is a large part of flourishing.

Combining an authentic self with connections to language learning can form an authentic positive L2 self. In a different context, Hudson (2008) suggested that in addition to having lots of fun in a language lesson we should also "make the language itself interesting" (p. 111). Academic linguists "almost by definition are driven by interest in language" (p. 111) and are also a potential source of ideas. Even for people who are not learning a language, there is a general interest in learning about language as attested by popular books such as, for example, *Mother Tongue* (Bryson, 1990), *The Story of English* (McCrum, MacNeil, & Cran, 2002), *The Stories of English* (Crystal, 2004), or *The Adventure of English: The Biography of a Language* (Bragg, 2003). The different skills and aspects that make up the language system, for instance, the sound system, grammar, vocabulary, et cetera, make it difficult to learn but also provide an endless source of novel things to learn. When language is coupled with other topics, as in the many language-and-fields, for example, language-and-history, language-and-culture, language-and-learning, language-and-interaction, language-and-psychology, et cetera, the possibilities for fun and interesting avenues to discover and explore becomes even greater. Learner connections to the language can be made in a variety of ways. To sum up, in other words, an authentic positive L2 self approach means becoming meaningfully involved and engaged in the here-and-now with a topic that can provide fascinating unlimited learning, that is, language.

Instead of focusing on a narrow future ideal L2 self, students can be helped to develop a more expansive hopeful self, and a present interest in an L2 so that with an appropriate level of language competence and challenge they continue to be interested in the L2; another approach would be to develop harmonious L2 passion through meeting student needs of competence, autonomy, and relationships with others; teachers can help learners develop mastery or learning goal orientation by not making relative comparisons of learners but stressing the importance of all students making incremental gains and giving feedback that develops competence (Da Silva, 2007; Dweck, 2000). In addition to focusing on learning goals, teachers can help learners develop practices that directly relate to learning, such as breaking up a distal goal into proximal sub-goals, practicing, time management techniques, staying aware of the importance of persistent effort over long time periods. Teachers can help students increase their self-efficacy by regularly measuring student progress through sensitive assessments and giving feedback that demonstrates competence, thus making learning gains salient (Brown & Hudson, 2002; Da Silva & McInerney, 2008; Rouault, 2007).

The measures in this study all have variances that are then correlated with other measures; in other words, there are students that are high or low on a measure and many of these students are correspondingly high or low on other measures as suggested by the correlations. So implicit in this study is that students have a wide range of motivational levels running in parallel. The correlations show that those higher in the positive L2 self constructs are higher in global positive self and the motivational variables. This shows that many students are highly intrinsically motivated. Teachers and researchers need to be sensitive to differences among students and not lump them all into a demotivated category (Da Silva & McInerney, 2008). Clearly, again as the correlations show, some are much less motivated than others and teachers could potentially help students such as these.

As mentioned in the literature review in Chapter 2, there are many gender differences in how education and motivation is perceived in Japan (Okano, 2009; Sugimoto, 2010). A university education for female students in Japan might be important for its symbolic value of higher social status and cultural sophistication (Okano & Tsuchiya, 1999). An English language identity especially might be valued by adding an element of international sophistication. For some Japanese women, the symbolic value can facilitate a practical form of resistance to traditional expectations. Japanese women can use an L2 self to resist gender expectations, create additional paths to pursue and provide more flexible standards and lifestyles than have been traditionally available to them (Kelkey, 2001). On the other hand, an L2 self can also conflict with values held to be important to some women. L2 identity or ego has long been recognized to relate to differing levels of L2 achievement (Brown, 1973). Considering this gender background in Japan and the female participants of this study, this might have contributed to the relationship between a positive L2 self and L2 motivation and L2 proficiency.

Also implicit in the results of this study are the importance of competence, novelty/challenge, and goals. Competencies are part of the evaluations that produce such beliefs of self-efficacy, interest, and self-esteem. Elliot and Dweck (2005) note that competence is not only the core of achievement motivation "but also a foundational building block for any theory of personality, development, and well-being" (p. 8). Novelty/challenge is an important component of mastery goal orientation, interested L2 self, and passion for L2. Goals are important because they give direction to generated motivational energy and guide self-regulation (Zimmerman, 2008; Zimmerman & Schunk, 2011). Why certain goals are held also signals important information about implicit self-beliefs and potential strategies for the goals (Dweck, 2000; Dweck & Leggett, 1988; Mercer & Ryan, 2010).

Chapter Summary

This chapter discussed the research questions and some implications of this study. The first section of the chapter discussed research question 1, which showed that a composite positive self-concept could be constructed from components of curiosity, flourishing, and two components of agentic hope and pathways hope. The second section discussed research question 2, showing that a composite construct of positive L2 self could be constructed from components of interest in L2 self, passion for L2 learning, and mastery goal orientation. In section three, research question 2 was discussed, elaborating on its findings that L2 motivation could be constructed from components of L2 speaking self-efficacy, L2 listening self-efficacy, and L2 reading self-efficacy. Research question 4 in section four discussed relationships among the latent variables. Positive self-concept showed the strongest relationship with positive L2 self and a weak relationship to L2 motivation. Positive L2 self showed a stronger relationship to L2 motivation than did positive self-concept. These relationships confirmed that more proximal relationships have stronger relationships than the distal relationships. In the fifth section, the fifth research question was discussed with results that showed relationships among positive self-concept, positive L2 self, and L2 proficiency. Positive self-concept showed a strong relationship to positive L2 self but a non-significant relationship to L2 proficiency. Positive L2 self had a moderate relationship to L2 proficiency. Further validation was done through cross-validation of this model for invariance with two different groups that had taken different proficiency measures, the TOEIC Bridge and the TOEIC. The tests of measurement invariance indicated that all factor loadings, structural paths, residuals of the measured variables and factors were equivalent across the two proficiency groups and were well fitting. The next section was a general discussion that related to the self-focused nature of this study and the role of goals. Theoretical implications were discussed in the next section and offered technical implications for positive psychology models. The final section discussed practical implications for learners and teachers, including how the particular features of Japanese society and gender relations need to be considered. It is important to build competence in learners so that competence beliefs are developed leading to even higher levels of competence. Over time and effort more stable dispositional identity beliefs can develop a positive L2 self. Teachers and learners can be joyful to know that in addition to a second language, it is also possible to develop a positive L2 self, and this might contribute toward an overall positive self-concept that helps them flourish in other aspects of life.

CHAPTER 9

CONCLUSION

In this chapter a summary of the results are presented, limitations of the study are discussed, and suggestions for future research are given. Finally, some concluding thoughts are provided.

Summary of the Results

This study addressed multiple research questions with the overarching aim of demonstrating how constructs from positive psychology can be integrated with second language learning motivation theory and research. An important principle underlying the study is that the specificity of the constructs and measurement variables needs to be clarified and then modeled appropriately. For this study, specificity levels included a general global level, where positive self-concept referenced the whole person; a more specific domain level, where positive L2 self referenced the person within the L2 domain; and even more specific L2 motivations and proficiency.

Composite latent variables for the different levels were constructed with measured variables that went through a rigorous validation process. Internal validity evidence was gathered through Rasch analysis that looked at item and scale characteristics to produce measures with sound psychometric properties. External validity evidence was gathered through correlational analysis to show convergent and divergent relationships and to situate the variables into a nomological network of variables at the same level of specificity. In other words, a nomological network shows both the distinctiveness of variables and how they relate to similar or dissimilar variables in related research traditions.

Confirmatory factor analysis indicated that the latent variables fit a measurement model with the measured variables. To answer the first three research questions CFA was done on three composite latent variables of positive self-concept, positive L2 self, and L2 motivation. Two models showed adequate fit, one with L2 motivation as the outcome latent variable and one with L2 proficiency as the outcome latent variable.

Structural models then revealed more clearly and rigorously the relationships among the constructs of positive self-concept, positive L2 self, and L2 motivation, as well as L2 proficiency in place of L2 motivation to answer research questions 4 and 5. For the L2 proficiency model, a cross-validation study demonstrated model equivalence/invariance. This series of structural equation models provides strong support for the generalizability of the findings to similar populations.

Overall, this study showed how constructs from positive psychology can be integrated with second language motivation theory. Positive psychology is a rapidly growing area in the general field of psychology and educational psychology. This study shows one way it might be applied in the field of second language learning. In addition, this study developed many measures as both main and peripheral variables. These variables can provide a useful starting point for exploring other research avenues in positive psychology and L2 motivation and learning.

Limitations

One limitation is that this study's use of many scales that have previously been used in positive psychology research might have sometimes limited measurement precision. Although their overall performance showed that established previous research is applicable to this sample in Japan, the scales that were developed for this study tended to have better psychometric properties. Some of the established scales can be improved in the future by adding more items. Creating a longer instrument can produce improved measurement.

Another limitation of this study was that two of the scales had poorly performing items. Rasch analysis detected an item in the Satisfaction with Life scale that had poor fit statistics. As noted in Chapter 4 it is known that the satisfaction with life scale contains a minor secondary dimension (Diener, Inglehart, & Tay, 2013; Slocum-Gori & Zumbo, 2011; Slocum-Gori, Zumbo, Michalos, & Diener, 2009). In the Subjective Happiness scale, Rasch analysis detected an item that had poor fit statistics. As with the previous item, while it was not detrimental to measurement, it also did not add useful information to the scale and in future studies with similar samples it can be deleted. These two examples show that it is important to check item functioning, even with well-established scales.

My use of the hopelessness in achievement and hopelessness in relationships were another limitation of this study. Hopelessness is often used with distressed patients in psychiatric settings (Beck, Steer, Beck, & Newman, 1993; Beck, Steer, & Garbin, 1988). Hopelessness has been shown to predict suicide (Beck, Brown, & Steer, 1989). Unless one is doing research on distressed students or has the means to provide interventions, then it would be best to seek a measure more appropriate to the particular study to provide validity evidence. This conclusion was also reached by Steed (2001) who recommended that the hopelessness scale should not be used with normal populations.

The use of L2 speaking self-efficacy as a component of L2 motivation might represent a limitation as speaking can interact with other constructs in ways that were not modeled in this study. For example, speaking is a productive skill that might be easier for those in social situations, that is, those that are more extroverted. People can identify themselves as more positive at both the self level and positive L2 self level if they are more extroverted. In this study, the personality trait of extroversion was not measured or modeled.

Finally, as mentioned in Chapter 2, there are many constructs that are included under the umbrella of positive psychology. The present study modeled only a small subset of possible variables. Many other models are possible, both with the variables used in this study and additional variables not presented.

Suggestions for Future Research

A longitudinal study can provide additional useful insights. Long term follow up studies could be done to see how developments in emerging adulthood create more lasting effects for more lifelong learning and life goals. Over periods of time the path directionality should be reciprocal rather than unidirectional as in this study. In the future, a longitudinal study could be done to model the reciprocal effects of positive self with positive L2 self and positive L2 self with achievement measures or to motivational variables proximal to achievement. In other words, over time successful L2 learning should build positive competence beliefs and a more stable L2 domain self.

This study relied heavily on quantitative data. Qualitative data for a mixed methods study would have provided a richness beyond what the numbers alone could supply. It would be useful to explore the reasons students were strong or weak in particular dimensions. It would also be interesting to learn what particular goals students have for themselves and where they see themselves in relation to the goals. A deeper understanding could help teachers and councilors with better guidance to students in the future.

Studies on different populations would give additional insights into such things as gender or age effects. Further studies could be done in other countries where the culture is different from Japan.

If a longitudinal study were done that showed that L2 learning contributes toward a positive L2 self and in turn can contribute toward a positive self-concept, this would have many implications for learning and teaching. It would mean that learners are not only learning content but also developing positive identities. Just as positive self-concept is not determined by positive L2 self, but might contribute toward the positive L2 self, a positive self-concept might contribute toward other life domains. L2 teachers might also find this motivating because it would give them new perspectives on the potential importance of the day's lesson on, for example, some grammatical point or set of vocabulary words. If over time these lessons build competence and competence beliefs, and these in turn develop more dispositional selves then the importance and value of teaching can become greater.

This study was done in the second language domain. Researchers and teachers in other academic domains could do similar studies. For example, in math research could be done to see if positive self-concept, positive math self-concept, and math motivation were related. Studies on developmental expertise could also be done to compare related development in self or identity.

In addition, the peripheral variables used to support the main variables in this study are interesting in their own right. For example, a study relating grit, persistent effort and language learning could be explored and a model could be created.

Some variables that were measured at the trait level, such as hope, could be explored at the state level. For example, it might be fruitful to explore state hope with motivation and learning.

Furthermore, similar studies done on specific skills would give students and teachers more precise knowledge of how L2 self-variables interact with L2 motivation and skill learning. One such initial study was done for L2 reading (Lake, 2014). In the longitudinal study over fifteen weeks, students doing extensive reading were found to gain in reading speed, in L2 reading self-efficacy and positive L2 reading self. Studies such as this could be done in other L2 skill areas.

Final Comments

In the last few years, there have been many popular books in English published about education, learning, and motivation. Many of these books are by journalists or writers who are summarizing some of the recent research done in the field of positive psychology. Examples of such books include: *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character* by Paul Tough (2012); *Outliers: The Story of Success* by Malcolm Gladwell (2008); *Drive: The Surprising Truth about What Motivates Us* by Daniel Pink (2009); *The Smartest Kids in the World and How They Got That Way* by Amanda Ripley (2013). These books attest to the current public interest in personal growth, improvement of education and student achievement.

These bestselling books report on research based on what are often called: non-cognitive skills or factors (e.g., Heckman & Rubinstein, 2001; Heckman, Stixrud, & Urzua, 2006); character or character strengths (e.g., Peterson & Seligman, 2004); 21st century skills (e.g., Autor, Levy, & Murnane, 2003; Marzano

& Heflebower, 2012); or self-theories or mindsets (e.g., Dweck, 2000, 2006). These other terms are more often used in different or narrower contexts even though there might be considerable conceptual overlap and use. For example, non-cognitive factors are terms often used in economics research and 21st century skills are terms often used in the context of the rapidly changing technological future. Thomas Friedman (2007) writing about the future of education created the equation "CQ + PQ > IQ", where CQ and PQ equals curiosity and passion quotient and that they matter more than intelligence quotient. He stated that, "Curious, passionate kids are self-educators and self-motivators" (p. 314). In this study the terms used were framed in the context of positive psychology and second language motivation. This study provides one way for concerns at different levels to be linked and understood.

Learning a second language is a good example of learning a complex skill through much time and effort. Teachers and administrators tend to focus on how the content, communicating in a second language, is important for current and future academic purposes and future career development. This study shows that the beyond academic and career purposes, learning processes can help in developing positive identities for personal growth and lead to flourishing lives.

REFERENCES

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358-372.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.

Adler, A. (1926). *The neurotic constitution*. (B. Glueck & J. Lind, Trans.) New York, NY: Dodd, Mead. (Original work published in 1912)

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review, 18*, 391-405. doi: 10.1007/s10648-006-9033-0

Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher, 32*(8), 10-14. doi: 10.3102/0013189X032008010

Alexander, P. A. (2004). A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 273-298). Mahwah, NJ: Erlbaum.

Alexander, P. A., & Winne, P. H. (Eds.). (2006). *Handbook of educational psychology* (2nd ed.). Mahwah, NJ: Erlbaum.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Student learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260-267.

Anderman, E. M., & Wolters, C. A. (2006). Goals, values, and affects: Influences on student motivation. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed.) (pp. 369-389). Mahwah, NJ: Erlbaum.

Ansbacher, H. L. (1970). Alfred Adler: A historical perspective. *American Journal of Psychiatry, 127*, 777-782.

Arbuckle, J. L. (2007). AMOS (Version 16.0). Chicago, IL: SPSS.

Aristotle (1974). *The Nichomacean ethics.* (J. A. K. Thomson, Trans.) Harmondsworth, NY: Penguin. (Original work written in 350 B.C.)

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*, 469-480. doi:10.1037//0003-066X.55.5.469

Arnett, J. J. (2004). *Emerging adulthood: The winding road from the late teens through the twenties*. New York, NY: Oxford University Press.

Arnold, J., & Brown, H. (1999). A map of the terrain. In J. Arnold (Ed.), *Affect in language learning* (pp. 1-24). Cambridge: Cambridge University Press.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*, 359-372.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin, 120*, 338-375.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics, 118*(4), 1279–1334. doi:10.3386/w8337

Bachman, L. F. (1989). The development and use of criterion-referenced tests of language ability in language program evaluation. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 242-258). Cambridge, England: Cambridge University Press.

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

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26. doi:10.1146/annurev.psych.52.1.1

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*(2), 164-180. doi:10.1111/j.1745-6916.2006.00011.x

Bandura, A. (2008). An agentic perspective on positive psychology. In S. J. Lopez (Ed.), *Positive psychology: Discovering human strengths* (pp. 167-196). New York, NY: Praeger.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest though proximal self-motivation. *Journal of Personality and Social Psychology, 41*, 586-598.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462-479.

Barnette, J. J. (2000). Effect of stem and Likert response option reversals on survey on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*(3), 361-370. doi:10.1177/00131640021970592

Baumeister, R. F. (2005). *The cultural animal: Human nature, meaning, and social life.* New York, NY: Oxford University Press.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4,* 1-44. doi:10.1111/1529-1006.01431

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2005). Exploding the self-esteem myth. *Scientific American, 292,* 84-92. doi:10.1038/scientificamerican0105-84

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497-529.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103,* 5-33.

Baumeister, R. F., & Tierney, J. (2011). *Willpower: Rediscovering the greatest human strength*. New York, NY: Penguin.

Baumeister, R. F., & Vohs, K. D. (Eds.). (2004). *Handbook of self-regulation: Research, theory, and applications*. New York, NY: Guilford.

Beck, A. T., Steer, R. A., Beck, J. S., & Newman, C. F. (1983). Hopelessness, depression, suicidal ideation and clinical diagnosis of depression. *Suicide and Life-Threatening Behavior, 23*(2), 139-145.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The hopelessness scale. *Journal of Consulting and Clinical Psychology, 42*, 861-865.

Beck, D. A., & Koenig, H. G. (1996). Minor depression: A review of the literature. *International Journal of Psychiatry in Medicine, 26*, 177-209.

Bembenutty, H. (2009). An interview with Herbert W. Marsh: A leading voice on self-concept, teaching effectiveness, and a force in quantitative analysis (Part 1). *Journal of Advanced Academics, 20*, 546-545. doi:10.1177/1932202X0902000307

Bergin, D. (1999). Influences on classroom interest. *Educational Psychologist, 34*, 87-98.

Berliner, D. C., & Calfee, R. C. (1996). *Handbook of educational psychology*. New York, NY: Macmillan.

Berlyne, D. E. (1949). "Interest" as a psychological concept. *British Journal of Psychology, 39*, 186-195.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity.* New York, NY: McGraw-Hill.

Berwick, R., & Ross, S. J. (1989) Motivation after matriculation: Are Japanese learners of English still alive after exam hell? *JALT Journal*, *11,* 193-210.

Blatt, S. J. (1995). The destructiveness of perfectionism: Implications for the treatment of depression. *American Psychologist, 50*, 1003-1020.

Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation*. San Diego, CA: Academic Press.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Dordrecht, Germany: Springer.

Bown, J., & White, C. J. (2010). Affect in a self-regulatory framework for language learning. *System, 38*, 432-443. doi:10.1016/j.system.2010.03.016

Bracken, B. A. (1996). Clinical applications of a multidimensional, context-dependent model of self-concept. In B. A. Bracken (Ed.), *Handbook of self concept: Developmental, social, and clinical considerations* (pp. 463-505). New York, NY: Wiley.

Bracken, B. A. (2009). Positive self-concept. In R. Gilman, E. S. Huebner, & M. J. Furlong (Eds.). *Handbook of positive psychology in schools* (pp. 89-106). New York, NY: Routledge.

Bracken, B. A., & Lamprecht, M. S. (2003). Self-concept: An equal opportunity construct. *School Psychology Quarterly*, *18*, 103-121. doi:10.1521/scpq.18.2.103.21859

Brennan, R. L. (2001). *Generalizability theory.* New York, NY: Springer-Verlag.

Brennen, R. L., & Kane, M. T. (1977). Signal/noise ratios for domain-referenced tests. *Psychometrika, 42*, 609-625.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*, 185–216.

Brophy, J. (2004). *Motivating students to learn* (2nd ed.). Mahwah, NJ: Erlbaum.

Brophy, J. (2010). *Motivating students to learn* (3rd ed.). New York, NY: Routledge.

Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston, MA: Heinle & Heinle.

Brown, J. D. (1995) Differences between norm-referenced and criterion-referenced tests. In J. D. Brown & S. O. Yamashita, (Eds.), *Language testing in Japan* (pp. 12-19). Tokyo, Japan: Japan Association for Language Teaching.

Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment* (New ed.). New York, NY: McGraw-Hill.

Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge, England: Cambridge University Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (136–162). Newbury Park, CA: Sage.

Burrows, L. (2009). Assessing speaking self-efficacy: Constructing a scale. *Kinki University English Journal, 4*, 1-12.

Butler, R. (2007). Teachers' achievement goals and associations with teachers' help-seeking: Examination of a novel approach to teacher motivation. *Journal of Educational Psychology, 99*, 241-252. doi:10.1037/0022-0663.99.2.241

Butler, R., & Shibaz, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction, 18*, 453-467. doi:10.1016/j.learninstruc.2008.06.004

Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation*. Washington, DC: American Psychological Association.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin, 92*, 111-135.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press.

Chang, E. C. (1998). Hope, problem-solving ability, and coping in a college student population: Some implications for theory and practice. *Journal of Clinical Psychology, 54*, 953-962.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. doi:10.1207/S15328007SEM0902_5

Chulef, A. S., Read, S. J., & Walsh, D. A. (2001). A hierarchical taxonomy of human goals. *Motivation and Emotion, 25*, 191-232. doi:10.1023/A:1012225223418

Ciarrochi, J., Heaven, P. C. L., & Davies, F. (2007). The impact of hope, self-esteem, and attributional style on adolescents' school grades and emotional well-being: A longitudinal study. *Journal of Research in Personality, 41*, 1161-1178. doi:10.1016/j.jrp.2007.02.001

Cooper, H., Okamura, L., & Gurka, V. (1992). Social activity and subjective well-being. *Personality and Individual Differences, 13*, 573-583.

Cornette, M. M., Strauman, T. J., Abramson, L. Y., & Busch, A. M. (2009). Self-discrepancy and suicidal ideation, *Cognition and Emotion, 23*, 504-527. doi: 10.1080/02699930802012005

Coryn, C. L. S., Spybrook, J. K., Evergreen, S. D. H., & Blinkiewicz, M. (2009). Development and evaluation of the social-emotional learning scale. *Journal of Psychoeducational Assessment, 27*, 283-295. doi:10.1177/0734282908328619

Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. New York, NY: Cambridge University Press.

Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology, 51,* 171-200. doi:10.1146/annurev.psych.51.1.171

Covington, M. V., & Dray, E. (2002). The developmental course of achievement motivation: A need-based approach, In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 33-56). San Diego, CA: Academic Press.

Craven, R. G., & Marsh, H. W. (2008). The centrality of self-concept for psychological well-being and unlocking human potential: Implications for child and educational psychologists. *Educational and Child Psychology, 25*, 104-118.

Crocker, J., & Park, L. E. (2003). Seeking self-esteem: Maintenance, enhancement, and protection of self-worth. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 291-313). New York, NY: Guilford.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin, 130,* 392-414. doi:10.1037/0033-2909.130.3.392

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108,* 593-623. doi:10.1037/0033-295X.108.3.593

Cronbach, L. J., & Gleser, G. C. (1964). The signal/noise ratio in the comparison of reliability coefficients. *Educational and Psychological Measurement, 3*, 467-480.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Crookes, G., & Schmidt, R. (1991). Motivation: Reopening the research agenda. *Language Learning, 41*, 469-512.

Csikszentmihalyi, M. (1990). *Flow*. New York, NY: Harper & Row.

Csikszentmihalyi, M. (1993). *The evolving self*. New York, NY: HarperCollins.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (Eds.). (2006). *A life worth living: Contributions to positive psychology*. New York, NY: Oxford University Press.

Csikszentmihalyi, M., & Nakamura, J. (2011). Positive psychology: Where did It come from, where is it going? In M. K. Sheldon, T. B. Kashdan, & M. F. Steger (Eds.), *Designing positive psychology: Taking stock and moving forward*. (pp. 3-8). New York, NY: Oxford University Press.

Csizer, K., & Kormos, J. (2008). Modeling the role of inter-cultural contact in the motivation of learning English as a foreign language. *Applied Linguistics, 30*, 166-185. doi:10.1093/applin/amn025

Csizer, K., & Kormos, J. (2009). Attitudes, selves and motivated behavior: A comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 98-119). Clevedon, England: Multilingual Matters.

Curry, L. A., Snyder, C. R., Cook, D. L., Ruby, B. C., & Rehm, M. (1997). Role of hope in academic and sport achievement. *Journal of Personality and Social Psychology, 73*(6), 1257-1267.

Custers, R., & Aarts, H. (2007). Goal-discrepant situations prime goal-directed actions if goals are temporarily or chronically accessible. *Personality and Social Psychology Bulletin, 33*, 623-633. doi:10.1177/0146167206297404

Dai, D. Y., & Sternberg, R. J. (Eds.). (2004). *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development*. Mahwah, NJ: Erlbaum.

Daniel, M. H. (1999). Behind the scenes: Using new measurement methods on the DAS and KAIT. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every psychologist and educator should know*. (pp. 37-63). Mahwah, NJ: Erlbaum.

Da Silva, D. (2007). Self concept: What is it, and why is it important. In K. Bradford-Watts (Ed.), *JALT 2006 Conference Proceedings*. Tokyo: JALT.

Da Silva, D., & McInerney, D. M. (2008). Motivational and self goals of female students in contemporary Japan. In O. S. Tan, D. M., McInerney, A. D. Liem & A-G. Tan (Eds.), *What the West can learn from the East: Asian perspectives on the psychology of learning and motivation* (pp. 191-216). Charlotte, NC: Information Age.

Davey, T., & Pitoniak, M. J. (2006). Designing computerized adaptive tests. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 543-573). Mahwah, NJ: Erlbaum.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York, NY: Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227-268. doi:10.1207/S15327965PLI1104_01

Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester.

DeKeyser, R. M. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124*, 197-229.

Dewaele, J.-M. (2002). Psychological and sociodemographic correlates of communicative anxiety in L2 and L3 production. *The International Journal of Bilingualism, 6*, 23-39. doi:10.1177/13670069020060010201

Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching, 4*(2), 237-274. doi:10.14746/ssllt.2014.4.2.5

Dewey, J. (1895). *Interest as related to will*. Chicago, IL: University of Chicago Press.

Dewey, J. (1913). *Interest and effort in education*. Boston, MA: Houghton Mifflin.

Dewey, J. (1916). *Democracy and education*. New York, NY: Macmillan.

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research, 31,* 103-157.

Diener, E. (2009). Satisfaction with life scales. Retrieved from http://www.psych.uiuc.edu/~ediener/SWLS.html

Diener, E. (2013). The remarkable changes in the science of subjective well-being. *Perspectives on Psychological Science, 8*, 663-666. doi:10.1177/1745691613507583

Diener, E., & Biswas-Diener, R. (2008). *Happiness: Unlocking the mysteries of psychological wealth*. Malden, MA: Blackwell.

Diener, E., & Diener, C. (1996). Most people are happy. *Psychological Science, 7*, 181-185.

Diener, E., & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology, 47*, 1105-1117.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49,* 71-75.

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and validity of life satisfaction scales. *Social Indicators Research, 112*, 497-527. doi:10.1007/s11205-012-0076-y

Diener, E., Kanazawa, S., Suh, E. M., & Oishi, S. (2014). Why people are in a generally good mood. *Personality and Social Psychology Review*, 1-22. doi:10.1177/1088868314544467

Diener, E., & Seligman, M. E. P. (2002). Very happy people. *Psychological Science, 13*, 81-84. doi:10.1111/1467-9280.00415

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276-302.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener (2010). New well-being measures: Short scales to assess flourishing and positive negative feelings. *Social Indicators Research*, *97*, 143-156. doi:10.1007/s11205-009-9493-y

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling, 13*(3), 440-464. doi:10.1207/s15328007sem1303_6

Doran, H. C. (2005). The information function for the one-parameter logistic model: Is it reliability? *Educational and Psychological Measurement, 65*(5), 665-675. doi:10.1177/0013164404272500

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge, England: Cambridge University Press.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Erlbaum.

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). Clevedon, England: Multilingual Matters.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford, England: Oxford University Press.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589-630). Malden, MA: Blackwell.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.) New York, NY: Routledge.

Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*. Clevedon, England: Multilingual Matters.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Harlow, England: Longman.

Duckworth, A. L., & Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science, 23*, 319-325. doi:10.1177/0963721414541462

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*, 1087-1101. doi:10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment, 91*, 166-174. doi:10.1080/00223890802634290

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescence. *Psychological Science, 16*, 939-944. doi:10.1111/j.1467-9280.2005.01641.x

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*(1), 198-208. doi:10.1037/0022-0663.98.1.198

Dweck, C. S. (1992). The study of goals in psychology. *Psychological Science, 3*, 165-167.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development.* Philadelphia, PA: Taylor and Francis.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.

Dweck, C. S., Higgins, E. T., & Grant, H. (2003). Self-systems give unique meaning to self variables. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 239-252). New York, NY: Guilford.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.

Dweck, C. S., & Wortman, C. B. (1982). Learned helplessness, anxiety, and achievement motivation: Neglected parallels in cognitive, affective, and coping response. In H. W. Krohne, & L. Laux (Eds.), *Achievement, stress, and anxiety*. New York, NY: Academic Press.

Eccles, J. S. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75-146). San Francisco, CA: Freeman.

Eccles, J. S. (2006). A motivational perspective on school achievement: Taking responsibility for learning, teaching, and supporting. In R. J. Sternberg & R. F. Subotnik (Eds.), *Optimizing student success in school with the other three Rs: Reasoning, resilience, and responsibility* (pp. 199-224). Greenwich, CT: Information Age.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109-132. doi:10.1146/annurev.psych.53.100901.135153

Egbert, J. (2003). A study of flow theory in the foreign language classroom. *The Modern Language Journal, 87*(4), 499-518. doi:10.1111/1540-4781.00204

Eid, M., & Larsen, R. J. (Eds.). (2008). *The science of subjective well-being*. New York, NY: Guilford.

Eidswick, J. (2010). Interest and prior knowledge in second language reading comprehension. *JALT Journal, 32*(2), 149-168.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169-200.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169-189.

Elliot, A. J. (2005). A conceptual history of the achievement goal construct. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 52-72). New York, NY: Guilford.

Elliot, A. J. (Ed.). (2008). *Handbook of approach and avoidance motivation*. New York, NY: Psychology Press.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*, 5-12.

Elliot, A. J., & Dweck, C. S. (Eds.). (2005). *Handbook of competence and motivation*. New York, NY: Guilford.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80,* 501-519. doi:10.1037/0022-3514.80.3.501

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology, 100*, 613-628. doi:10.1037/0022-0663.100.3.613

Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 525-551). Malden, MA: Blackwell.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Emmons, R. A. (1986). Personal strivings: An approach to personality and subjective well-being. *Journal of Personality and Social Psychology, 51*, 1058-1068.

Emmons, R. A. (1999). *The psychology of ultimate concerns: Motivation and spirituality in personality*. New York, NY: Guilford Press.

Engelhard, G. (2013). *Invariant measurement*. New York, NY: Routledge.

Ericsson, D. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725-747.

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. New York, NY: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Feldman, D. B., & Snyder. C. R. (2005). Hope and the meaningful life: Theoretical and empirical associations between goal-directed thinking and life meaning. *Journal of Social and Clinical Psychology, 24*, 401-421. doi:10.1521/jscp.24.3.401.65616

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Flett, G. L., Hewitt, P. L., Blankstein, K. R., & Gray, L. (1998). Psychological distress and the frequency of perfectionistic thinking. *Journal of Personality and Social Psychology, 75*, 1363-1381.

Ford, M. E. (1992). *Motivating humans: Goals, emotions, and personal agency beliefs*. Newbury Park, CA: Sage.

Ford, M. E., & Ford, D. H. (Eds.). (1987). *Humans as self-constructing systems: Putting the framework to work*. Hillsdale, NJ: Erlbaum.

Ford, M. E., & Nichols, C. W. (1991). Using goal assessments to identify motivational patterns and facilitate behavioral regulation. In M. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement, Vol. 7: Goals and self-regulatory processes* (pp. 57-84). Greenwich, CT: JAI Press.

Ford, M. E., & Smith, P. R. (2007). Thriving with social purpose: An integrative approach to the development of optimal human functioning, *Educational Psychologist, 42*, 153-171. doi:10.1080/00461520701416280

Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology, 2*, 300-319.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218-226. doi:10.1O37//0OO3-O66X.56.3.218

Fredrickson, B. (2009). *Positivity: Groundbreaking research reveals how to embrace the hidden strength of positive emotions, overcome negativity, and thrive*. New York, NY: Crown.

Friedman, T. L. (2007). *The world is flat: A brief history of the twenty-first century*. (3rd ed.). New York, NY: Picador.

Frisch, M. B., Clark, M. P., Rouse, S. V., Rudd, M. D., Paweleck, J. K., Greenstone, A., & Kopplin, D. A. (2005). Predictive and treatment validity of life satisfaction and the quality of life inventory. *Assessment, 12*(1), 66-78. doi:10.1177/1073191104268006

Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research, 14*, 449-468.

Fujimura-Fanselow, K. (1995). College women today: Options and dilemmas. In K. Fujimura-Fanselow & A. Kameda (Eds.), *Japanese women: New feminist perspectives on the past, present, and future* (pp. 125-154). New York, NY: The Feminist Press**.**

Furlong, A. (2008) The Japanese hikikomori phenomenon: Acute social withdrawal among young people, *The Sociological Review, 56*, 309–325. doi:10.1111/j.1467-954X.2008.00790.x

Gable, S. L., & Haidt, J. (2005). What (and why) is positive psychology. *Review of General Psychology, 9*(2), 103-110. doi: 10.1037/1089-2680.9.2.103

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London, England: Edward Arnold.

Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 1-20). Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawaii Press.

Gardner, R. C. (2010). *Motivation and second language acquisition: The socio-educational model*. New York, NY: Peter Lang.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House.

Gilbert, D. (2006). *Stumbling on happiness*. New York, NY: Knopf.

Gilman, R., Dooley, J., & Florell, D. (2006). Relative levels of hope and their relationship with academic and psychological indicators among adolescents. *Journal of Social and Clinical Psychology, 25*, 166-178. doi:10.1521/jscp.2006.25.2.166

Gilman, R., Furlong, M. J., & Huebner, E. S. (Eds.). (2009). *Handbook of positive psychology in schools*. New York, NY: Routledge.

Gilman, R., & Huebner, E. S. (2006). Characteristics of adolescents who report very high life satisfaction. *Journal of Youth and Adolescence, 35*, 311-319. doi: 10.1007/s10964-006-9036-7

Gladwell, M. (2008). *Outliers: The story of success*. New York, NY: Little, Brown.

Gollwitzer, P. M. (1986). Striving for specific identities: The social reality of self-symbolizing. In R. F. Baumeister (Ed.), *Public self and private self* (pp. 143-159). New York, NY: Springer-Verlag.

Gollwitzer, P. M. (1987). The implementation of identity intentions: A motivational-volitional perspective on symbolic self-completion. In F. Halisch & J. Kuhl (Eds.), *Motivation, intention and volition* (pp. 349-369). Heidelberg, Germany: Springer-Verlag.

Gollwitzer, P. M., Fujita, K., & Oettinggen, G. (2004). Planning and the implementation of goals. In R. Baumeister & K. Vohs (Eds.), *Handbook of self-regulation research* (pp. 211-228). New York, NY: Guilford.

Gollwitzer, P. M., & Kirchhof, O. (1998). The willful pursuit of identity. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 389-423). New York: Cambridge University Press.

Gollwitzer, P. M., & Wicklund, R. A. (1985). Self-symbolizing and the neglect of others' perspectives. *Journal of Personality and Social Psychology, 48*, 702-715.

Goodman, R. (2003). The why, what and how of educational reform in Japan. In R. Goodman & D. Phillips (Eds.). *Can the Japanese change their education system?*, (pp. 7-30). Wallingford, CT: Symposium Books.

Graham, S. (2006). Listening comprehension: The learners' perspective. *System, 34*, 165–182. doi:10.1016/j.system.2005.11.001

Graham, S. (2007). Learner strategies and self-efficacy: Making the connection. *Language Learning Journal, 35*, 81-93. doi:10.1080/09571730701315832

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning, 58*(4), 747-783. doi:10.1111/j.1467-9922.2008.00478.x

Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 63-84). New York, NY: Macmillan.

Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *23*, 443-451.

Gregersen, T., & MacIntyre, P. D. (2014). *Capitalizing on language learners' individuality: From premise to practice*. Bristol, UK: Multilingual Matters.

Gustafsson, J., & Aberg-Bengtsson, L. (2010). Unidimensionality and interpretability of psychological instruments. In S. E. Embretson (Ed.), *Measuring psychological constructs: Advances in model-based approaches* (pp. 97-121). Washington, DC: APA.

Hamachek, D. E. (1978). Psychodynamics of normal and neurotic perfectionism. *Psychology, 15*, 27-33.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research, 52,* 123-142.

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, E. A., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest and performance. *Journal of Educational Psychology, 100*, 105-122. doi:10.1037/0022-0663.100.1.105

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology, 17,* 300-312.

Harter, S. (1999). *The construction of the self*. New York, NY: Guilford.

Harter, S. (2012). *The construction of the self: Developmental and sociocultural foundations* (2nd ed.). New York, NY: Guilford.

Hattie, J. A. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Oxford, England: Routledge.

Heckhausen, J. (1999). *Developmental regulation across adulthood: Age-normative and sociostructural constraints as adaptive challenges*. New York, NY: Cambridge University Press.

Heckhausen, J., & Heckhausen, H. (Eds.). (2008). *Motivation and action*. New York, NY: Cambridge University Press.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145-149.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411-482. doi:10.3386/w12006

Helgesen, M. (2005). Can we teach our students to be happy? ELT News Think Tank. Retrieved May 5, 2010, from http://www.eltnews.com/discussions/thinktank/2005/06/what_role_do_the_senses_play_i.html#01

Helgesen, M. (2006). ELT and the "science of happiness". *The Language Teacher, 30*, 28-30.

Henning, G. (1989). Meanings and implications of the principle of local independence. *Language Testing, 6*, 95-108.

Henning, G. (1992). Dimensionality and construct validity of language tests. *Language Testing, 9*, 1-11.

Hewitt, P. L., & Flett, G. L. (1991). Perfectionism in the self and social contexts: Conceptualizations, assessment, and association with psychopathology. *Journal of Personality and Social Psychology, 60*, 456-470.

Hidaka, Y., Operario, D., Takenaka, M., Omori, S., Ichikawa, S., & Shirasaka, T. (2008). Attempted suicide and associated risk factors among youth in urban Japan. *Social Psychiatry and Psychiatric Epidemiology, 43*, 752-757. doi:10.1007/s00127-008-0352-y

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60,* 549-571.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41,* 111-127. doi:10.1207/s15326985ep4102_4

Higgins, E. T. (1987). Self discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319-340.

Higgins, E. T. (1996). The 'self-digest': Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology, 71*, 1062-1083.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology, 30*, 1-46.

Higgins, E. T. (2012). *Beyond pleasure and pain: How motivation works*. New York, NY: Oxford University Press.

Higgins, E. T., Klein, R., & Strauman, T. (1985). Self-concept discrepancy theory: A psychological model for distinguishing among different aspects of depression and anxiety. *Social Cognition, 3*, 51-76.

Hitokoto, H., & Uchida, Y. (2015). Interdependent happiness: Theoretical importance and measurement validity. *Journal of Happiness Studies, 16*, 211-239. doi:10.1007/s10902-014-9505-8

Horan, P. M., DiStefano, C., & Motl, R. W. (2003).Wording effects in self esteem scales: Methodological artifact or response style? *Structural Equation Modeling, 10,* 444-455. doi:10.1207/S15328007SEM1003_6

Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*, 112-126. doi:10.1017/S0267190501000071

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching, 43*(2), 154-167. doi:10.1017/S026144480999036X

Howell, A. J. (2009). Flourishing: Achievement-related correlates of students' well-being. *The Journal of Positive Psychology, 4*, 1-13. doi:10.1080/17439760802043459

Howell, R., Kern, M. L., & Lyubomirsky, S. (2007). Health benefits: Meta-analytically determining the impact of well-being on objective health outcomes. *Health Psychology Review, 1*, 83-136. doi:10.1080/17437190701492486

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1-55.

Huebner, E. S. (1994). Preliminary development and validation of a multidimensional life satisfaction scale for children. *Psychological Assessment, 6*, 149-158.

Huebner, E. S., Gilman, R., & Laughlin, J. E. (1999). A multimethod investigation of the multidimensionality of children's well-being reports: Discriminant validity of life satisfaction and self-esteem. *Social Indicators Research*, *46*, 1-22.

Huebner, E. S., Laughlin, J. E., Ash, C., & Gilman, R. (1998). Further validation of the multidimensional students' life satisfaction scale. *Journal of Psychoeducational Assessment, 16*, 118-134.

Humphreys, L. G. (1962). The organization of human abilities. *American Psychologist, 17,* 475-483.

Hunter, J. P., & Csikszentmihalyi, M. (2003). The positive psychology of interested adolescents. *Journal of Youth and Adolescence, 32*, 27-35. doi:10.1023/A:1021028306392

Huppert, F. A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being, 1*, 137-164. doi:10.1111/j.1758-0854.2009.01008.x

Huynh, H. (1994) On equivalence between a partial credit item and a set of independent Rasch binary items. *Psychometrika, 59*, 111-119.

Huynh, H. (1996) Decomposition of a Rasch partial credit item into independent binary and indecomposable trinary items. *Psychometrika, 61*(1), 31-39.

Imai, Y. (2010). Emotions in SLA: New insights from collaborative learning for an EFL classroom. *Modern Language Journal, 94*(2), 278-292. doi:10.1111/j.1540-4781.2010.01021.x

Jiao, H., Wang, S., & Kamata, A. (2007). Modeling local item dependence with the hierarchical generalized linear model. In E. V. Smith & R. M. Smith (Eds.), *Rasch measurement: Advanced and specialized applications* (pp. 390-404). Maple Grove, MN: JAM Press.

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology, 87*, 797-807. doi:10.1037//0021-9010.87.4.797

Jung, C. G. (1933). *Modern man in search of a soul*. New York, NY: Harcourt Brace.

Jung, C. G. (1957). *The undiscovered self*. New York, NY: Mentor.

Jung, C. G. (1964). *Man and his symbols*. New York, NY: Dell.

Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: The foundations of hedonic psychology*. New York, NY: Sage.

Kaneko, S. (2006). Japan's socially withdrawn youths and time constraints in Japanese society: Management and conceptualization of time in a support group for "hikikomori". *Time and Society, 15*, 233-249. doi:10.1177/0961463X06067034

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-Y., & Camerer, C. F. (2009). The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science, 20,* 963–973. doi:10.1111/j.1467-9280.2009.02402.x

Kaplan, A., & Maehr, M. L. (1999). Achievement goals and student well-being. *Contemporary Educational Psychology, 24*, 330-358.

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19,* 141-184. doi:10.1007/s10648-006-9012-5

Kaplan, A., Middleton, M. J., Urdan, T., & Midgley, C. (2002). Achievement goals and goal structures. In C. Midgley (Ed.), *Goals, goal structures and patterns of adaptive learning* (pp. 21-53). Mahwah, NJ: Erlbaum.

Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Newbury Park, CA: Sage.

Kashdan, T. B. (2004). Curiosity. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification* (pp. 125-141). New York, NY: Oxford University Press.

Kashdan, T. B. (2009). *Curious? Discover the missing ingredient to a fulfilling life*. New York, NY: William Morrow.

Kashdan, T., & Biswas-Diener, R. (2014). *The upside of your dark side*. New York, NY: Hudson Street Press.

Kashdan, T., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and initial psychometrics. *Journal of Research in Personality, 43*, 987-998. doi:10.1016/j.jrp.2009.04.011

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment, 82*, 291-305. doi:10.1207/s15327752jpa8203_05

Kashdan, T. B., & Silvia, P. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (2nd ed.) (pp. 367-375). New York, NY: Oxford University Press.

Kashdan, T. B., & Yuen, M. (2007). Whether highly curious students thrive academically depends on perceptions about the school learning environment: A study of Hong Kong students. *Motivation and Emotion, 31*, 260–270. doi:10.1007/s11031-007-9074-9.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), 486-507. doi:10.1177/0049124114543236

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*, 333-3511. doi:10.1207/S15328007SEM1003_1

Keyes, C. L. (1998). Social well-being. *Social Psychology Quarterly, 61*(2), 121-140.

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*, 207-222. doi:10.2307/3090197

Keyes, C. L. M. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology, 73*, 539-548. doi:0022-006X.73.3.539

Keyes, C. L. M. (2007). Promoting and protecting mental health as flourishing. *American Psychologist, 62*, 95-108. doi:10.1037/0003-066X.62.2.95

Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology, 82*, 1007-1022. doi:10.1037/0022-3514.82.6.1007

Kikuchi, K. (2013). Demotivators in the Japanese EFL context. In M. T. Apple, D. Da Silva & T. Fellner (Eds.). *Language learning motivation in Japan* (pp. 206-224). Bristol, UK: Multilingual Matters.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford.

Klinger, E. (1975). Consequences of commitment to and disengagement from incentives. *Psychological Review*, *82*, 1-25.

Kok, B. E., Catalino, L. I. & Fredrickson, B. L. (2008). The broadening, building, buffering effects of positive emotions. In S. J. Lopez (Ed.), *Positive psychology: Exploring the best of people, Vol. 3: Capitalizing on emotional experiences* (pp. 1-19). Westport, CT: Greenwood.

Krapp, A. (2002). An educational-psychological theory of interest and its relation to SDT. In E. L. Deci & R. M. Ryan (Eds.), *The handbook of self-determination research* (pp. 405-427). Rochester, NY: University of Rochester Press.

Krosnick, J. A. (1999). Survey methodology. *Annual Review of Psychology, 50*, 537-567.

Krosnick, J. A., & Fabrigar, L. R. (1997). *Designing rating scaling for effective measurement in surveys. Survey measurement and process quality*. New York, NY: Wiley.

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In J. D. Wright & P. V. Marsden (Eds.), *Handbook of survey research* (2nd ed., pp. 263-313). West Yorkshire: Emerald.

Kunter, M., Baumert, J., & Koller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*, 494-509. doi:10.1016/j.learninstruc.2007.09.002

Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology, 95*, 66-75. doi:10.1037/0022-3514.95.1.66

Lake, J. (2014). Curious readers and interesting reads: Developing a positive L2 reading self and motivation through extensive reading. *Journal of Extensive Reading, 2*, 13-27.

Latham, G. P. (2007). *Work motivation: History, theory, research, and practice*. Thousand Oaks, CA: Sage.

Lau, I. C. Y., Yeung, A. S., Jin, P., & Low, R. (1999). Toward a hierarchical, multidimensional English self-concept. *Journal of Educational Psychology, 91*, 747-755.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology, 58*, 317-344. doi:10.1146/annurev.psych.58.110405.085658

Leary, M. R., & Tangney, J. P. (2003). The self as an organizing construct in the behavioral and social science. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 3-14). New York, NY: Guilford.

Leung, J. and Leung, K. (1992). Life satisfaction, self-concept, and relationship with parents in adolescence. *Journal of Youth and Adolescence*, *21*, 653-665.

Lewin, K. (1935). *A dynamic theory of personality*. New York, NY: McGraw-Hill.

Lewis, A. D., Huebner, E. S., Reschly, A. L., & Valois, R. F. (2009). The incremental validity of positive emotions in predicting school functioning, *Journal of Psychoeducational Assessment, 27*, 397-408. doi:10.1177/0734282908330571

Lewis, A. D., Huebner, E. S., Malone, P. S., & Valois, R. F. (2011). Life satisfaction and student engagement in adolescents. *Journal of Youth and Adolescence*, 40, 249-262. doi:10.1007/s10964-010-9517-6

Linacre, J. M. (1994). Sample size and item calibration stability, *Rasch Measurement Transactions, 7*(4), 328.

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement, 3*(2), 193-122.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*, 85-106.

Linacre, J. M. (2006). Dichotomous equivalents to rating scales. *Rasch Measurement Transactions, 20*(1), 1052.

Linacre, J. M. (2011). *WINSTEPS Rasch measurement computer program*. [Computer software]. Chicago: Winsteps.com

Linacre, J. M. (2011). *A user's guide to WINSTEPS: Rasch-model computer program*. Chicago: Winsteps.com.

Linley, P. A. (2008). *Average to A+: Realising strengths in yourself and others*. Coventry: CAPP Press.

Little, B. R. (1983). Personal projects: A rationale and method for investigation. *Environment and Behavior, 15*, 273-309.

Little, B. R., Salmela-Aro, K., & Phillips, S. D. (Eds.). (2007). *Personal project pursuit: Goals, action, and human flourishing*. Mahwah, NJ: Erlbaum.

Little, T. D., & Slegers, D. W. (2005). Factor analysis: Multiple groups with means. In B. Everitt & D. Howell (Eds.) & D. Rindskopf (Section Ed.), *Encyclopedia of statistics in behavioral science* (pp. 617–623). Chichester, UK: Wiley.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational behavior and human performance, 3*(2), 157-189.

Locke, E. A. (2005). Setting goals for life and happiness. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 299-312). New York, NY: Oxford University Press.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*, 705-717. doi:10.1037/0003-066X.57.9.705

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science, 15*, 265-268. doi:10.1111/j.1467-8721.2006.00449.x

Locke, E. A., & Latham, G. P. (Eds.). (2013). *New developments in goal setting and task performance*. New York, NY: Routledge.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*, 635-694.

Loevinger, J. (1965). Person and population as psychometric concepts. *Psychological Review, 72*, 143-155.

Lopez, S. J. (2013). *Making hope happen: Create the future you want for yourself and others*. New York, NY: Atria.

Lopez, S. J., Pedrotti, J. T., & Snyder, C. R. (2014). *Positive psychology: The scientific and practical explorations of human strengths* (3rd ed.). Thousand Oaks, CA: Sage.

Lopez, S. J., & Snyder, C. R. (Eds.). (2003). *Positive psychological assessment: A handbook of models and measures*. Washington, DC: American Psychological Association.

Lopez, S. J., & Snyder, C. R. (Eds.). (2009). *Oxford handbook of positive psychology* (2nd ed.). New York, NY: Oxford University Press.

Lucas, R. E., Diener, E., Grob, A., Suh, E. M., & Shao, L. (2000). Cross-cultural evidence for the fundamental features of extraversion. *Journal of Personality and Social Psychology, 79*, 452-468. doi:10.1037/0022-3514.79.3.452

Lucas, R. E., Diener, E., & Suh, E. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology, 71,* 616-628.

Luo, G., & Andrich, D. (2005). Information functions for the general dichotomous unfolding model. In S. Alagumalai, D. Curtis, & N. Hungi (Eds.), *Applied Rasch measurement: A book of exemplars: Papers in honour of John P. Keeves* (pp. 309-328). Dordrecht, Netherlands: Springer.

Lyubomirsky, S. (2007). *The how of happiness: A scientific approach to getting the life you want*. New York, NY: Penguin.

Lyubomirsky, S. (2013). *The myths of happiness*. New York, NY: Penguin.

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*, 137-155.

Lyubomirsky, S., King, L. A., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131*, 803-855. doi:10.1037/0033-2909.131.6.803

MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and second language. *Language Learning, 41*, 513-534.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*, 283-305.

MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching 2*(2), 193-213.

MacLeod, A. K., & Conway, C. (2007). Well-being and the anticipation of future positive experiences: The role of income, social networks and planning ability. *Cognition and Emotion, 18*, 357-374. doi:10.1080/02699930441000247

Maddux, J. E. (2002). Self-efficacy: The power of believing you can. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 277-287). New York, NY: Oxford.

Maddux, J. E., & Gosselin, J. T. (2012). Self-efficacy. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (2nd ed.) (pp. 198-224). New York, NY: Guilford.

Maehr, M. L. (1984). Meaning and motivation: Toward a theory of personal investment. In R. Ames & C. Ames (Eds.), *Research on motivation in education* (pp. 115-144). New York, NY: Academic Press.

Maehr, M. L., & Midgley, C. (1991). Enhancing student motivation: A schoolwide approach. *Educational Psychologist, 26*, 399-427.

Maehr, M. L., & Midgley, C. (1996). *Transforming school cultures*. Boulder, CO: Westview.

Makarova, V., & Rodgers, T. (Eds.). (2004). *English language teaching: The case of Japan.* Munich: Lincom.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*, 954-969.

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, *82*, 623-636.

Marsh, H. W. (1990). The causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82,* 646-656.

Marsh, H. W. (1992). Content specificity of relations between academic achievement and academic self-concept. *Journal of Educational Psychology*, *84*, 35-42.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70,* 810-819.

Marsh, H. W. (2006). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, England: British Psychological Society.

Marsh, H. W. & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*(2), 133-163. doi:10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Craven, R. G., & McInerney, D. M. (Eds.). (2005). *New frontiers for self research*. Greenwich, CT: Information Age.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*, 59-77. doi:10.1348/000709910X503501

Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, *20*, 107-125.

Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research, 79*, 327-365. doi:10.3102/0034654308325583

Martin, R., & Barresi, J. (2006). *The rise and fall of soul and self: An intellectual history of personal identity*. New York, NY: Columbia University Press.

Marzano, R. J., & Heflebower, T. (2012). *Teaching and assessing 21st century skills*. Bloomington, IN: Marzano Research Laboratory.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370-396.

Maslow, A. (1954). *Motivation and personality*. New York, NY: Harper.

Maslow, A. H. (1955). Deficiency motivation and growth motivation. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation, 3*, 1-30. Lincoln, NB: University of Nebraska.

Maslow, A. (1968). *Toward a psychology of being*. New York, NY: Van Nostrand.

Maslow, A. (1970). *Religions, values, and peak experiences*. New York, NY: Viking.

Maslow, A. (1971). *The farther reaches of human nature*. New York, NY: Viking.

Massey, E. K., Gebhardt, W. A., & Garnefski, N. (2008). Adolescent goal content and pursuit: A review of the literature from the past 16 years. *Developmental Review, 28*, 421-460. doi:10.1016/j.dr.2008.03.002

Mathews, G. (1996). *What makes life worth living? How Japanese and Americans make sense of their worlds*. Berkeley, CA: University of California Press.

Mathews, G., & Izquierdo, C. (Eds.). (2009). *Pursuits of happiness: Well-being in anthropological perspective.* New York, NY: Berghahn.

May, R. (1975). *The courage to create*. New York, NY: Norton.

May, R. (Ed.). (1969). *Existential psychology*. New York, NY: Random House.

McClelland, D. C. (1987). *Human motivation*. Cambridge: Cambridge University Press.

McInerney, D. M., & Liem, A. D. (2008). Motivation theory and engaged learning. In P. A. Towndrow, C. Koh, & T. H. Soon (Eds.), *Motivation and practice for the classroom* (pp. 11-36). The Netherlands: Sense.

McInerney, D. M., & Van Etten, S. (Eds.). (2004). *Big theories revisited*. Greenwich, CN: Information Age.

McMahon, D. (2006). *Happiness: A history*. New York, NY: Atlantic Monthly.

McVeigh, B. J. (1997) *Life in a Japanese Women's College: Learning to be ladylike.* London: Routledge.

McVeigh, B. J. (2002). *Japanese higher education as myth*. Armonk, NY: M. E. Sharpe.

Melges, F. T., & Bowlby, J. (1969). Types of helplessness in psychopathological process. *Archives of General Psychiatry, 20*, 690-699.

Mercer, S. (2011a). *Towards an understanding of language learning self-concept*. Dordrecht, Germany: Springer.

Mercer, S. (2011b). Language learner self-concept: Complexity, continuity and change. *System, 39*(3), 335-346. doi:10.1016/j.system.2011.07.006

Mercer, S. (2012a). Self-concept: Situating the self. In S. Mercer, R. Ryan, & M. Williams (Eds.). *Psychology for language learning: Insights from research, theory and practice* (pp. 10-25). Basingstoke: Palgrave Macmillan.

Mercer, S. (2012b). The dynamic nature of a tertiary learner's foreign language self-concept. In M. Pawlak (Ed.), *New perspective on individual differences in language learning and teaching* (pp. 201-215). Dordrecht, Germany: Springer.

Mercer, S., & Ryan, S. (2010). A mindset for EFL: Learners' beliefs about the role of natural talent. *ELT Journal, 64*(4), 436-444. doi:10.1093/elt/ccp083

Mercer, S., Ryan, S., & Williams, M. (Eds.). (2012). *Psychology for language learning: Insights from research, theory and practice*. Basingstoke: Palgrave Macmillan.

Mercer, S., & Williams, M. (Eds.). (2014). *Multiple perspectives on the self in SLA*. Bristol: Multilingual Matters.

Meese, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual review of psychology, 57*, 487-504). doi:10.1146/annurev.psych.56.091103.070258

Midgley, C. (Ed.). (2002). *Goals, goal structures, and patterns of adaptive learning*. Mahwah, NJ: Erlbaum.

Midgley, C., & Urdan, T. (2001). Academic self-handicapping and performance goals: A further examination. *Contemporary Educational Psychology, 26*, 61-75. doi:10.1006/ceps.2000.1041

Miller, C. W., & Roloff, M. E. (2006). The perceived characteristics of irresolvable, resolvable and resolved intimate conflicts: Is there evidence of intractability? *International Journal of Conflict Management, 17*, 291-315. doi:10.1108/10444060610749464

Miller, C. W., Roloff, M. E., & Resnik, R. M. (2014). Hopelessness and interpersonal conflict: Antecedents and consequences of losing hope. *Western Journal of Communication, 78*, 563-585. doi:10.1080/10570314.2014.896026

Mills, N. A. (2009). A ‘Guide du Routard’ simulation: Enhancing the Standards through project-based learning. *Foreign Language Annals, 42*(4), 607-639. doi:10.1111/j.1944-9720.2009.01046.x

Mills, N. (2014). Self-efficacy in foreign language acquisition. In S. Mercer & M. Williams (Eds.), *Multiple Perspectives on the Self in SLA* (pp. 6-22). Bristol, UK: Multilingual Matters.

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*, 276-295. doi:10.1111/j.1944-9720.2006.tb02266.x

Minkoff, K., Bergman, E., Beck, A. T., & Beck, R. (1973). Hopelessness, depression and attempted suicide. *American Journal of Psychiatry, 130*, 455-459.

Mischel, W., & Morf, C. C. (2003). The self as a psycho-social dynamic processing system: A meta-perspective on a century of the self in psychology. In M. Leary & J. Tangney (Eds.) *Handbook of self and identity* (pp. 15-43). New York, NY: Guilford.

Moskowitz, G. B., & Grant, H. (Eds.). (2009). *The psychology of goals*. New York, NY: Guilford.

Motl, R. W., & DiStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equation Modeling, 9,* 562-578. doi:10.1207/S15328007SEM0904_6

Mruk, C. J. (2006). *Self-esteem research, theory, and practice: Toward a positive psychology of self-esteem* (3rd ed.). New York, NY: Springer.

Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivational terminology. *Contemporary Educational Psychology, 25*, 3-53. doi:10.1006/ceps.1999.1019

Murray, H. A. (1938). *Explorations in personality*. New York, NY: Oxford University Press.

Murray, G., Goa, X., & Lamb, T. (Eds.). (2011). *Identity, motivation and autonomy in language learning*. Bristol, UK: Multilingual Matters.

Nakao, M., & Takeuchi, T. (2006). The suicide epidemic in Japan and strategies of depression screening for its prevention. *Bulletin of the World Health Organization, 84*(6), 492-493.

Nakata, Y. (2006). *Motivation and experience in foreign language learning*. Bern: Peter Lang.

Nesse, R. M., & Ellsworth, P. C. (2009). Evolution, emotions, and emotional disorders. *American Psychologist, 64*, 129-139. doi:10.1037/a0013503

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91*, 328-346.

Noels, K. A. (2001). New orientations in language learning motivation: Towards a model of intrinsic, extrinsic, and integrative orientations and motivation. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 43-68). Honolulu: University of Hawai'i Press.

Noels, K. A. (2009). Identity and the internalization of language learning into the self-concept. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language, identity and the L2 self* (pp. 295-313). Clevedon: Multilingual Matters.

Norton, B. (2001). Non-participation, imagined communities and the language classroom. In M. P. Breen (Ed.), *Learner contributions to language learning: New directions in research* (pp. 159-171). Harlow, England: Longman.

Notterman, J. M. (2004). Persistent conceptual issues in psychology. *Theory & Psychology, 14*(2), 239-260. doi:10.1177/0959354304042019

Nugent, S. A. (2000). Perfectionism: Its manifestations and classroom-based interventions. *Journal of Secondary Gifted Education, 11*, 215-221. doi:10.4219/jsge-2000-630

Nurmi, J.-E. (1991). How do adolescents see their future? A review of the development of future orientation and planning. *Developmental Review, 11*, 1-59.

Nurmi, J.-E. (1992). Age differences in adult life goals, concerns, and their temporal extension: A life course approach to future-oriented motivation. *International Journal of Behavioral Development, 15*, 487-508.

Okano, K. (2009). *Young women in Japan: Transitions to adulthood*. New York, NY: Routledge.

Okano, K., & Tsuchiya, M. (1999). *Education in contemporary Japan: Inequality and diversity*. Cambridge: Cambridge University Press.

Ong, A., & Van Dulmen, M. H. M. (Eds.). (2007). *Oxford handbook of methods in positive psychology*. New York, NY: Oxford University Press.

Ostini, R., & Nering, M. (2006). *Polytomous item response theory models.* Thousand Oaks, CA: Sage.

Oxford, R. L. (1994). Where are we regarding language learning motivation? *The Modern Language Journal, 78*, 512-514.

Oxford, R. L., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *Modern Language Journal, 78*, 12-28.

Pacht, A. R. (1984). Reflections on perfectionism. *American Psychologist, 39*, 386-390.

Pajares, F. (2001). Toward a positive psychology of academic motivation. *Journal of Educational Research, 95*, 27-35. doi:10.1080/00220670109598780

Pajares, F., & Urdan, T. (Eds.). (2002). *Academic motivation of adolescents*. Greenwich, CT: Information Age.

Pajares, F., & Urdan, T. (Eds.). (2006). *Self-efficacy beliefs of adolescents*. Greenwich, CT: Information Age.

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment, 5*, 164-172.

Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *Journal of Positive Psychology, 3*, 137-152. doi:10.1080/17439760701756946

Perry, N. E., Turner, J. C., & Meyer, D. K. (2006). Classrooms as contexts for motivating learning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed.) (pp. 327-348). Mahwah, NJ: Erlbaum.

Peterson, C. (2006). *A Primer in positive psychology*. New York, NY: Oxford University Press.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook of classification.* New York, NY: Oxford University Press.

Pink, D. (2009). *Drive: The surprising truth about what motivates us*. New York, NY: Riverhead.

Pintrich, P. R. (2000a). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544-555. doi:10.1037/0022-0663.92.3.544

Pintrich, P. (2000b). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner, (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic Press.

Pintrich, P. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*, 667-686. doi: 10.1037/0022-0663.95.4.667

Pintrich, P. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385-407. doi:1040-726X/04/1200-0385/0

Pintrich, P. R., Conley, A. M., & Kempler, T. M. (2003). Current issues in achievement goal theory and research. *International Journal of Educational Research, 39,* 319-337. doi:10.1016/j.ijer.2004.06.002

Popham, W. J. (1978). *Criterion-referenced measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Popham, W. J. (1981). *Modern educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Pressman, S. D., & Cohen, S. (2005). Does positive affect influence health? *Psychological Bulletin, 131*, 925-971. doi:10.1037/0033-2909.131.6.925

Proctor, C. L., Linley, P. A., & Maltby, J. (2009). Youth life satisfaction: A review of the literature. *Journal of Happiness Studies, 10*, 583-630. doi:10.1007/s10902-008-9110-9

Pruitt, D. G., & Olczak, P. V. (1995). Beyond hope: Approaches to resolving seemingly intractable conflict. In B. B. Bunker & J. Z. Rubin (Eds.), *Conflict, cooperation, and justice: Essays inspired by the work of Morton Deutsch* (pp. 59-92). San Francisco, CA: Jossey-Bass.

Pullmann, H., & Allik, J. (2008). Relations of academic and general self-esteem to school achievement. *Personality and Individual Differences, 45*, 559-564. doi:10.1016/j.paid.2008.06.017

Rasch, G. (1960/1993). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: MESA.

Reeve, J. (2005). *Understanding motivation and emotion* (4th ed.). Hoboken, NJ: Wiley.

Reeve, J. (2009). *Understanding motivation and emotion* (5th ed.). Hoboken, NJ: Wiley.

Reschly, A. L., Huebner, E. S., Appleton, J. J., & Antaramian, S. (2008). Engagement as flourishing: The contribution of positive emotions and coping to adolescents' engagement at school and with learning. *Psychology in the Schools, 45*, 419-431. doi:10.1002/pits.20306

Ripley, A. (2013). *The smartest kids in the world: And how they got that way*. New York, NY: Simon & Schuster.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale, *Personality and Social Psychology Bulletin, 27*, 151-161. doi:10.1177/0146167201272002

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The reverse of social anxiety is not always the opposite: The reverse-scored items of the Social Interaction Anxiety Scale do not belong. *Behavior Therapy, 38*, 192-206. doi:10.1016/j.beth.2006.08.001

Rogers, C. (1951). *Client-centered therapy: Its current practice, implications and theory*. Boston, MA: Houghton Mifflin.

Rogers, C. (1961). *On becoming a person*. Boston, MA: Houghton Mifflin.

Rogers, C. (1980). *A way of being*. Boston, MA: Houghton Mifflin.

Romero, E., Villar, P., Luengo, M. A., & Gomez-Fraguela, J. A. (2009). Traits, personal strivings and well-being. *Journal of Research in Personality, 43*, 535-546. doi:10.1016/j.jrp.2009.03.006

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rosenberg, M. (1979). *Conceiving the self.* New York, NY: Basic Books.

Rosenberg, M. (1989). Self-concept research: A historical overview. *Social Forces, 68*(1), 34-44.

Rouault, G. (2007). Developing a criterion-referenced self-assessment. In K. Bradford-Watts (Ed.), *JALT 2006 Conference Proceedings*. Tokyo: JALT.

Rubio, F. (Ed.). (2007). *Self-esteem and foreign language learning*. Newcastle, UK: Cambridge Scholars Publishing.

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. doi:10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudemonic well-being. *Annual Review of Psychology, 52*, 141-166. doi:10.1146/annurev.psych.52.1.141

Ryan, S. (2005). Language learning motivation within the context of globalization: An L2 self within an imagined global community. *Critical Inquiry in Language Studies: An International Journal, 4*(2), 23-45. doi:10.1207/s15427595cils0301_2

Ryan, S. (2008). *The ideal L2 selves of Japanese learners of English*. Unpublished doctoral dissertation, University of Nottingham.

Ryan, S. (2009). Self and identity in L2 motivation in Japan: The ideal L2 self and Japanese learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 120-143). Bristol, UK: Multilingual Matters.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological wellbeing. *Journal of Personality and Social Psychology, 57*, 1069-1081.

Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*, 719-727.

Ryff, C. D., & Singer, B. (1998). The contours of positive human health. *Psychological Inquiry, 9*, 1-28.

Ryff, C. D., & Singer, B. (2000). Interpersonal flourishing: A positive health agenda for the new millennium. *Personality and Social Psychology Review, 4*(1), 30-44. doi:10.1207/S15327957PSPR0401_4

Ryff, C. D., & Singer, B. (2008). Know thyself and become what you are: A eudaimonic approach to psychological well-being. *Journal of Happiness Studies, 9*, 13-39. doi:10.1007/s10902-006-9019-0

Salmela-Aro, K., Aunola, K., & Nurmi, J-E. (2007). Personal goals during emerging adulthood: A 10-year follow up. *Journal of Adolescent Research, 22*, 690-715. doi:10.1177/0743558407303978

Sato, A., & Yasuda, A. (2001). Development of the Japanese version of positive and negative affect schedule (PANAS) scales [In Japanese]. *Japanese Journal of Personality, 9*, 138-139.

Sawyer, R. K. (Ed.). (2006). *The Cambridge handbook of the learning sciences*. Cambridge, England: Cambridge University.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion*. Oxford: Oxford University Press.

Schlegel, R. J., & Hicks, J. A. (2011). The true self and psychological health: Emerging evidence and future directions. *Social and Personality Psychology Compass, 5*, 989-1003. doi:10.1111/j.1751-9004.2011.00401.x

Schlegel, R. J, Hicks, J. A., Arndt, J., & King, L. A. (2009). Thine own self: True self-concept accessibility and meaning in life. *Journal of Personality and Social Psychology, 96*, 473-490. doi:10.1037/a0014060

Schmid, K. L., Phelps, E., Kiely, M. K., Napolitano, C. M., Boyd, M. J., & Learner, R. M. (2011). The role of adolescents' hopeful futures in predicting positive and negative developmental trajectories: Findings from the 4-H study of positive youth development. *The Journal of Positive Psychology, 6*, 45-56. doi:10.1080/17439760.2010.536777

Schmukle, S. C., Egloff, B., & Burns, L. R. (2002). The relationship between positive and negative affect in the Positive and Negative Affect Schedule. *Journal of Research in Personality, 36*, 463-475. doi:10.1016/S0092-6566(02)00007-7

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Routledge.

Schumann, J. H. (1997). *The neurobiology of affect in language learning*. Malden, MA: Blackwell.

Schumann, J. H. (2001). Learning as foraging. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 21-28). Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Schumann, J. H., & Wood, L. A. (2004). The neurobiology of motivation. In J. H. Schumann, S. E. Crowell, N. E. Jones, N. Lee, S. A. Schuchert & L. A. Wood (Eds.), *The neurobiology of learning.* (pp. 23-42). London: Erlbaum.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 26*, 207-231.

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

Schunk, D. H., & Zimmerman, B. J. (2006). Competence and control beliefs: Distinguishing the means and ends. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed.) (pp. 349-367). Mahwah, NJ: Erlbaum.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (2008). *Motivation and self-regulated learning: Theory, research, and applications*. New York, NY: Erlbaum.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco, CA: Freeman.

Seligman, M. E. P. (1999). President's address. *American Psychologist, 54*, 559-562.

Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York, NY: Free Press.

Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being.* New York, NY: Free Press.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55,* 5-14. doi:10.1037/0003-066X.55.1.5

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist, 60*, 410-421. doi:10.1037/0003-066X.60.5.410

Senko, C., Durik, A. M., & Harackiewicz, J. M. (2008). Historical perspectives and new directions in achievement goal theory: Understanding the effects of mastery and performance-approach goals. In J. Y Shah & W. Gardner (Eds.), *Handbook of motivational science* (pp. 100-113). New York, NY: Guilford.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46,* 407-441.

Shavelson, R. J., & Webb, N. M. (1991) *Generalizability theory: A primer*. Newbury Park, CA: Sage.

Sheldon, K. M. (2002). The self-concordance model of healthy goal-striving: When personal goals correctly represent the person. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 65-86). Rochester, NY: University of Rochester Press.

Sheldon, K. M. (2004). *Optimal human being: An integrated multi-level perspective*. Mahwah, NJ: Erlbaum.

Sheldon, K. M., & Elliott, A. J. (1999). Goal striving, need-satisfaction, and longitudinal well-being: The Self-concordance model. *Journal of Personality and Social Psychology, 76*, 482-497.

Sheldon, K. M., & Elliott, A. J. (2000). Personal goals in social roles: Divergences and convergences across roles and levels of analysis. *Journal of Personality, 68*, 51-84. doi:10.1111/1467-6494.00091

Sheldon, K. M., Ryan, R. M., Deci, E. L., & Kasser, T. (2004). The independent effects of goal contents and motives on well-being: It's both what you pursue *and* why you pursue it. *Personality and Social Psychology Bulletin, 30,* 475-486. doi:10.1177/0146167203261883

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

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*, 158-176. doi:10.1521/scpq.18.2.158.21860

Silvia, P. J. (2001). Interest and interests: The psychology of constructive capriciousness. *Review of General Psychology, 5*, 270-290. doi:10.1037/1089-2680.5.3.270

Silvia, P. J. (2003). Self-efficacy and interest: Experimental studies of optimal incompetence. *Journal of Vocational Behavior, 62*, 237-249. doi:10.1016/S0001-8791(02)00013-1

Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion, 5*, 89-102. doi:10.1037/1528-3542.5.1.89

Silvia, P. J. (2006). *Exploring the psychology of interest.* New York, NY: Oxford University Press.

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science, 17,* 57-60. doi:10.1111/j.1467-8721.2008.00548.x

Silvia, P. J., & Warburton, J. B. (2006). Positive and negative affect: Bridging states and traits. In J. C. Thomas & D. L. Segal (Eds.), *Comprehensive handbook of personality and psychopathology: Vol. 1. Personality and everyday functioning* (pp. 268-284). New York, NY: Wiley.

Slocum-Gori, S. L., & Zumbo, B. D. (2011). Assessing the unidimensionality of psychological scales: Using multiple criteria from factor analysis. *Social Indicators Research, 102*, 443-461. doi:10.1007/s11205-010-9682-8

Slocum-Gori, S. L., Zumbo, B. D., Michalos, A. C., & Diener, E. (2009). A note on the dimensionality of quality of life scales: An illustration with the Satisfaction with Life Scale (SWLS). *Social Indicators Research, 92*, 489-496. doi:10.1007/s11205-008-9303-y

Snyder, C. R. (1994). *The psychology of hope: You can get there from here*. New York, NY: Free Press.

Snyder, C. R. (2000a). *Handbook of hope*. San Diego, CA: Academic Press.

Snyder, C. R. (2000b). The past and future of hope. *Journal of Social and Clinical Psychology, 19*, 11-28. doi:10.1521/jscp.2000.19.1.11

Snyder, C. R. (2002). Hope theory: Rainbows in the mind. *Psychological Inquiry, 13*, 249-275. doi:10.1207/S15327965PLI1304_01

Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L, Gibb, J., Langelle, C., & Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60,* 570-585.

Snyder, C. R., & Lopez, S. J. (Eds.). (2002). *Handbook of positive psychology*. New York, NY: Oxford University Press.

Snyder, C. R., & Lopez, S. J. (2007). *Positive psychology: The scientific and practical explorations of human strengths*. Thousand Oaks, CA: Sage.

Snyder, C. R., Shorey, H. S., Cheavens, J., Pulvers, K., Adams, V. H., & Wiklund, C. (2002). Hope and academic success in college. *Journal of Educational Psychology, 94*(4), 820-826. doi:10.1037/0022-0663.94.4.820

Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher, 27*, 11-20.

Sternberg, R. J., & Grigorenko, E. L. (Eds.). (2003). *The psychology of abilities, competencies and expertise*. Cambridge, England: Cambridge University Press.

Sternberg, R. J., & Subotnik, R. F. (2006). *Optimizing student success in school with the other three Rs: Reasoning, resilience, and responsibility*. Greenwich, CT: Information Age.

Stipek, D. J. (2002). *Motivation to learn: Integrating theory and practice* (4th ed.). Boston, MA: Allyn and Bacon.

Strauman, T. J. (1989). Self-discrepancies in clinical depression and social phobia: Cognitive structures that underlie emotional disorders? *Journal of Abnormal Psychology, 98*, 5-14.

Strauman, T. J., & Higgins, E. T. (1988). Self-discrepancies as predictors of vulnerability to distinct syndromes of chronic emotional distress. *Journal of Personality, 56*, 685-707.

Sugimoto, Y. (2010). *An introduction to Japanese society* (3rd ed.). Cambridge: Cambridge University.

Suldo, S. M., & Huebner, E. S. (2006). Is extremely high life satisfaction during adolescence advantageous?. *Social Indicators Research*, *78*, 179-203. doi:10.1007/s11205-005-8208-2

Suldo, S. M., Shaffer, E. J., & Riley, K. N. (2008). A social-cognitive-behavioral model of academic predictors of adolescents' life satisfaction. *School Psychology Quarterly, 23*, 56-69. doi:10.1037/1045-3830.23.1.56

Suldo, S. M., & Shaffer, E. J. (2007). Evaluation of the self-efficacy questionnaire for children in two samples of American adolescents. *Journal of Psychoeducational Assessment, 24*, 341-355. doi:10.1177/0734282907300636

Swann, W. B., & Bosson, J. (2010). Self and identity. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed.) (pp. 589-628). New York, NY: McGraw-Hill.

Swann, W. B., Chang-Schneider, C., McClarty, K. L. (2007). Do people's self-views matter?: Self-concept and self-esteem in everyday life. *American Psychologist, 62*(2), 84-94. doi:10.1037/0003-066X.62.2.84

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Tafarodi, R. W., & Swann, W. B., Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment*, *65*, 322-342.

Taguchi, T., Magid, M., & Papi, M. (2009). The L2 motivational self system among Japanese, Chinese, and Iranian learners of English: A comparative study. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 66-97). Bristol: Multilingual Matters.

Takahira, M. (1998). Reliability and validity of an extended Japanese version of hopelessness scale (In Japanese). *The Japanese Journal of Personality, 7*, 1-10.

Teo, A. R. (2010). A new form of social withdrawal in Japan: A review of *hikikomori. International Journal of Social Psychiatry, 56*, 178-185. doi:10.1177/0020764008100629

Thissen, D., & Orlando, M. (2001). Item response theory for items scored in two categories. In D. Thissen & H. Wainer (Eds.), *Test scoring* (pp. 73-140). Mahwah, NJ: Erlbaum.

Thompson, B. (Ed.). (2003). *Score reliability: Contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage.

Thorndike, E. L. (1904). *An introduction to the theory of mental and social measurements.* New York, NY: Teachers College, Columbia University.

Tong, E. M. W., Fredrickson, B. L., Chang, W., & Lim, Z. X. (2010). Re-examining hope: The roles of agency thinking and pathways thinking. *Cognition and Emotion*, *24*(7), 1207-1215. doi:10.1080/02699930903138865

Tough, P. (2012). *How children succeed: Grit, curiosity, and the hidden power of character.* New York, NY: Houghton Mifflin Harcourt.

Tourangeau, R., Couper, M. P., & Conrad, F. (2007). Colors, labels, and interpretive heuristics for response scales. *Public Opinion Quarterly, 71*(1), 91-112. doi:10.1093/poq/nfl046

Twenge, J. (2006). *Generation me: Why today's young Americans are more confident, assertive, entitled—And more miserable than ever before*. New York, NY: Free Press.

Twenge, J. (2008). Generation me, the origins of birth cohort differences in personality traits, and cross-temporal meta-analysis. *Social and Personality Compass, 2*, 1440-1454. doi:10.1111/j.1751-9004.2008.00094.x

Twenge, J., & Campbell, W. K. (2009). *The narcissism epidemic: Living in the age of entitlement*. New York, NY: Free Press.

Urdan, T., Midgley, C., & Anderman, E. A. (1998). The role of classroom goal structure in students' use of self-handicapping. *American Educational Research Journal, 35*, 101-122.

Urdan, T., Ryan, A. M., Anderman, E. M., Gheen, M. H. (2002). Goals, goal structures, and avoidance behaviors. In C. Midgley (Ed.), *Goals, goal structures and patterns of adaptive learning* (pp. 55-83). Mahwah, NJ: Erlbaum.

Ushioda, E. (2008). Motivation and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 19-34). Cambridge, England: Cambridge University Press.

Ushioda, E. (2013). Foreign language motivation research in Japan: An insider perspective from outside Japan. In M. Apple, D. Da Silva & T. Fellner (Eds.). *Language Learning Motivation in Japan*, (pp. 1-14). Bristol: Multilingual Matters.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review. *Educational Psychologist, 39,* 111-133. doi:10.1207/s15326985ep3902_3

Vallerand, R. J. (2010). On passion for life activities: The dualistic model of passion. *Advances in Experimental Social Psychology, 42*, 97–193. doi:10.1016/S0065-2601(10)42003-1

Vallerand, R. J., Blanchard, C. M., Mageau, G. A., Koestner, R., Ratelle, C., Léonard, M., & Gagné (2003). Les passions de l'âme: On obsessive and harmonious passion. *Journal of Personality and Social Psychology*, *85*, 756-767. doi:10.1037/0022-3514.85.4.756

Vallerand, R. J., Salvy, S. J., Mageau, G. A., Elliot, A. J., Denis, P., Grouzet, F. M. E., & Blanchard, C. B. (2007). On the role of passion in performance. *Journal of Personality, 75*, 505-533. doi:10.1111/j.1467-6494.2007.00447.x

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations. *Organizational Research Methods 3*(1), 4-70, doi:10.1177/109442810031002

Vincent, P. J., Boddana, P., & MacLeod, A. K. (2004). Positive life goals and plans in parasuicide. *British Journal of Clinical Psychology, 11*, 90-99. doi:10.1002/cpp.394

von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*, 574-588. doi:10.1177/1745691611421204

Wainer, H., & Thissen, D. (1996). How is reliability related to the quality of test scores? What is the effect of local dependence on reliability? *Educational Measurement: Issues and Practice, 15*, 22-29.

Walker, C. J., & Symons, C. (1997). The meaning of human motivation. In J. L. Bess (Ed.), *Teaching well and liking it: Motivating faculty to teach effectively* (pp. 3-18). Baltimore, MD: John Hopkins University Press.

Walsh, R. (1993). The transpersonal movement: A history and state of the art. *Journal of Transpersonal Psychology, 25*, 123-140.

Walsh, R., & Vaughan, F. (Eds.). (1993). *Paths beyond ego: The transpersonal vision*. Los Angeles, CA: Tarcher.

Waterman, A. S. (Ed.). (2013a). *The best within us: Positive psychology perspectives on eudaimonia*. Washington, DC: American Psychological Association.

Waterman, A. S. (2013b). The humanistic psychology–positive psychology divide: Contrasts in philosophical foundations. *American Psychologist*, *68*(3), 124-133. doi:10.1037/a0032168

Watts, R. E., & Critelli, J. W. (1997). Roots of contemporary cognitive theories in the individual psychology of Alfred Adler. *Journal of Cognitive Psychotherapy, 11*, 147-156.

Watson, D. (2000). *Mood and temperament.* New York, NY: Guilford.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*, 346-353.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*, 820-838.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548-573.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York, NY: Springer-Verlag.

Weiner, B. (1990). History of motivational research in education. *Journal of Educational Psychology, 82*, 616-622.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: Bradford.

Wendler, C. L. W., & Walker, M. E. (2006). Practical issues in designing and maintaining multiple test forms for large-scale programs. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 445-467). Mahwah, NJ: Erlbaum.

Wentzel, K. R. (1989). Adolescent classroom goals, standards for performance, and academic achievement: An interactionist perspective. *Journal of Educational Psychology, 81,* 131-142.

Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development, 62*, 1066-1078.

Wentzel, K. R. (1993). Social and academic goals at school: Motivation and achievement in early adolescence. *Journal of Early Adolescence, 13*, 4-20.

Wentzel, K. R. (1999). Social-motivational processes and interpersonal relationships: Implications for understanding students' academic success. *Journal of Educational Psychology, 91*, 76-97.

Wentzel, K. R. (2000). What is it that I'm trying to achieve? Classroom goals from a content perspective. *Contemporary Educational Psychology, 25*, 105-115. doi: 10.1006/ceps.1999.1021

Wentzel, K. R. (2005). Peer relationships, motivation, and academic performance at school. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 279-296). New York, NY: Guilford.

Wentzel, K. R., & Wigfield, A. (Eds.). (2009). *Handbook of motivation at school*. New York, NY: Routledge.

Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine*. Cambridge, MA: MIT Press.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*, 265-310.

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91-120). San Diego, CA: Academic Press.

Wilber, K. (2000). *Integral psychology: Consciousness, spirit, psychology, therapy*. Boston, MA: Shambhala.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Erlbaum.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.

Wilson, T. D. (2003). Knowing when to ask: Introspection and the adaptive unconscious. *Journal of Consciousness Studies, 10*, 131-140.

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science, 14*, 131-134. doi:10.1111/j.0963-7214.2005.00355.x

Woodrow, L. (2006). A model of adaptive language learning. *The Modern Language Journal, 90*(3), 297-319. doi:10.1111/j.1540-4781.2006.00424.x

Woodrow, L. (2008). *A model of adaptive language learning: The case of EAP students.* Saarbrucken, Germany: Verlag Dr Muller.

Woodrow, L. (2012). Goal orientations: Three perspectives on motivation goal orientations. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for language learning* (pp. 188-202). Basingstoke, England: Palgrave Macmillan.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement, 14*(2), 97-116.

Wright, B. D. (1996). Reliability and separation. *Rasch Measurement Transactions, 9*(4), 472.

Wright, B. D., & Masters, J. (1982). *Rating scale analysis*. Chicago, IL: Mesa.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: Mesa.

Wright, T. A. (2005). The role of “happiness” in organizational research: Past, present and future directions. In P. L. Perrewe & D. C. Ganster (Eds.), *Research in occupational stress and well-being* (pp. 221-264). Amsterdam: JAI Press.

Wright, T. A. (2010). Much more than meets the eye: The role of psychological well-being in job performance, employee retention and cardiovascular health. *Organizational Dynamics, 39*, 13-23. doi:10.1016/j.orgdyn.2009.10.009

Yalom, I. D. (1980). *Existential psychotherapy*. New York, NY: Basic.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 5,* 245-262.

Yen, W. M. (1993). Scaling performance assessments: Strategies for managing LID. *Journal of Educational Measurement, 30*, 187-213.

Yeung, A. S., Shui, H. -S., Lau, I. C.–Y., McInerney, D. M., Russell-Bowie, D., & Suliman, R. (2000). Where is the hierarchy of academic self-concept? *Journal of Educational Psychology, 92*, 556-567. doi:10.1037/0022-0663.92.3.556

Yeung, A. S., & Wong, E. K. P. (2004). Domain-specificity of trilingual teachers' verbal self-concepts. *Journal of Educational Psychology, 96*, 360-368. doi:10.1037/0022-0663.96.2.360

Young, M. A., Halper, I. S., Clark, D. C., Scheftner, W., & Fawcett, J. (1992). An item-response theory evaluation of the Beck hopelessness scale. *Cognitive Therapy and Research, 16*, 579-587.

Zelenski, J. M., Murphy, S. A., & Jenkins, D. A. (2008). The happy-productive worker thesis revisited. *Journal of Happiness Studies, 9*, 521-537. doi:10.1007/s10902-008-9087-4

Zimmerman, B. J. (2008). Goal setting: A key proactive source of academic self-regulation. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 267-295). New York, NY: Erlbaum.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Erlbaum.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2004). Self-regulating intellectual processes and outcomes: A social cognitive perspective. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 323-349). Mahwah, NJ: Erlbaum.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York, NY: Routledge.

APPENDICES

APPENDIX A

COMPLETE SURVEY

(The complete survey given to students is on the following four pages. Following the survey is a translation of the instructions into English that was not given to students.)

英語学習者に関するアンケート
このアンケートの目的は、学生が人生や教育、英語に対して一般的に持っている考えや感情をより良く理解することにあります。この結果が単位取得や成績に反映したり、担当の先生に知られたりすることは全くありません。分析が終了次第、アンケートは処分されます。正確なデータを得るために、よく考えて、記入漏れのないように記入してください。学生番号の記入もよろしくお願いします。また、「こうあるべき」、あるいは「こうありたい」、といった答え方ではなく、できる限り事実にあてはまるものを正直に解答してください。

あなた自身について御伺いします。当てはまるところを〇で囲むか、空欄に記入をお願いします。

学生番号 （　　　　　　　　　　　　　　　　　）

性別　男　女　　　　学年　1年　2年　3年　4年

年齢　18　19　20　21　22　23　その他（　　　　）

Part 1
一般的な自分の気持ちについて、どの程度それぞれの気分を感じていますか。
当てはまるものを一つ選び、番号を〇で囲んでください。

全く感じない	めったに感じない	ときどき感じる	しばしば感じる	ほとんどいつも感じる	いつも感じる
1	2	3	4	5	6

例） いつも冷静で、あまり緊張しない人

	落ち着いた	1	2	3	4	5	⑥
	どきどきした	1	②	3	4	5	6
1	うれしい	1	2	3	4	5	6
2	わくわくした	1	2	3	4	5	6
3	恐れている	1	2	3	4	5	6
4	熱狂した	1	2	3	4	5	6
5	おびえた	1	2	3	4	5	6
6	恥じた	1	2	3	4	5	6
7	強気な	1	2	3	4	5	6
8	怒った	1	2	3	4	5	6
9	びくびくした	1	2	3	4	5	6
10	肯定的な	1	2	3	4	5	6
11	誇らしい	1	2	3	4	5	6
12	心配した	1	2	3	4	5	6
13	幸せな	1	2	3	4	5	6
14	苦悩した	1	2	3	4	5	6
15	機敏な	1	2	3	4	5	6
16	気分がいい	1	2	3	4	5	6
17	ぴりぴりした	1	2	3	4	5	6
18	きっぱりとした	1	2	3	4	5	6

19	悲しい	1	2	3	4	5	6
20	いらだった	1	2	3	4	5	6
21	うろたえた	1	2	3	4	5	6
22	否定的な	1	2	3	4	5	6
23	満足した	1	2	3	4	5	6
24	気分が悪い	1	2	3	4	5	6
25	陽気な	1	2	3	4	5	6
26	活気のある	1	2	3	4	5	6
27	不愉快な	1	2	3	4	5	6
28	気合いの入った	1	2	3	4	5	6

Part 2　自分自身について、あなたが一般的に感じたり、行動したりすることについて、また英語の学習について聞きます。

全く当てはまらない	当てはまらない	どちらかというと当てはまらない	どちらかというと当てはまる	当てはまる	非常によく当てはまる
1	2	3	4	5	6

	例）カレーがとても好きで、ハンバーグはあまり好きではなく、ピーマンが人						
	カレーが好きだ	1	2	3	4	5	⑥
	ハンバーグが好きだ	1	2	③	4	5	6
	ピーマンが好きだ	①	2	3	4	5	6
1	いろいろといい素質を持っている。	1	2	3	4	5	6
2	私は英語学習に夢中だ。	1	2	3	4	5	6
3	何かをするとき、一般的に自分は一生懸命にするタイプだ。	1	2	3	4	5	6
4	英語の歌を聴いて主な内容を理解できる。	1	2	3	4	5	6
5	仕事（学業も含む）の将来を考えると、楽しさよりむしろ不愉快さばかりを感じる。	1	2	3	4	5	6
6	過去に感じて来た感情と比べると、今自分はとても幸せだ。	1	2	3	4	5	6
7	人は自分を尊敬している。	1	2	3	4	5	6
8	英語のテレビ番組が見られるならいつも見るだろう。	1	2	3	4	5	6
9	私はこの授業で難しいことを学ぶのが好きだ。	1	2	3	4	5	6
10	数学の理解度は高い。	1	2	3	4	5	6
11	自分自身が問われ、人として成長するような機会をしばしば求めている。	1	2	3	4	5	6
12	夢が実現すれば、将来英語を効果的に使うだろうと思う。	1	2	3	4	5	6
13	私には、必要なときには話せる人がたくさんいる。	1	2	3	4	5	6
14	英語の活動で難しいものがあったときには、より理解できるように時間をもっとかける。	1	2	3	4	5	6
15	新しいことを学ぶためにクラスメートと協力する。	1	2	3	4	5	6
16	英語の新聞記事を読んで理解することができる。	1	2	3	4	5	6
17	将来、人間関係で本当に満足することなどほとんどありえない。	1	2	3	4	5	6
18	学校で自分は社会の一員だと感じる。	1	2	3	4	5	6
19	英語を学んで発見する新しいことのおかげで、英語学習がますますいいことだと感じる。	1	2	3	4	5	6
20	自分には、自慢できることがよくある。	1	2	3	4	5	6
21	学校で友達が問題を抱えているときに自分は友達を助ける。	1	2	3	4	5	6
22	自分は目標を一度設定したら、達成するまで突き進む。	1	2	3	4	5	6
23	ほとんどの面で、私の人生は私の理想に近い。	1	2	3	4	5	6

24	英語を母国語とする人たちの短い会話の、細かい所まで聞き取り理解することができる。	1	2	3	4	5	6
25	一般的に言って、自分はとても幸せだと思う。	1	2	3	4	5	6
26	私は望むような人間関係を全くもっていないし、この先持てるとも思わない。	1	2	3	4	5	6
27	自分の人生に生き甲斐を感じている。	1	2	3	4	5	6
28	正直に言って自分は英語をマスターするために本当によくがんばっていると思う。	1	2	3	4	5	6
29	数学の勉強は楽しい。	1	2	3	4	5	6
30	違う集団や文化の人たちと、私はうまく交流することができる。	1	2	3	4	5	6
31	自分は英語学習を習慣にしようとしている。	1	2	3	4	5	6
32	英語の簡単なウェブページを読んで理解することができる。	1	2	3	4	5	6
33	自分が始めたことは何でもちゃんと終わらせる。	1	2	3	4	5	6
34	私は人生で最高の人間関係を失ってしまい、この先も事態はよくならないだろう。	1	2	3	4	5	6
35	自分は役に立つ人間だと思う。	1	2	3	4	5	6
36	自分は努力家だ。	1	2	3	4	5	6
37	英語の授業で学んでいることは大切だ。	1	2	3	4	5	6
38	友達のほとんどと比べて、自分は幸せだと思う。	1	2	3	4	5	6
39	自分はクラスメートに受け入れられていると感じる。	1	2	3	4	5	6
40	どこに行っても、新しいことや新しい経験を期待している。	1	2	3	4	5	6
41	数学は自分の得意教科の一つだ。	1	2	3	4	5	6
42	私は私の友人達のことを信頼できる。	1	2	3	4	5	6
43	英語の旅行パンフレットを読んで理解することができる。	1	2	3	4	5	6
44	学校で友達が物事がうまくいっていないとき、自分は友達を励ます。	1	2	3	4	5	6
45	自分は善良な人間で、いい人生を送っている。	1	2	3	4	5	6
46	私はこの授業でできるだけ多くのことを学びたい。	1	2	3	4	5	6
47	英語で聞かれた簡単な質問を聞き取り理解することができる。	1	2	3	4	5	6
48	私の人生は、 とてもすばらしい状態だ。	1	2	3	4	5	6
49	英語に関して何らかの問題が生じたら、うまくできるようになるまでより多くの時間をかける。	1	2	3	4	5	6
50	将来、人間関係がよくなる見込みはないと思う。	1	2	3	4	5	6
51	電話で英語で話すことができる。	1	2	3	4	5	6
52	先の読めないことをする方がわくわくするので好きだ。	1	2	3	4	5	6
53	数学の内容をすぐに身につける。	1	2	3	4	5	6
54	将来のことを考えると英語を使うことは大切だと思う。	1	2	3	4	5	6
55	将来の計画について英語で話すことができる。	1	2	3	4	5	6
56	英語のメニューを読んで理解することができる。	1	2	3	4	5	6
57	クラスメートが自分に賛成しないときに、自分の意見を言うことができる。	1	2	3	4	5	6
58	私はこの授業で間違いから学ぶことが好きだ。	1	2	3	4	5	6
59	英語で簡単な質問に答えることができる。	1	2	3	4	5	6
60	数学の成績は良かった。	1	2	3	4	5	6
61	自分は目的のある、意味のある人生を送っている。	1	2	3	4	5	6
62	洋服を買うことについて英語で話すことができる。	1	2	3	4	5	6
63	私は毎日の生活でおこる思いがけない出来事を心から楽しむタイプだ。	1	2	3	4	5	6
64	もし英語の講座があれば将来受講したい。	1	2	3	4	5	6
65	自分はクラスメートに理解されていると感じる。	1	2	3	4	5	6
66	自分は英語の授業で学ぶ内容は個人的に有意義だと思う。	1	2	3	4	5	6

67	英語を学ぶことで様々な経験ができる。	1	2	3	4	5	6
68	もう一度人生をやり直せるとしても、ほとんど何も変えないだろう。	1	2	3	4	5	6
69	英語の映画を字幕なしで見て、細かい内容を英語を聞き取り理解することができる。	1	2	3	4	5	6
70	友人との関係は、私が望むようにはいかないだろう。	1	2	3	4	5	6
71	英語の小説を読んで理解することができる。	1	2	3	4	5	6
72	英語を話せるようになっている自分をよく想像する。	1	2	3	4	5	6
73	人生において自分に大切なことを身につける方法をたくさん思いつくことができる。	1	2	3	4	5	6
74	自分が日常生活の中でしていることに没頭し、興味を持っている。	1	2	3	4	5	6
75	きつい状況は自分が成長したり学んだりするチャンスだと思う。	1	2	3	4	5	6
76	難しい話題について、英語で自分の意見を述べることができる。	1	2	3	4	5	6
77	自分にとって英語を勉強することはとても大切なことだ。	1	2	3	4	5	6
78	仕事（学業も含む）でうまくやろうとしても多分成功しないだろうからやってもむだだ。	1	2	3	4	5	6
79	私の目標は授業中に示される内容を全て理解することだ。	1	2	3	4	5	6
80	ピンチを切り抜けるために、自分は多くの方法を思いつくことができる。	1	2	3	4	5	6
81	少し怖いと思うことをすることがしばしば好きだ。	1	2	3	4	5	6
82	重要な課題を達成するために挫折を乗り越えたことがある。	1	2	3	4	5	6
83	この授業で私は英語の力を伸ばすことに集中している。	1	2	3	4	5	6
84	他の人との付き合いでは、本当に欲しいものは手に入らないだろう。	1	2	3	4	5	6
85	人は親切だと私は信じている。	1	2	3	4	5	6
86	英語はとても面白い学問分野だ。	1	2	3	4	5	6
87	全体的に、私は今とても幸せな人生を送っている。	1	2	3	4	5	6
88	だいたいにおいて、自分に満足している。	1	2	3	4	5	6
89	英語は意欲的に勉強している。	1	2	3	4	5	6
90	自分は自分の目標に向かってエネルギッシュに取り組む。	1	2	3	4	5	6
91	英語の授業はあっという間に一時間が終わる。	1	2	3	4	5	6
92	英語を学ぶことで思い出深い経験ができる。	1	2	3	4	5	6
93	自分が自分自身や世の中をどう思っているかを試されるような経験をすることをいつも探している。	1	2	3	4	5	6
94	数学が得意だ。	1	2	3	4	5	6
95	他人との人間関係により、自分は支えられており、嬉しいと感じる。	1	2	3	4	5	6
96	英語学習は自分の人生の他の活動と関連している。	1	2	3	4	5	6
97	これまでの人生はかなりうまくいってきている。	1	2	3	4	5	6
98	「とても幸せで、何が起こっていても人生を楽しみ、多くのことを得る人がいます。」この描写に自分は当てはまると思う。	1	2	3	4	5	6
99	英語を母国語とする人たちの短い会話の、主な内容を聞き取り理解することができる。	1	2	3	4	5	6
100	英語の交通標識を見て理解することができる。	1	2	3	4	5	6
101	どんな問題でも解決する方法が多くあると自分は思う。	1	2	3	4	5	6
102	私は何年もかかって目標を達成したことがある。	1	2	3	4	5	6
103	学校や仕事のことを英語で話すことができる。	1	2	3	4	5	6
104	自分で決めた目標は達成する。	1	2	3	4	5	6
105	私は自分の人生に満足している。	1	2	3	4	5	6
106	空港で英語のアナウンスを聞き取り理解することができる。	1	2	3	4	5	6
107	この先自分の仕事(学業も含む)に本当に満足することなどほとんどありえない。	1	2	3	4	5	6
108	私はこの授業で新しい内容を学ぶのが好きだ。	1	2	3	4	5	6
109	英語の映画を字幕なしで見て、主な内容を英語を聞いて理解できる。	1	2	3	4	5	6

110	自分にとって大切な活動をする力量と能力が自分にはある。	1	2	3	4	5	6
111	他の人たちがやる気をなくしていても、自分は問題を解決する方法を見つけることができる。	1	2	3	4	5	6
112	人生がうまくいっていると思うことがよくある。	1	2	3	4	5	6
113	英語を使ってレストランで注文ができる。	1	2	3	4	5	6
114	私は新しい状況ではできる限りの情報を活発に得る。	1	2	3	4	5	6
115	いつも数学が得意だった。	1	2	3	4	5	6
116	私は将来、より良い社会を作るのに貢献するだろう。	1	2	3	4	5	6
117	毎週英語は何時間も自分で勉強している。	1	2	3	4	5	6
118	英語で演説ができる。	1	2	3	4	5	6
119	仕事（学業も含む）は、私が望むようにはいかないだろう。	1	2	3	4	5	6
120	自分がなりたいタイプの人間に自分はなるだろうという希望に満ちている。	1	2	3	4	5	6
121	英語の授業を楽しみにしている。	1	2	3	4	5	6
122	外国人の友達と英語で話している自分を思い浮かべることができる。	1	2	3	4	5	6
123	ラジオで英語の曲を聴くときには注意深く聞いて歌詞を理解しようとしている。	1	2	3	4	5	6
124	英語にとても魅力を感じる。	1	2	3	4	5	6
125	将来英語を使って仕事をしている自分をよく想像する。	1	2	3	4	5	6
126	あることへの興味を何年も持続できる。	1	2	3	4	5	6
127	英語の演説を聞いて主な考えを理解できる。	1	2	3	4	5	6
128	見知らぬ人、出来事、場所との出会いに喜んで応じるタイプだ。	1	2	3	4	5	6
129	将来、仕事（学業も含む）で成功する見込みはあまりないように思う。	1	2	3	4	5	6
130	英語を理解するときに問題が生じたら、わかるまで努力を続ける。	1	2	3	4	5	6
131	私はこの講座の内容をできるだけ完全に理解しようと努力している。	1	2	3	4	5	6
132	私は他人との間に暖かく信頼できる人間関係を経験して来た。	1	2	3	4	5	6
133	英語の習得のためによく努力する方だ。	1	2	3	4	5	6
134	私は自分の将来について楽観的だ。	1	2	3	4	5	6
135	最近自分は人生をとても楽しんでいるとしばしば感じる。	1	2	3	4	5	6
136	過去の経験のおかげで自分の将来に対する心構えが充分にできた。	1	2	3	4	5	6
137	終わるまでに何ヶ月もかかる学習課題に集中できる。	1	2	3	4	5	6
138	自分について好きなところに一つは英語を学んでいるところだ。	1	2	3	4	5	6
139	英語のラジオ番組を聞いて主な内容を理解できる。	1	2	3	4	5	6
140	物事を人並みにはうまくやれる。	1	2	3	4	5	6
141	数学は自分にとって簡単だ。	1	2	3	4	5	6
142	将来のやりたいことのためには英語を話すことが必要だ。	1	2	3	4	5	6
143	パーティーで自己紹介を英語でできる。	1	2	3	4	5	6
144	英語に対してとてもわくわくする。	1	2	3	4	5	6
145	学校で英語の授業がなかったらどこかほかに英語の勉強をしにいく。	1	2	3	4	5	6
146	私は英語学習に対して強い関心がある。	1	2	3	4	5	6
147	ある学習課題に取り組んでいるときに、新しい考えや他の課題に気が散ることはない。	1	2	3	4	5	6
148	少なくとも人並みには価値のある人間だと思う。	1	2	3	4	5	6
149	ほとんどの人は私のことを思いやりのある人間だと思っている。	1	2	3	4	5	6
150	私の目標はこの授業でできるだけ多くのことを学ぶことだ。	1	2	3	4	5	6
151	英語の授業は楽しい。	1	2	3	4	5	6
152	教室内での英語の勉強に、自分はついていけており、取り残されてはいない。	1	2	3	4	5	6
153	私が本当に望む仕事上（学業も含む）の目標は、決して手に入らないだろう。	1	2	3	4	5	6
154	複雑なこと、骨の折れることをすることは自分に合っている。	1	2	3	4	5	6

155	英文のＥメールの主な内容を読んで理解することができる。	1	2	3	4	5	6
156	私はこれまで、自分の人生に求める大切なものを得てきた。	1	2	3	4	5	6
157	私は他人の幸福に積極的に貢献している。	1	2	3	4	5	6
158	たいてい、自分はとても幸せだと感じている。	1	2	3	4	5	6
159	今よりも良い人間に自分はなると期待している。	1	2	3	4	5	6

ご協力ありがとうございました！

English Learner Questionnaire (Instructions translated from Japanese)

This survey is conducted to better understand the general thoughts and feelings students have of their life, education, and studying English. Your results of this survey will not affect your grades nor be given to teachers. After the analysis is done, this questionnaire will be disposed. In order to get precise data, please carefully consider and answer all the questions. Please write down your matriculation number. Please give your answer sincerely, not because you feel that you should answer a particular way, but because this expresses your honest response.

This section asks you about yourself. Please fill out this form by circling a number or writing down in parentheses.

Your matriculation number (   )

Sex: Male Female   Year: 1 2 3 4

Age: 18 19 20 21 22 23 other (   )

Part 1
In general, how often do you have the feelings below? Choose the most appropriate one and circle the number.

Never	Hardly ever	Sometimes	Often	Almost always	Always
1	2	3	4	5	6

Example: a person who is always calm and who hardly ever gets nervous

calm	1 2 3 4 5 ⑥
heart races with excitement	1 ② 3 4 5 6

Part 2
This section asks you about how you generally feel or act, or about learning English and so on. Please choose the most appropriate one and circle the number.

Definitely not true of me	Not true of me	Slightly not true of me	Slightly true of me	True of me	Definitely true of me
1	2	3	4	5	6

Example: a person who loves curry rice, doesn't like hamburger so much and hates bell peppers

I like curry rice	1 2 3 4 5 ⑥
I like hamburgers	1 2 ③ 4 5 6
I like bell peppers	① 2 3 4 5 6

APPENDIX B

CURIOSITY AND EXPLORATION INVENTORY-II

	English	Japanese	Item #
1.	I actively seek as much information as I can in new situations.	私は新しい状況ではできる限りの情報を活発に得る。	114
2.	I am the type of person who really enjoys the uncertainty of everyday life.	私は毎日の生活でおこる思いがけない出来事を心から楽しむタイプだ。	63
3.	I am at my best when doing something that is complex or challenging.	複雑なこと、骨の折れることをすることは自分に合っている。	154
4.	Everywhere I go, I am out looking for new things or experiences.	どこに行っても、新しいことや新しい経験を期待している。	40
5.	I view challenging situations as an opportunity to grow and learn.	きつい状況は自分が成長したり学んだりするチャンスだと思う。	75
6.	I like to do things that are a little frightening.	少し怖いと思うことをすることがしばしば好きだ。	81
7.	I am always looking for experiences that challenge how I think about myself and the world.	自分が自分自身や世の中をどう思っているかを試されるような経験をすることをいつも探している。	93
8.	I prefer jobs that are excitingly unpredictable.	先の読めないことをする方がわくわくするので好きだ。	52
9.	I frequently seek out opportunities to challenge myself and grow as a person.	自分自身が問われ、人として成長するような機会をしばしば求めている。	11
10.	I am the kind of person who embraces unfamiliar people, events, and places.	見知らぬ人、出来事、場所との出会いに喜んで応じるタイプだ。	128

APPENDIX C

HOPE SCALE

	English	Japanese	Item #
1.	I can think of many ways to get out of a jam.	ピンチを切り抜けるために、自分は多くの方法を思いつくことができる。	80
2.	There are lots of ways around any problem.	どんな問題でも解決する方法が多くあると自分は思う。	101
3.	I can think of many ways to get the things in life that are important to me.	人生において自分に大切なことを身につける方法をたくさん思いつくことができる。	73
4.	Even when others get discouraged, I know I can find a way to solve the problem.	他の人たちがやる気をなくしていても、自分は問題を解決する方法を見つけることができる。	111
5.	I energetically pursue my goals.	自分は自分の目標に向かってエネルギッシュに取り組む。	90
6.	My past experiences have prepared me well for my future.	過去の経験のおかげで自分の将来に対する心構えが充分にできた。	136
7.	I've been pretty successful in life.	これまでの人生はかなりうまくいってきている。	97
8.	I meet the goals that I set for myself.	自分で決めた目標は達成する。	104
9.	I feel hopeful about being a better person than I am now.	今よりも良い人間に自分はなると期待している。	159
10.	I feel hopeful about becoming the type of person I aspire to become.	自分がなりたいタイプの人間に自分はなるだろうという希望に満ちている。	120

APPENDIX D

FLOURISHING SCALE

	English	Japanese	Item #
1.	I lead a purposeful and meaningful life.	自分は目的のある、意味のある人生を送っている。	61
2.	My social relationships are supportive and rewarding.	他人との人間関係により、自分は支えられており、嬉しいと感じる。	95
3.	I am engaged and interested in my daily activities.	自分が日常生活の中でしていることに没頭し、興味を持っている。	74
4.	I actively contribute to the happiness and well-being of others.	私は他人の幸福に積極的に貢献している。	157
5.	I am competent and capable in the activities that are important to me.	自分にとって大切な活動をする力量と能力が自分にはある。	110
6.	I am a good person and live a good life.	自分は善良な人間で、いい人生を送っている。	45
7.	I am optimistic about my future.	私は自分の将来について楽観的だ。	134
8.	People respect me.	人は自分を尊敬している。	7
X.	I feel *ikigai* (生き甲斐) in my life.	自分の人生に生き甲斐を感じている。	27

Note. Item # = Item number in final instrument; X = item eliminated during study.

APPENDIX E

SELF-ESTEEM SCALE

	English	Japanese	Item #
1.	I feel that I'm a person of worth, at least on an equal plane with others.	少なくとも人並みには価値のある人間だと思う。	148
2.	I feel that I have a number of good qualities.	いろいろといい素質を持っている。	1
3.	I often think my life is going well.	人生がうまくいっていると思うことがよくある。	112
4.	I am able to do things as well as most other people.	物事を人並みにはうまくやれる。	140
5.	I often feel I have much to be proud of.	自分には、自慢できることがよくある。	20
6.	On the whole, I am satisfied with myself.	だいたいにおいて、自分に満足している。	88
7.	I think I am useful person at times.	自分は役に立つ人間だと思う。	35

APPENDIX F

SATISFACTION WITH LIFE SCALE

	English	Japanese	Item #
1.	In most ways my life is close to ideal.	ほとんどの面で、私の人生は私の理想に近い。	23
2.	The conditions of my life are excellent.	私の人生は、とてもすばらしい状態だ。	48
3.	I am satisfied with my life.	私は自分の人生に満足している。	105
4.	So far I have gotten the important things I want in life.	私はこれまで、自分の人生に求める大切なものを得てきた。	156
5.	If I could live my life over, I would change almost nothing.	もう一度人生をやり直せるとしても、ほとんど何も変えないだろう。	68

APPENDIX G

POSITIVE AND NEGATIVE AFFECT SCALE

	English	Japanese	Item #
	Positive Affect Scale		
1.	Active	活気のある	26
2.	Proud	誇らしい	11
3.	Strong	強気な	7
4.	Attentive	気合いの入った	28
5.	Determined	きっぱりとした	18
6.	Excited	わくわくした	2
7.	Alert	機敏な	15
8.	Enthusiastic	熱狂した	4
	Negative Affect Scale		
1.	Scared	びくびくした	9
2.	Afraid	おびえた	5
3.	Upset	うろたえた	21
4.	Nervous	心配した	12
5.	Distressed	苦悩した	14
6.	Jittery	ぴりぴりした	17
7.	Ashamed	恥じた	6
8.	Irritable	いらだった	20

APPENDIX H

POSITIVE AND NEGATIVE FEELINGS SCALE

	English	Japanese	Item #
Positive Feelings Scale			
1.	Positive	肯定的な	10
2.	Good	気分がいい	16
3.	Pleasant	うれしい	1
4.	Happy	幸せな	13
5.	Joyful	陽気な	25
6.	Contented	満足な	23
Negative Feelings Scale			
1.	Negative	否定的な	22
2.	Bad	気分が悪い	24
3.	Unpleasant	不愉快な	27
4.	Sad	悲しい	19
5.	Afraid	恐れている	3
6.	Angry	怒った	8

APPENDIX I

SUBJECTIVE HAPPINESS SCALE

	English	Japanese	Item #
1.	In general, I consider myself a very happy person.	一般的に言って、自分はとても幸せだと思う。	25
2.	I'm happy compared to most of my peers.	友達のほとんどと比べて、自分は幸せだと思う。	38
3.	Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. This describes myself.	「とても幸せで、何が起こっていても人生を楽しみ、多くのことを得る人がいます。」この描写に自分は当てはまると思う。	98
4.	I am happy now compared to how I have felt in the past.	過去に感じて来た感情と比べると、今自分はとても幸せだ。	6
5.	On the whole, I am living a very happy life now	全体的に、私は今とても幸せな人生を送っている。	87
6.	Usually, I feel extremely happy.	たいてい、自分はとても幸せだと感じている。	158
7.	I often feel that I am really enjoying life these days.	最近自分は人生をとても楽しんでいるとしばしば感じる。	135

APPENDIX J

POSITIVE SOCIAL RELATIONSHIPS SCALE

	English	Japanese	Item #
1.	I have many people who I can talk to when I need to.	私には、必要なときには話せる人がたくさんいる。	13
2.	I can trust my friends.	私は私の友人達のことを信頼できる。	42
3.	Most people see me as a compassionate person.	ほとんどの人は私のことを思いやりのある人間だと思っている。	149
4.	I have experienced warm and trusting relationships with others.	私は他人との間に暖かく信頼できる人間関係を経験して来た。	132
5.	I believe that people are kind.	人は親切だと私は信じている。	85
6.	I will contribute to making society better in the future.	私は将来、より良い社会を作るのに貢献するだろう。	116
7.	I can interact well with people from different groups or cultures.	違う集団や文化の人たちと、私はうまく交流することができる。	30

APPENDIX K

GRIT SCALE

	English	Japanese	Item #
1.	Once I set a goal, I pursue it until finished.	自分は目標を一度設定したら、達成するまで突き進む。	22
2.	I do not get distracted by new ideas or new projects when working on a project.	ある学習課題に取り組んでいるときに、新しい考えや他の課題に気が散ることはない。	147
3.	I can focus on projects that take months to complete.	終わるまでに何ヶ月もかかる学習課題に集中できる。	137
4.	I can maintain my interest in topics for years.	あることへの興味を何年も持続できる。	126
5.	I have achieved a goal that took years of work.	私は何年もかかって目標を達成したことがある。	102
6.	I have overcome setbacks to complete an important challenge.	重要な課題を達成するために挫折を乗り越えたことがある。	82
7.	I finish whatever I begin.	自分が始めたことは何でもちゃんと終わらせる。	33
8.	I am a hard worker.	自分は努力家だ。	36
9.	When I do things, generally I am a hard worker.	何かをするとき、一般的に自分は一生懸命にするタイプだ。	3

APPENDIX L

ACHIEVEMENT HOPELESSNESS SCALE

	English	Japanese	Item #
1.	There's no use trying to do well on my work because it won't be successful.	仕事（学業も含む）でうまくやろうとしても多分成功しないだろうからやってもむだだ。	78
2.	There is almost no possibility for me to be really satisfied with my work.	この先自分の仕事(学業も含む)に本当に満足することなどほとんどありえない。	107
3.	In the future, I don't think I'll be successful at my work.	将来、仕事（学業も含む）で成功する見込みはあまりないように思う。	129
4.	I'll never achieve my goal I really want to reach in my work.	私が本当に望む仕事上（学業も含む）の目標は、決して手に入らないだろう。	153
5.	My work will not be like what I want.	仕事（学業も含む）は、私が望むようにはいかないだろう。	119
6.	When I think of my future, I feel rather upset than pleased.	仕事（学業も含む）の将来を考えると、楽しさよりむしろ不愉快さばかりを感じる。	5

APPENDIX M

RELATIONSHIP HOPELESSNESS SCALE

	English	Japanese	Item #
1.	I don’t think I can get what I really want when it comes to relationships with others.	他の人との付き合いでは、本当に欲しいものは手に入らないだろう。	84
2.	I don’t have nice relationships with others and I don’t expect to in the future.	私は望むような人間関係を全くもっていないし、この先持てるとも思わない。	26
3.	I don’t think I’ll be satisfied with my relationships with others in the future.	将来、人間関係で本当に満足することなどほとんどありえない。	17
4.	I’ve lost my best relationship with my friend and I don’t think it will be fixed in the future.	私は人生で最高の人間関係を失ってしまい、この先も事態はよくならないだろう。	34
5.	I don’t think I’ll have better relationships with others in the future.	将来、人間関係がよくなる見込みはないと思う。	50
6.	I don’t think my relationships with my friends will become what I want.	友人との関係は、私が望むようにはいかないだろう。	70

APPENDIX N

INTEREST-IN-L2 SELF SCALE

	English	Japanese	Item #
1.	I'm looking forward to taking English classes.	英語の授業を楽しみにしている。	121
X.	One English class passes like a moment.	英語の授業はあっという間に一時間が終わる。	91
2.	English lessons are enjoyable.	英語の授業は楽しい。	151
3.	English is a very interesting field of study.	英語はとても面白い学問分野だ。	86
4.	English is very fascinating.	英語にとても魅力を感じる。	124
5.	I'm very excited about English.	英語に対してとてもわくわくする。	144
6.	What I learn in English lessons is important.	英語の授業で学んでいることは大切だ。	37
7.	I personally think what I learn in English class is meaningful.	自分は英語の授業で学ぶ内容は個人的に有意義だと思う。	66

Note. Item # = Item number in final instrument; X = item eliminated during study.

APPENDIX O

PASSION FOR L2 LEARNING SCALE

	English	Japanese	Item #
1.	Learning English allows me to live a variety of experiences.	英語を学ぶことで様々な経験ができる。	67
2.	The new things that I discover through learning English allow me to appreciate learning English even more.	英語を学んで発見する新しいことのおかげで、英語学習がますますいいことだと感じる。	19
3.	Learning English allows me to live memorable experiences.	英語を学ぶことで思い出深い経験ができる。	92
4.	One of the qualities I like about myself is learning English.	自分について好きなところに一つは英語を学んでいるところだ。	138
5.	Learning English is related to other activities in my life.	英語学習は自分の人生の他の活動と関連している。	96
6.	I am passionate about learning English.	私は英語学習に対して強い関心がある。	146
7.	I am completely absorbed with learning English.	私は英語学習に夢中だ。	2

APPENDIX P

MASTERY GOAL ORIENTATION SCALE

	English	Japanese	Item #
X.	My aim is to completely master the material presented in this class.	私は英語学習に夢中だ。私の目標は授業中に示される内容を全て理解することだ。	79
1.	I am striving to understand the content of this course as thoroughly as possible.	私はこの講座の内容をできるだけ完全に理解しようと努力している。	131
2.	My goal is to learn as much as possible in this class.	私の目標はこの授業でできるだけ多くのことを学ぶことだ。	150
3.	I want to learn as much as possible in this class.	私はこの授業でできるだけ多くのことを学びたい。	46
4.	I like learning difficult things in this class.	私はこの授業で難しいことを学ぶのが好きだ。	9
5.	I like learning from my mistakes in this class.	私はこの授業で間違いから学ぶことが好きだ。	58
6.	In this class I focus on developing my English language skills.	この授業で私は英語の力を伸ばすことに集中している。	83
7.	I like to study new topics in this class.	私はこの授業で新しい内容を学ぶのが好きだ。	108

Note. Item # = Item number in final instrument; X = item eliminated during study.

APPENDIX Q

IDEAL L2 SELF SCALE

	English	Japanese	Item #
1.	The things I want to do in the future require me to speak English.	将来のやりたいことのためには英語を話すことが必要だ。	142
2.	Whenever I think of my future career, I imagine myself being able to use English.	将来英語を使って仕事をしている自分をよく想像する。	125
3.	I often imagine myself as someone who is able to speak English.	英語を話せるようになっている自分をよく想像する。	72
4.	If my dreams come true, I will use English effectively in the future.	夢が実現すれば、将来英語を効果的に使うだろうと思う。	12
5.	I can imagine speaking English with international friends.	外国人の友達と英語で話している自分を思い浮かべることができる。	122
6.	When I think about my future, it is important that I use English.	将来のことを考えると英語を使うことは大切だと思う。	54

APPENDIX R

PROSOCIALITY SCALE

	English	Japanese	Item #
1.	I help other school friends when they have a problem.	学校で友達が問題を抱えているときに自分は友達を助ける。	21
2.	I cheer up school friends when something has gone wrong.	学校で友達が物事がうまくいっていないとき、自分は友達を励ます。	44
3.	I cooperate with my classmates to learn new things.	新しいことを学ぶためにクラスメートと協力する。	15
4.	I can express my opinions when my classmates disagree with me.	クラスメートが自分に賛成しないときに、自分の意見を言うことができる。	57
5.	In college I feel a sense of social belonging.	学校で自分は社会の一員だと感じる。	18
6.	I feel I am accepted by my classmates.	自分はクラスメートに受け入れられていると感じる。	39
7.	I feel understood by my classmates.	自分はクラスメートに理解されていると感じる。	65

APPENDIX S

MATH SELF-CONCEPT SCALE

	English	Japanese	Item #
1.	I am very good at mathematics.	数学が得意だ。	94
2.	I have always been good at mathematics.	いつも数学が得意だった。	115
3.	Mathematics work is easy for me.	数学は自分にとって簡単だ。	141
4.	I get good grades in mathematics.	数学の成績は良かった。	60
5.	I learn things quickly in mathematics.	数学の内容をすぐに身につける。	53
6.	Mathematics is one of my best subjects.	数学は自分の得意教科の一つだ。	41
7.	I enjoy learning mathematics.	数学の勉強は楽しい。	29
8.	I have a good understanding of mathematics.	数学の理解度は高い。	10

APPENDIX T

SPEAKING SELF-EFFICACY SCALE

	English	Japanese	Item #
1.	I can speak English to order a meal in a restaurant.	英語を使ってレストランで注文ができる。	113
2.	I can answer simple questions by speaking in English.	英語で簡単な質問に答えることができる。	59
3.	I can introduce myself at a party speaking in English.	パーティーで自己紹介を英語でできる。	143
4.	I can talk about shopping for clothes in English.	洋服を買うことについて英語で話すことができる。	62
5.	I can talk about my school or job in English.	学校や仕事のことを英語で話すことができる。	103
6.	I can talk about my future plans in English.	将来の計画について英語で話すことができる。	55
7.	I can talk on the telephone in English.	電話で英語で話すことができる。	51
8.	I can give my opinions and talk about difficult topics in English.	難しい話題について、英語で自分の意見を述べることができる。	76
9.	I can give a speech in English.	英語で演説ができる。	118

APPENDIX U

LISTENING SELF-EFFICACY SCALE

	English	Japanese	Item #
1.	I can listen to and understand the main ideas in short English native speaker conversations.	英語を母国語とする人たちの短い会話の、主な内容を聞き取り理解することができる。	99
2.	I can listen to and understand announcements made in English at the airport.	空港で英語のアナウンスを聞き取り理解することができる。	106
3.	I can listen to and understand simple questions asked in English.	英語で聞かれた簡単な質問を聞き取り理解することができる。	47
4.	I can listen to and understand most details of short English native speaker conversations.	英語を母国語とする人たちの短い会話の、細かい所まで聞き取り理解することができる。	24
5.	I can understand the main ideas when listening to and watching English movies without Japanese subtitles.	英語の映画を字幕なしで見て、主な内容を英語を聞いて理解できる。	109
6.	I can understand the main ideas when listening to English songs.	英語の歌を聴いて主な内容を理解できる。	4
7.	I can listen to and understand the main ideas in an English speech.	英語の演説を聞いて主な考えを理解できる	127
8.	I can listen to and understand the main ideas in radio shows in English.	英語のラジオ番組を聞いて主な内容を理解できる。	139
9.	I can understand most of the details when listening to and watching English movies without Japanese subtitles.	英語の映画を字幕なしで見て、細かい内容を英語を聞き取り理解することができる。	69

APPENDIX V

READING SELF-EFFICACY SCALE

	English	Japanese	Item #
1.	I can read and understand the main ideas of emails in English.	英文のＥメールの主な内容を読んで理解することができる。	155
2.	I can read and understand a menu in English.	英語のメニューを読んで理解することができる。	56
3.	I can read and understand English traffic signs.	英語の交通標識を見て理解することができる。	100
4.	I can read and understand a travel pamphlet in English.	英語の旅行パンフレットを読んで理解することができる。	43
5.	I can read and understand newspaper articles in English.	英語の新聞記事を読んで理解することができる。	16
6.	I can read and understand novels in English.	英語の小説を読んで理解することができる。	71
7.	I can read and understand simple web pages in English.	英語の簡単なウェブページを読んで理解することができる。	32

APPENDIX W

INTENDED LEARNING EFFORT SCALE

	English	Japanese	Item #
1.	I am working hard at learning English.	英語は意欲的に勉強している。	89
2.	It is extremely important for me to learn English.	自分にとって英語を勉強することはとても大切なことだ。	77
3.	If an English course was offered in the future, I would like to take it.	もし英語の講座があれば将来受講したい。	64
4.	When I hear an English song on the radio, I listen carefully and try to understand all the words.	ラジオで英語の曲を聴くときには注意深く聞いて歌詞を理解しようとしている。	123
5.	I can honestly say that I am really doing my best to learn English.	正直に言って自分は英語をマスターするために本当によくがんばっていると思う。	28
6.	If I could have access to English-speaking TV stations, I would try to watch them often.	英語のテレビ番組が見られるならいつも見るだろう。	8
7.	I am the kind of person who makes great efforts to learn English.	英語の習得のためによく努力する方だ。	133
8.	If English were not taught in school, I would try to go to English classes somewhere else.	学校で英語の授業がなかったらどこかほかに英語の勉強をしにいく。	145

APPENDIX X

PERSISTENT EFFORT AT L2 LEARNING SCALE

	English	Japanese	Item #
1.	When I have a problem understanding English, I keep trying until I understand.	英語を理解するときに問題が生じたら、わかるまで努力を続ける。	130
2.	I keep up and don't get behind with my English schoolwork.	教室内での英語の勉強に、自分はついていけており、取り残されてはいない。	152
3.	I study English by myself hours each week.	毎週英語は何時間も自分で勉強している。	117
4.	If I have trouble learning something about English, I spend more time until I do better.	英語に関して何らかの問題が生じたら、うまくできるようになるまでより多くの時間をかける。	49
5.	When an English activity is difficult for me, I spend extra time so that I understand it well.	英語の活動で難しいものがあったときには、より理解できるように時間をもっとかける	14
6.	I try to make my English study a habit.	自分は英語学習を習慣にしようとしている。	31