

UNIVERSITY OF CALGARY

Examining Youth Leadership Through the Development and Validation
of a Self-Report Measure: The Youth Leadership Questionnaire (YLQ)

by

Yvonne Hindes

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DIVISION OF APPLIED PSYCHOLOGY

CALGARY, ALBERTA

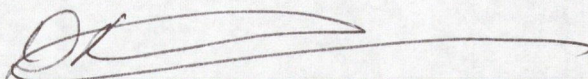
MARCH, 2011

© Yvonne Hindes 2011

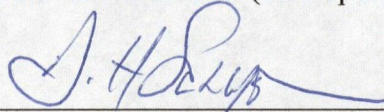
UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Examining Youth Leadership Through the Development and Validation of a Self-Report Measure: The Youth Leadership Questionnaire (YLQ)" submitted by Yvonne Hindes in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

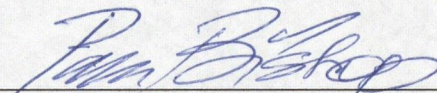
Dr. Vicki Schwean (Supervisor), Division of Applied Psychology



Dr. David Nordstokke (Co-Supervisor), Division of Applied Psychology



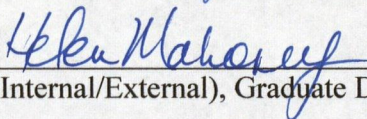
Dr. Don Saklofske (Supervisory Committee), Division of Applied Psychology



Dr. Pam Bishop (Supervisor Committee), Graduate Division of Educational Research



Dr. Michelle Drefs (Internal), Division of Applied Psychology



Dr. Helen Mahoney (Internal/External), Graduate Division of Educational Research

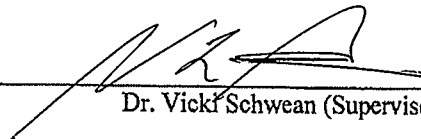
Dr. Dorothy Andrews (External), University of Southern Queensland

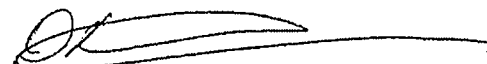
March 21 11

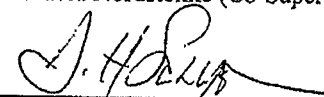
Date

UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Examining Youth Leadership Through the Development and Validation of a Self-Report Measure: The Youth Leadership Questionnaire (YLQ)" submitted by Yvonne Hindes in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

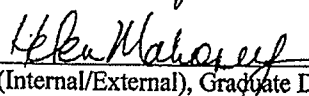

Dr. Vicki Schwean (Supervisor), Division of Applied Psychology

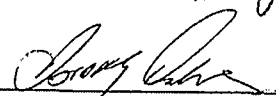

Dr. David Nordstokke (Co-Supervisor), Division of Applied Psychology


Dr. Don Saklofske (Supervisory Committee), Division of Applied Psychology


Dr. Pam Bishop (Supervisor Committee), Graduate Division of Educational Research


Dr. Michelle Drefs (Internal), Division of Applied Psychology


Dr. Helen Mahoney (Internal/External), Graduate Division of Educational Research


Dr. Dorothy Andrews (External), University of Southern Queensland

March 2 / 11
Date

Abstract

A measure of youth leadership (YLQ) was developed and validated with a sample of high school students ($N = 694$) and undergraduate university students ($N = 73$). Empirical analysis and theoretical review reduced a 25-item pool to an 18-item self-report questionnaire with 3 factors: Self-Concept, Emotional Intelligence, and Self-Efficacy. Internal consistency reliability and test-retest reliability were at an exceptional level. Evidence for construct validity and the fit of the model were also adequate. The results were preliminary and represented the initial steps in the construction of the YLQ. With further development, the YLQ could provide researchers, educators, and clinicians with a useful tool to use when answering various questions related to youth leadership.

Acknowledgements

The completion of this dissertation represents an achievement of a goal and a dream, and I would like to take this opportunity to acknowledge and thank those individuals who played an instrumental role in my success.

I would like to express my sincere and great appreciation to my supervisors, Drs. Vicki Schwean and David Nordstokke, for their ongoing direction, support, supervision, encouragement, and dedication throughout this endeavour. I would also like to thank Drs. Don Saklofske and Pam Bishop for being a part of my committee and providing me with additional support and guidance through the entire process. I am very fortunate to have worked with all of you, and you have definitely made this experience a positive one. Additionally, I would like to thank Drs. Michelle Drefs, Helen Mahoney, and Dorothy Andrews for agreeing and taking the time to be examiners for my dissertation defense. Your questions, comments, and support were very much appreciated.

I would like to thank the Social Sciences and Humanities Research Council (SSHRC) and the Werklund Foundation for funding this research. I also would like to extend my gratitude to the Werklund Foundation, the Rapport Teen Leadership Breakthrough Program, the Calgary Board of Education, and the University of Calgary for supporting and assisting with the recruitment and collection of data. Their interest in my research and contributions to the study were exemplary and made this project possible. Moreover, I would like to acknowledge and thank the participants for their willingness to take part in this research.

Lastly, I would like to express my thanks and gratitude to my family and friends who have shown me much patience, support, motivation, and encouragement every step of the way throughout my entire education.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vii
List of Figures.....	viii
 CHAPTER ONE: INTRODUCTION.....	 1
Leadership.....	3
Theoretical models.....	3
Trait theory.....	3
Behavioural theory.....	4
Skills theory.....	5
Constructivist developmental theory.....	5
Transactional theory.....	6
Transformational theory.....	6
Issues with the construct of leadership.....	7
Conceptual definition of youth leadership.....	7
Conceptual model of youth leadership.....	8
Leadership and Adolescence.....	10
Youth leadership programs.....	12
Summary.....	12
 CHAPTER TWO: LITERATURE REVIEW.....	 14
Test Construction.....	14
Stages of test construction.....	15
Choosing the rating scale.....	15
Item development and selection.....	17
Validity.....	19
Models of validity.....	19
Criterion model.....	19
Content model.....	20
Construct model.....	22
Current notion of validity.....	24
Reliability.....	25
Summary.....	27
 CHAPTER THREE: METHODS.....	 29
Objectives of the Current Study	29
Methods.....	30
Participants.....	30
Measures.....	32
The Youth Leadership Questionnaire.....	32
The Resiliency Scale for Children and Adolescents-Adult Revision	32
The Satisfaction With Life Scale.....	33

Procedure.....	34
Stage one: Test construction.....	34
Stage two: Pilot study.....	36
Stage three: Test validation.....	36
Factor analysis.....	38
Reliability.....	40
Construct validity.....	40
Summary.....	41
CHAPTER FOUR: RESULTS.....	42
Results.....	42
Exploratory factor analysis.....	42
Confirmatory factor analysis.....	43
Control group factor analysis.....	44
Reliability.....	45
Construct validity.....	46
Summary.....	46
CHAPTER FIVE: DISCUSSION.....	48
Significance/Implications.....	51
Limitations.....	53
Future Directions.....	54
REFERENCES.....	57
TABLE 1.....	71
TABLE 2.....	72
TABLE 3.....	73
TABLE 4.....	76
TABLE 5.....	77
TABLE 6.....	78
TABLE 7.....	79
TABLE 8.....	80
TABLE 9.....	81
TABLE 10.....	82
TABLE 11.....	83

TABLE 12.....	84
TABLE 13.....	85
TABLE 14.....	86
TABLE 15.....	87
FIGURE 1.....	88
APPENDIX A: The YLQ.....	89
APPENDIX B: Standardized Residuals for Each of the Models.....	92

Lists of Tables

Table 1. University Participants' Area of Study/Degree Programs

Table 2. Descriptive Statistics of the Groups

Table 3. YLQ Items and Their Associated Factors

Table 4. Regression Equations for the Initial EFA

Table 5. Factor Loadings for the Initial EFA

Table 6. Factor Loadings for the EFA on the Revised YLQ

Table 7. Regression Equations for the EFA on the Revised YLQ

Table 8. Factor Loadings for the CFA

Table 9. Regression Equations for the CFA

Table 10. Factor Loadings for the Control Group

Table 11. Regression Equations for the Control Group

Table 12. Coefficient Alpha for the YLQ

Table 13. Inter-Item Correlations for the YLQ

Table 14. Item Total Correlations for the YLQ

Table 15. Construct Validity Correlations of the YLQ With Other Measures

List of Figures

Figure 1. Conceptual Model of Youth Leadership

Chapter One- Introduction

The belief that more effective leaders play a vital role in building a better society and world suggests that the enhancement of leadership skills and the empowerment of students should be a critical and integral part of the learning and educational process (Astin & Astin, 2000), starting with the educational experience of children and youth.

Peace, democracy, and sustainability are three abiding global themes for the 21st century. To achieve any one of these throughout the world will require extraordinary efforts on behalf of humankind and the planet. To achieve all three will require unprecedented levels of commitment, collaboration, and sound leadership from the world's nations. Indeed, because of the complexity of this ambitious quest, it will require a breadth and depth of leadership capacity that is greater than has ever been developed. Leadership from all areas of societies will be needed. Youth leadership, in particular, will be a key part of the expansive leadership required to achieve peace, democracy, and sustainability. (Schwean, Bishop & McKeough, 2009, p.2)

To meet the significant challenge of nurturing leadership in our youth, a proliferation of programs has entered the market; however, most of these programs lack empirical validation, and there is very little evidence of their capacity to produce long-term effects (Allio, 2005). Thus, can we say with any confidence that these programs are effective? Unfortunately, one of the most significant factors contributing to this paucity of empirical data is the lack of psychometrically sound measures designed to evaluate the acquisition of those skills and traits associated with

positive leadership in youth. Indeed, a comprehensive review of the youth literature did not identify any psychometrically sound measures specifically assessing the factors identified with leadership in youth. The purpose of this study was to construct such a measure on youth leadership.

A University of Calgary Applied Psychological and Educational Services (U-CAPES) contract existed with the Werklund Foundation to (a) develop and test an instrument to evaluate youth leadership [The Youth Leadership Questionnaire (YLQ)] and (b) use that instrument to evaluate the Rapport Teen Leadership Breakthrough program (TLB). This dissertation encompassed part (a) and consisted of two components: (a) test development and (b) evaluating the psychometric properties of the YLQ using data from the TLB program. The major research questions were: (a) what domains are associated with youth leadership and should be included in the YLQ? (b) does the proposed three factor model (Self-Concept, Self-Efficacy, & Emotional Intelligence) provide a good fit? (c) what is the reliability of the YLQ? and (d) what is the degree of validity, particularly construct validity, of the YLQ?

The objective of this particular study was to develop a psychometrically sound measure of youth leadership that encompassed domains and competencies that have been shown to be both theoretically and empirically important to leadership. To identify what qualities, skills, characteristics, and other factors are important to leadership and inform scale development, literature related to the theoretical underpinnings of leadership was reviewed. Given that the population participating in the study were within the adolescent age range, it was important to briefly review the theoretical and research literature on youth leadership development to elucidate on those factors that may impact the expression of leadership during these years. Another purpose of the present study was to use the data from the TLB program to address issues

related to the factor structure, reliability, and validity, particularly construct validity, of the YLQ. Methodological literature related to test construction and the establishment of the psychometric properties of an instrument will be reviewed in a subsequent chapter.

Leadership

In general, leadership is defined as a social process that involves influencing and motivating the actions of others (Dickmann & Stanford-Blair, 2009). Leadership is a process/interaction between people, with the leaders being the agents of change who carry out the process and apply their knowledge and skills (Northouse, 2007). Leaders are often referred to as those individuals who have the ability to think for themselves, communicate their thoughts and feelings, and help others understand and act on their own thoughts, feelings, and beliefs (Halloran & Benton, 1987; van Linden & Fertman, 1998; Wallace, Foster, & DaCosta, 2007). However, given that leadership is a complex social phenomenon, efforts to understand leadership and how it develops have led to a wide array of conceptualizations and theoretical frameworks (McCauley, Drath, Palus, O'Connor, & Baker, 2006; Northouse, 2007; Wallace et al., 2007). A review of some of the past and current theories was conducted to inform model construction for the YLQ. The theories reviewed below do not exhaust the list of theoretical perspectives on leadership, but were believed to be some of the most closely linked/relevant theories to this study.

Theoretical models.

Trait theory. Trait theory is one of the earliest approaches to conceptualizing leadership. Within this theoretical framework, inherent individual characteristics and personality traits are hypothesized to differentiate leaders from nonleaders and account for effective leadership performance (Hollander, 1978; Lord, Brown, & Freiberg, 1999). A few of the reoccurring traits

that have been repeatedly identified as fundamental to leadership include intelligence/cognitive ability, persistence, and motivation (Kirkpatrick & Locke, 1991; Lord, DeVader, & Alliger, 1986; Stogdill, 1974). Research findings have reported that higher cognitive capabilities, including verbal and perceptual reasoning abilities, are positively correlated with leadership (Zaccaro, Kemp, & Bader, 2004). Moreover, studies have revealed that the Big Five personality factors (neuroticism, extraversion, openness, agreeableness, and conscientiousness) are related to leadership, with extraversion and conscientiousness being the most strongly correlated with the construct (Judge, Bono, Ilies, & Gerhardt, 2002).

Behavioural theory. Behavioural theories of leadership focus on the behaviours of the leader and imply that leaders are not born but made (Dickmann & Stanford-Blair, 2009; Lussier & Achua, 2004). Using an empirical framework, behavioural theories expanded the understanding of leadership by examining the relationship between particular behaviors related to perceived leadership and the antecedent positive or negative consequences associated with them. Results of this research contributed to the development of a broad conceptual map used to understand the complexities of leadership (Northouse, 2007). Through analyzing the relationships between the behavior and the response of the environment, individuals can enhance or nurture their leadership style. For example, leaders can learn about the effect of their behaviors by relating them to the tasks they are undertaking and through their interactions and relationships with others (Lussier & Achua, 2004). Although there is evidence to support the role of operant conditioning principles in the development of leadership behaviors, one criticism of behavioural theories is their lack of ability to identify universal behaviours or styles that are associated with and deemed to be critical to effective leadership (Northouse, 2007).

Skills theory. Skills theory is similar to trait theory in that it focuses primarily on the leader, but it differs from trait theory in that it views leadership not as innate characteristics but as abilities that can be learned and developed (Northouse, 2007). Moreover, skills theory is similar to behavioural models because they both recognize that leadership can be learned; however, it is distinct in that it focuses on the underlying or latent characteristics related to behavior in contrast to the particular behavior itself. Latent skills that have been frequently linked with leadership include self-concept, self-efficacy, and intra- and interpersonal competencies including emotional intelligence (Bass, 1990; Conner & Strobels, 2007; Judge et al., 2002). More contemporary skills models of leadership also take into account how individual attributes, such as intelligence, both cognitive and social, and environmental factors affect the development and expression of the aforementioned leadership skills (Mumford, Zaccaro, Harding, Jacobs, & Fleishman, 2000).

Constructivist developmental theory. To deepen understanding of leadership, theories of development need to be considered because leadership capacities unfold over time and are thought to be dependent on the developmental stage of the individual (Northouse, 2007; van Linden & Fertman, 1998). Constructive developmental theory was first proposed by Kegan in 1980 and identified important patterns in the manner in which individuals mature, such that earlier ways of thinking and meaning-making are integrated into more complex and comprehensive ways later on in life (Berg & Sternberg, 2003). With reference to leadership, these theories provide a description of how individuals develop more complex and comprehensive ways of thinking, acting, and making sense of themselves and their experiences across the lifespan (Hoare, 2006; McCauley et al., 2006; Thomas, 2005), and how these developments, in turn, influence the expression of leadership skills.

Transactional theory. Transactional theory of leadership involves an exchange between the leader and the followers, and focuses primarily on what leaders do. It involves skills or tasks, such as public speaking or decision making, in an attempt to motivate others and advance one's own as well as others' agendas (Holander, 1986). Reward and punishment are prominent ways to motivate people, and are thus, often used by transactional leaders to achieve their goals. Within transactional leadership, clear agreements and expectations are conveyed to the followers (Avolio, Bass, & Jung, 1999), and followers who comply with the leader receive praise or rewards in exchange for their agreement and performance (Bass, 1985). Leadership, from this perspective, requires making decisions, using standards and principles to guide decision making, and taking control of situations (van Linden & Fertman, 1998).

Transformational theory. Transformational theory, a more contemporary approach to understanding and conceptualizing leadership, refers to it as a process that transforms and changes individuals. It challenges leaders to be creative in their thinking (Avolio, 1999). Unlike transactional leadership which involves more external motivation and emphasizes the desire to satisfy needs, transformational leaders are more internally and self-motivated (Avolio & Bass, 1995; Howell & Avolio, 1993). More specifically, according to this framework, leadership involves engaging and connecting with others to raise the level of motivation, morality, empowerment, and idealism in both the leader and the followers (Northouse, 2007), which are all fundamental to successful leadership (Dickmann & Stanford-Blair, 2009). Transformational leadership involves believing in one's self and goals (Bass, Avolio, Jung, & Berson, 2003). Unlike some of the other aforementioned approaches, this model recognizes the important role of the situation in leadership. A leader's behaviour is embedded within a particular context, and this

context needs to be considered when looking at or evaluating leadership performance (Avolio & Bass, 1995).

Issues with the construct of leadership. Research has deepened and broadened individuals' understandings and insights into the complexity of leadership (Dickmann & Stanford-Blair, 2009). A range of theoretical perspectives is represented in the research on leadership (Wallace et al., 2007). Traditional approaches presume that leadership rests in individual characteristics, personalities, perceptions, and behaviours (Hollander, 1978; Lord et al., 1999). Other, more recent approaches include the social and relational processes underlying leadership (Dachler, & Hosking, 1995; Uhl-Bien, 2006). Moreover, some researchers propose that leadership encompasses self-reflection and self-knowledge (Mohamed & Wheeler, 2001; Nagle, 2003; van Linden & Fertman, 1998), while still others emphasize collaboration, mentorship, self-assertion, or decision-making (Des Marais, Yang, & Farzanehkia, 2000; McLaughlin, 2000; Woyach, 1996). Previous attempts have been made in Canada to develop more systemic, integrated, and comprehensive models of leadership, but to date, none have been very successful (Burger, 2007). As a result of theoretical heterogeneity, the construct of leadership has been defined and operationalized in multiple ways, resulting in some vexing research and interpretative issues with respect to youth leadership.

Conceptual definition of youth leadership. An essential component of instrument construction is operationally defining the construct to be measured, in this case, youth leadership. Based on the aforementioned theories and literature, the following definition of leadership is being proposed for this study: leadership is a dynamic social and developmental process whereby an individual influences others to achieve common goals in a particular setting or context (in this study, the setting is school environment). In addition, it includes the development and utilization

of a set of skills that can be learned, developed, and practiced in order to transform one's own and others' behaviours and thinking as well as guide performance (Adapted from *Leadership: Theory and practice*, by P.G. Northouse, 2007). Overall, this definition draws from several of the previously referenced theoretical models (skill, developmental, and transformational theories) and as such, recognizes that environment/context, development, and personal skills and attributes all contribute to youth leadership.

Conceptual model of youth leadership. The above definition can be further exemplified through a conceptual model. The model is an integrative one, insofar as, it not only examines the relationship between leaders' knowledge, skills, and performance, but it also incorporates the interaction of personal attributes, development, and environment (Northouse, 2007).

 Insert Figure 1 Here

This model contends that leadership outcomes are the result of an individual's competencies/skills, including self-efficacy and emotional intelligence. It also illustrates how personal attributes, such as self-concept, can influence the skills, as well as indirectly impact performance as a leader. A leader's knowledge and skills contribute to the process of leadership, while other attributes provide leaders with characteristics that make them unique and may or may not impact their performance. Lastly, the model demonstrates how environmental and developmental factors can play a role in the establishment and enhancement of leadership skills and one's performance as a leader. All situations/contexts are different and can influence how leaders respond and perform (Northouse, 2007).

Competencies and personal attributes, such as self-concept, self-efficacy, and emotional intelligence, are key factors that were associated with leadership in a previous small scale leadership study (Hindes, Thorne, Schwean, & McKeough, 2008), and have been repeatedly shown in the theoretical and research literature to account for and influence leadership performance (Bandura, 1997; Maurer, 2001; Northouse, 2007). In some cases, a reciprocal effect occurs. For example, self-concept not only positively impacts leadership development (Hindes et al., 2008; Lussier & Achua, 2004), but leadership competencies also contribute to addressing the needs, values, and aspirations of youth, as well as their self-assurance about their abilities (Lussier & Achua, 2004; Northouse, 2007). Further, strong relationships exist between self-efficacy, achievement, and positive organizational behaviors (Bandura, 1997; Stajkovic & Luthans, 1998). Research in this area has shown that the belief that one has the personal capabilities and resources to meet the demands of a specific task influences one's goals, aspirations, motivation, and effort (Maurer, 2001), and ultimately contributes to leadership capacities. Moreover, highly efficacious individuals attempt leadership roles at a significantly greater frequency than individuals with low self-efficacy (Chemers, Watson & May, 2000). Finally, a number of contemporary studies have shown strong associations between emotional intelligence, emotion regulation, social interactions, life satisfaction, and leadership (Ciarrochi, Chan, & Caputi, 2000; Edward & Warelow, 2005; Hindes et al., 2008). These studies have demonstrated that the ability to manage and regulate one's emotions is significantly related to leadership competencies (Chan, 2007; Lopes, Salovey, Cote, & Beers, 2005). Given the strong relationships of these competencies to leadership, as depicted in the research and literature, these skills will be used to form the item clusters for the YLQ. Overall, these domains reflect some of the skills/competencies and individual attributes indicated in the previous conceptual model. The

other areas of the model (developmental and environmental factors) were viewed as precursors to the development of leadership skills and leadership performance, and also were taken into consideration in the construction of the YLQ.

Leadership and Adolescence

The youth of today are the leaders of tomorrow and the future. Youth have the power to lead and the potential to enforce change. Our rapidly changing society needs young leaders who are able to address complex and pressing social issues. Moreover, our youth need to learn general life and leadership skills, as these skills are needed and are important in virtually all areas of adolescent and adult life (Astin & Astin, 2000). Adolescence is a period in life involving challenges, transitions, opportunities, and enrichment. Leadership development is important because it can empower youth and give them a greater sense of control over their lives (Astin & Astin, 2000; Dahl, 2004; Frisen, Jonsson, & Persson, 2007; Hotton & Haans, 2004). Given that the population participating in the study was youth, a brief review of the research literature on adolescence and leadership was conducted to examine and explain those factors that may impact the development and expression of leadership skills during these years of life.

Adolescence is a period of time in which youth may be vulnerable to various developmental risks, as it is a time in which they are experiencing changes in their physical development, maturity, sense of self, behaviours, and understanding of the world (Dahl, 2004). Youth are struggling to gain a sense of self and control over their lives. While the early childhood years are defined by an external locus of control (that is, they locate success and failure outside of themselves), the adolescent years represent a striving for self-determination (an internal locus of control) (Lussier & Achua, 2004). Additionally, youth are attempting to sort out their lives and futures as they prepare to transition towards independence and adulthood (Dahl,

2004; van Linden & Fertman, 1998). As a result, developing leadership during the adolescent years may not be viewed by youth as a priority in their own lives. But adolescence is, in fact, a critical time for youth to develop and enhance leadership skills because leadership helps them realize and reach their potential. Strong leadership skills enable youth to have a voice and an opinion about decisions that affect their lives. It is important for youth to feel empowered and in control of their lives because it gives them the confidence to make decisions that will influence their own lives and the lives of others (Astin & Astin, 2000; van Linden & Fertman, 1998).

In the past, ideas about leadership involving “great/powerful people” and leadership being “born not made” permeated the thinking of youth. These perceptions often resulted in youth not identifying themselves as leaders. When asked about leadership, youth frequently associated it with certain characteristics such as being male, older, self-confident, reliable, wealthy, popular, intelligent, and having a desire to excel (van Linden & Fertman, 1998). Although some of these characteristics may play a role in leadership, they are not the sole determinants of leadership ability and effective performance (Northouse, 2007). Overall, youth tended to narrowly define leadership and frequently viewed it as something that was formal and difficult to attain. Youth will implicitly generate their conceptions of leadership from what is taught intentionally and unintentionally to them across the educational experience, and they will find it difficult to lead until they have experienced effective leadership as part of their education. Youth need to be taught and shown that leadership is no longer the province of the few, the privileged, or even the merely ambitious (Astin & Astin, 2000).

Some evidence exists that leadership in youth and adults is different, with today’s youth putting more emphasis on personal and relational aspects of leadership, rather than on task and systematic components (Dempster & Lizzio, 2007; Roach, 1999). However, other research

suggests that this divergence in conceptualization of leadership may not be as great as previously believed (Schneider, Ehrhart & Ehrhart, 2002). Research has shown that leadership skills develop and increase during post-secondary education (Pascarella & Terenzini, 2005), suggesting that the development of leadership skills, such as self-awareness, self-concept, and emotional intelligence, does not end at adolescents and extends into young adulthood.

Youth leadership programs.

Despite the inconsistency in defining leadership, an increasing number of youth leadership programs are emerging (Day, 2000; Day & O'Connor, 2003). One emerging youth leadership program is the TLB program that was developed by Rapport Leadership International. The TLB program engages youth in activities that focus on goal setting, self-awareness and reflection, teamwork, giving and receiving feedback, and public speaking. A small program evaluation study on the TLB program was previously completed, and it helped set the framework/context for this study (for more information on the TLB program and this previous study, refer to Hindes et al., 2008). Overall, the goals, processes, and activities of the TLB program focus on several competencies and domains that have been linked to youth development and leadership (Rapport Leadership International, N.D.), including those domains that were examined in this study.

Summary

Leadership is a complex construct. Most researchers generally agree that leadership is dependent on social, emotional, and intra- and inter-personal competencies including self-concept, self-efficacy, and emotional intelligence (Bass, 1990; Conner & Strobels, 2007; Judge et al., 2002; van Linden & Fertman, 1998; Zaccaro, Gilbert, Thor, & Mumford, 1991). These skills enable and influence youth's performance as leaders and have been associated with positive

outcomes (Lussier & Achua, 2004; Roth & Brooks-Gunn, 1998; Stajkovic & Luthans, 1998). The development and expression of leadership skills in youth are also influenced by personal attributes and context (Dickmann & Stanford-Blair, 2009; van Linden & Fertman, 1998; Zaccaro, 2007). Individual attributes may be viewed as precursors to effective leadership performance and can support youth as they apply their leadership competencies. Leadership is reflective of the people, situations/context, and tasks involved (Zaccaro, 2007). Consequently, a youth who is a leader in one context may not be a leader in another situation.

The current study proposed to develop a measure of youth leadership (YLQ) that encompassed domains and competencies that have been shown to be both theoretically and empirically important to leadership. Over the years, schools have become increasingly more involved in the promotion of youth mental health and positive well-being (Webber, & Mulford, 2007). In this study, the research and data were based on the Rapport TLB program, which was delivered within the school context, and hence examined leadership from that perspective. Thus, because the YLQ was constructed and evaluated within a school context, it should be considered from an educational perspective. Examination of the factor structure, reliability, and validity was also included in order to establish preliminary psychometric properties of the YLQ.

Chapter 2- Literature Review

Test Construction

Tests are pervasive in our world, and they play a significant role in all of our lives. Despite an increased awareness of the importance and need for valid assessments of cognitive, academic, social, emotional, and other life skills (McIntire & Miller, 2007; Messick, 1980; 1989), currently there are no known instruments that can be used to specifically evaluate leadership skills in youth. Therefore, measures of leadership need to be created and validated. To inform the development of the Youth Leadership Questionnaire (YLQ), this chapter reviews the literature on test construction and outlines the process of developing and validating a test.

Past strategies for the development of tests for children and adolescents gave precedence to developmental changes over theory (Anastasi, 1986). Tests were essentially designed to reflect the advances that children and youth made in the acquisition of certain skills and knowledge as a function of maturity. However, a shift towards a theoretical approach to test construction emerged (Loevinger, 1957), requiring researchers and test developers to gain a more articulated understanding of the interrelated theoretical concepts before engaging in scale development. Examining theory helps researchers to better conceptualize and define the construct of interest, including not only what it is, but also what it is not. This approach allows for a more precise and detailed account of the nature and range of the construct and its theoretical context (Clark & Watson, 1995). Contemporary literature advocates for an integration of both empirical data and theory because using both provides more total validity evidence for the newly formed measurement instrument (Anastasi, 1986). Inclusion of theory and statistical analyses in test construction, especially in item development, is important because it enhances the understanding

of the nature and structure of the test, its domains, and its items. Empirical data can also help identify which items are weak and should be eliminated (Clark & Watson, 1995).

Stages of test construction. Test construction is a multi-step procedure that initially begins with the identification, definition, and operationalization of the construct of interest (Clark & Watson, 1995; Dawis, 1987). Conception of the construct stems from the theoretical frameworks and orientations of the test developer, as well as from a thorough review of the literature and research on the construct and other areas associated with it (Anastasi, 1986; Clark & Watson, 1995; Dawis, 1987). The literature review not only provides information on the construct, but it also identifies the importance and need for developing a test to measure the proposed construct (Clark & Watson, 1995). After the scope and purpose of the test have been determined, selection of the scaling method and construction of the items can be completed.

Choosing the rating scale. Likert scales are commonly used to measure the frequency, degree, or agreement of respondents (Clark & Watson, 1995). There is debate as to whether or not Likert scales should consist of an even or an odd number of ratings. Using an odd number of ratings with a mid-point option allows respondents to have an intermediate or neutral opinion when they may truly not have an opinion or are unsure about the subject. However, one disadvantage of utilizing an odd scale is that the middle rating can reflect an uncertainty in responding. Some researchers prefer respondents to make a definite choice rather than choose neutral or intermediate positions on a scale (Garland, 1991). For this reason, a scale without a mid-point may be preferable. Although an even scale eliminates the aforementioned issues by forcing respondents to choose a side, this forced decision also presents a problem worthy of consideration as it may be seen as objectionable by some respondents. Using an even scale that forces respondents to choose an opinion or rating can bias the results (Tull & Hawkins, 1993).

Research has shown that when mid-points are removed from scales, respondents are pushed more to the extreme ends of the scale (Garland, 1991; Worcester & Burns, 1975).

Debate also exists around the effects of the number of response categories on the reliability and validity of the measure. Some researchers have demonstrated that the number of response categories has little impact on the reliability and degree of validity (Aiken, 1983; Clark & Watson, 1995). On the other hand, there is some evidence that the more scale points used, the more reliable the scale (Churchill & Peter, 1984). By increasing the number of response items the researcher/examiner can impact the likelihood of detecting a true change because the intervals between the points on a scale get smaller as the number of response options gets larger; thus, making it easier to find a statistical difference. However, having a highly sensitive scale can result in small effect sizes and interpretations that may only be due to measurement error (Friedman, Wilamowsky, & Friedman, 1981).

Overall, the debate surrounding the number of response options continues, and the answer is largely one of individual researcher preference. The most frequently used and recommended Likert scales consist of five to seven response options (Cox, 1980; Friedman & Friedman, 1986; Lehman & Hulbert, 1972). For the purpose of this study, an odd scale with a mid-point option was utilized because some respondents may legitimately have an intermediate level of skills, and using a scale without a mid-point could introduce respondent bias because respondents are forced to choose a more positive or negative response (Tull & Hawkins, 1993). In addition, a seven point scale was chosen because it provides sufficient sensitivity and decreases the reliance on the mid-point option. Research has shown that as the number of response options increases, respondents' use of the mid-point category decreases. For example, on three and five point scales, an average of twenty percent of respondents chose the mid-point

category; whereas this category was only utilized by an average of seven percent of respondents when a seven point scale was used (Matell & Jacoby, 1972).

Item development and selection. Item selection and writing can be a lengthy and on-going process. The initial pool of items should be broad and include those items that may be related, as well as potentially unrelated, to the construct (Zeidner & Most, 1992). The number of items must adequately sample the construct as well as the domains within the construct. Both theoretical and psychometric techniques should be used to identify those items that are and are not related to the construct, as well as to determine if there are a sufficient number of items (Clark & Watson, 1995). For multidimensional constructs, the sub-domains do not have to be entirely unrelated; however, too much overlap can be problematic and compromise inferences (Zeidner & Most, 1992). Therefore, careful consideration as to whether or not the construct is multidimensional and if the sub-domains are distinct is required.

Factor analytic techniques are used most frequently to determine if a construct is multidimensional and if subscales/sub-domains are desired (Clark & Watson, 1995). Factor analysis determines which items “hang together” and how many factors provide the best fit for the data/test items (Gregory, 2004; Zumbo, 2007). If the majority of the items all load on one factor, then the scale may be best described as unidimensional; however, if items load on different factors, the scale likely consists of sub-domains (Zumbo, 2007). Given that factor analysis is commonly used in item selection and scale development, it was the method chosen for this study.

Item analysis identifies those items that are internally consistent (measuring the same construct) and inter-correlated (related to other items). Correlations can be used to identify the relationship among the items (Zumbo, 2007). Items within one sub-domain should be more

correlated/related to one another than with items in other sub-domains. If all the items are inter-correlated, the scale is homogeneous or unidimensional (Clark & Watson, 1995). Those items that are not highly inter-correlated with other items should be re-worded or eliminated from the test (Zumbo, 2007). In this study, both factor analysis and correlations were used in the development, writing, and selection of the items for the YLQ.

Items included in a test should be written in a language that is simple, non-ambiguous, understandable, and appropriate for the target population (Clark & Watson, 1995). For example, colloquialisms, trendy expressions, and double barreled questions (e.g., is this tool interesting and useful?) should be avoided. Moreover, the items should be as free of bias as possible to avoid social desirability (most respondents choose the most positive rating or the one that they perceive the examiner wants them to choose), which can threaten the degree of validity. These types of items are undesirable because they convey little information and will usually correlate weakly with other items (Clark & Watson, 1995; Sireci, 1998). Items that result in social desirability should be removed. Examination of the shape of the distributions for each of the items can be used to identify socially desirable items as the distributions would not be normally distributed; instead they would be skewed. Moreover, combining statistical methods and theory, and careful review of the wording of the items with a group of experts in the area can help identify problematic or socially desirable items (Anastasi, 1986; Clark & Watson, 1995). Questions that virtually everyone or no one will endorse should only be included if you want some items to assess invalid responding (Clark & Watson, 1995).

Pilot studies on convenient samples, such as university students, are frequently conducted prior to the actual data collection to assist in item analysis (Anastasi, 1986; Clark & Watson, 1995). Subsequent analysis of the data from the pilot study examines the extent to which the

scale and its items performed as expected under practical conditions. Doing some preliminary empirical analyses on the structure and content of the scale provides an opportunity to identify questionable items that were not originally identified or recognized in the process of item writing. Additionally, it suggests the number of test items for the final version (Izard, 2005). Overall, pilot testing assists in determining the precision and structure of the test.

For this study, a pilot study involving the first 67 participants in the study was conducted. Basic descriptive (frequencies, means) analyses as well as factor analysis and correlations on the pilot data were computed to gain an understanding of the pattern of responding on each of the items. A review of the graphs that depicted the shape of the distributions was also conducted. Further, the items were reviewed by other professors to identify any items that were poorly worded or possibly problematic. Lastly, correlations and factor analyses were used to examine the relationship between the items and the structure of the YLQ.

Validity

One of the most important psychometric properties of a test that needs to be considered during test construction is validity (Messick, 1980; 1989), and therefore, a review of the models of validity is provided. In general terms, validity refers to the scientific inquiry into the meaning of test scores (Messick, 1989). Over the years, the concept of validity has continued to evolve, and there has been much debate and re-conceptualization of this concept.

Models of validity.

Criterion model. In general, the criterion model defines validity in terms of how well the test provides an estimation of an individual's performance on another outcome, either concurrently or in the future. Criterion validity, therefore, involves the comparison of test scores with one or more external variables or criteria that are known or believed to measure the attribute

under study (Gregory, 2004; Kane, 2001; Kerlinger & Lee, 2000). Criterion validity can be further broken down into two subtypes, concurrent and predictive validity (Angoff, 1988).

Concurrent validity is focused on the relationship between the scores on the test and the criterion measure, which are taken at the same time (Kane, 2006; Suen, 1990). Predictive validity, on the other hand, is the extent to which the test scores can be used to predict the scores from a criterion measurement procedure that will take place at some point in time in the future (Kerlinger & Lee, 2000; Suen, 1990).

In more operational terms, the criterion model describes validity as the correlation between scores on a test and some other objective measure or criterion that the test is used to evaluate (Angoff, 1988; Kane, 2006). Gathering evidence for criterion validity is generally a statistical process, and the degree of criterion validity is indicated by the strength of the correlation or relationship between the scores on a test and a chosen criterion measure of performance. The strength of the relationship is usually expressed and represented in the form of a Pearson's correlation coefficient (Suen, 1990). During the early part of the century, psychologists and test developers used correlations to learn about their tests, and they focused on the convergence between measures as evidence of validity. A test could, therefore, be valid if it was correlated with another construct tapped by a different measure (Angoff, 1988; Kane, 2001). Current conceptualizations of validity suggest that criterion validity is used to support the appropriate use of scores from the test as an indicator of performance on some criterion measure (Kane, 2001; Suen, 1990).

Content model. Content validity has its own history, but is most often associated with the tradition of achievement testing (Kerlinger & Lee, 2000). Content validity refers to the relevance and representativeness or sampling adequacy of the test items or topic of a measuring instrument.

In general, content validity is required for tests whose purpose is to describe an individual's estimated overall level of skill or performance in that subject or area (Gregory, 2004; Sireci, 1998; Suen, 1990). At one time, test developers believed that a domain could be defined so exactly that items could be sampled literally at random from it (Angoff, 1988; Shepard, 1993). Tests are composed of a limited number of items that are often thought to exhaust or sample the subject matter adequately. However, in the contemporary world, it has been determined that this exact defining and random sampling does not work, especially with more complex content domains because rarely do the items of a test actually exhaust the universe of all possible items, nor are items normally drawn randomly from some universe of items (Angoff, 1988; Kerlinger & Lee, 2000). Content test validation, therefore, seeks to answer the question: is the content of the measure and test items representative of the construct of interest that the test is purporting to measure? A clearly defined and detailed description of the content to be measured is important because it allows test developers to determine whether each test item lies within the boundaries of the universe of content (Kerlinger & Lee, 2000; Murphy & Davidshofer, 2005).

The evaluation of content validity requires determining and establishing whether the test items and the domains being sampled are appropriate and reasonable. Evidence of content validity is demonstrated when the test items form a demonstrative sample of the appropriate domain and are deemed to be relevant to the intended use of the score (Kerlinger & Lee, 2000). A test score can be considered relevant for its intended use when all items in the test are shown to be within the domain of interest (Murphy & Davidshofer, 2005; Shepard, 1993). This process of judgment can be formal or informal. The least formal method involves establishing a casual overall impression of whether or not a test appears to contain the appropriate items. This form of validity evidence is often referred to as face validity and is generally considered an unacceptable

approach to determining validity (Suen, 1990). Face validity has an inherent ambiguity and, as a result, test developers should regard it as a concession to gain acceptability rather than a serious psychometric effort to support validity (Angoff, 1988). A more formal and logical judgmental process that is frequently used today requires experts in the area judging whether the items of the test adequately sample the intended domains and universe, and whether different formats or modes of assessment might alter the content or construct meanings (Shepard, 1993). Each test item is weighed by experts for its presumed representativeness of the domain of interest (Kerlinger & Lee, 2000; Suen, 1990).

Construct model. Historically, construct validity was to be utilized whenever content or criterion validity were not appropriate or sufficient (Kane, 2001; Murphy & Davidshofer, 2005). One aspect that distinguishes construct validity from the other two approaches is its focus on theory and scientific empirical inquiry, as well as its emphasis on the testing of hypothesized relations and inferences. Construct validity was the first approach to test validation that linked psychometric notions and practices to theoretical variables (Kerlinger & Lee, 2000). Contemporary definitions of construct validity define it as the extent to which a test measures or correlates with a theoretically-based psychological construct (Kane, 2006). Moreover, it refers to the degree to which inferences can legitimately be made from the test scores to the theoretical constructs on which those test items are based. Construct validity has an advantage over the other validity models in that when construct validity is supported, inferences and conclusions can be made based on test score interpretation (Suen, 1990).

Construct validity can be further broken down into two subtypes- convergent and divergent validity. Convergent validity is used to demonstrate that tests which measure constructs that are hypothetically and theoretically interconnected to one another are actually

related to one another. Divergent validity, on the other hand, establishes that tests of hypothetically and theoretically unrelated constructs are not related to one another (Kerlinger & Lee, 2000). Through divergent validity evidence, test developers are able to empirically differentiate the construct of interest from other constructs.

To statistically estimate the convergent and divergent validity of two or more measures, correlation coefficients are commonly used because these relationships provide the patterns of intercorrelations among the measures (Kane, 2006; Zumbo, 2007). High correlations suggest that the measures are similar, and low correlations indicate that the measures are dissimilar. Correlational analyses among tests that support divergent validity are weakly, or not at all, correlated with one another (Kerlinger & Lee, 2000). However, a common problem encountered is what constitutes a high and low correlation. Generally, a correlation coefficient ranging from zero to about .20 is regarded as indicating no or negligible correlation; .20 to .40 is suggestive of a low degree of correlation; .40 to .60 is indicative of a moderate degree of correlation; .60 to .80 reveals a marked degree of correlation; and, .80 to 1.00 may be regarded as indicating high correlation (Christensen, 2001).

In addition to correlations, construct validity can also be determined by the homogeneity of the test items, theoretically consistent outcomes, and factor analysis of scores on test items to determine the extent to which of the items “hang together” or share common variance (Gregory, 2004; Zumbo, 2007). Factor analysis, a commonly used approach for test item selection and validation, determines how many factors are required to account for the inter-correlations of the test items (Zumbo, 2007). It is useful when the construct being examined is multi-dimensional; however, one disadvantage of factor analysis is it requires a large sample (Clark & Watson, 1995; Gregory, 2004). Some populations are often difficult to sample because they are not

readily available, are low incidence conditions, or are difficult to reach (e.g., individuals with rare disabling disorders).

Current notion of validity. In the past, the focus of test developers was on validating a test; however, today, it is evident that it is not the test that has validity, but rather the inferences and conclusions that test users make from the test scores (Angoff, 1988; Murphy & Davidshofer, 2005; Shepard, 1993; Sireci, 2007). Validity is, therefore, not a property of the test, but a function of what the test scores mean (Kane, 2006). Current notions of validity propose that it should focus on the meaningfulness and usefulness of the specific assumptions drawn from the test scores. It is, therefore, considered to be the overall evaluative judgment of the degree to which empirical evidence and theoretical rationale support the adequacy and appropriateness of inferences based on test scores and other modes of assessment (Messick, 1980; 1989).

Today, construct validity is now viewed as the one unifying and overarching framework for conceptualizing validity evaluations. Consequently, content and criterion validity are conceptualized as limited aspects of construct validity (Murphy & Davidshofer, 2005; Shepard, 1993; Sireci, 2007). Along with the unified perspective of validity, the current definitions also suggest that validity evidence should be continually collected over time, either enhancing or contradicting previous findings. As a result, test validation is a process of accumulating evidence and integrating statistical analyses with theory to support any particular inference or conclusion. Claims and interpretations need to be continually evaluated and defended because over time new evidence may be revealed that contradicts previous findings and inferences (Gregory, 2004; Kane, 2006; Murphy & Davidshofer, 2005; Zumbo, 2007).

Overall, current views of validity have departed from the static perspective of the past to a broader and more comprehensive approach (Cronbach, 1980; Shepard, 1993; Suen, 1990). It

has become increasingly clear that there is no such thing as a generically valid test, and it is impossible not only to prove that a test is valid but even that a test is measuring what test developers contend it is measuring (Kane, 2006). Consequently, it is now widely accepted that it is not the test that is being validated, but rather the inferences drawn from the test scores (Angoff, 1988; Murphy & Davidshofer, 2005; Shepard, 1993; Sireci, 2007).

In relation to this study, validity evidence was collected over time through different means, with an emphasis on the degree of construct validity of the YLQ. Factor analyses and correlations with other selected measures were computed on the YLQ in order to provide preliminary evidence of the degree of validity of the conclusions/inferences that can be drawn from participants' scores on the YLQ.

Reliability

Another important characteristic and property of tests that needs to be reviewed in relation to test construction is reliability (McIntire & Miller, 2007). Reliability refers to the consistency and stability of responses to the test items (Christensen, 2001). Reliability is important because it provides critical information on the internal consistency of the items, the homogeneity of the construct, and the error variance (DeVellis, 1991). Reliability can be evaluated by a) creating two parallel forms or alternate forms of the test; b) retesting respondents on the same test (test-retest reliability); c) dividing the items into two halves; or, d) using measures of internal consistency. For each of these methods, a correlation coefficient will determine the reliability of the test (Christensen, 2001). The same general guidelines for correlational coefficients, as noted above in the validity section, are used for determining the degree/strength of reliability; however, the standards of what constitutes an acceptable level of reliability are based on subjective decisions. For this study, a reliability correlation coefficient of

.70 or higher will be considered acceptable because correlations of .70 to 1.00 are often regarded as being respectable to very good (DeVellis, 1991).

Test-retest reliability was one of the methods chosen for this study, as it is important to ensure that constructs such as leadership remain stable over time when no other additional intervention or training is provided. Test-retest reliability was evaluated when the same respondents completed the same test at two different time periods. The assumption underlying this approach is that the construct being measured does not change between the two time periods. The more similar the results across the multiple time periods, the greater the test-retest reliability. To estimate the closeness of the results, correlation coefficients are used because these coefficients demonstrate the degree to which the responses from one time period are related to those from the second time period (DeVellis, 1991). When using test-retest reliability and examining the resulting coefficients, it is important to take into account the time gap between the testing periods. Shorter time gaps typically result in higher correlations, and as more time elapses between the testing periods, the correlations weaken (DeVellis, 1991; McIntire & Miller, 2007). One problem with this measure of reliability that needs to be considered is test effects. Test effects occur when respondents' answers to the questions the first time affect their responses on subsequent time periods (Christensen, 2001).

One of the more common measures of reliability is internal consistency. Internal consistency was also included within this study because although test-retest reliability is important for determining replicability of scores, it does not measure the relation of the test score to the true score. Consequently, internal consistency was utilized to determine whether several items that propose to measure the same general construct produce similar scores (Christensen, 2001). Internal consistency is also central to the theory of measurement error- the higher the

reliability, the smaller the error, and thus, greater accuracy of the relation of the test score to the true score (Zumbo, Gelin, & Hubley, 2002). Internal consistency entails statistically estimating how well the items that reflect the same construct yield similar results. Thus, it focuses on the consistency of the results for different items for the same construct within the measure (Christensen, 2001; Zumbo et al., 2002).

There are several methods for investigating internal consistency; however, one of the most common, and the one that was utilized in this study, is Coefficient Alpha (Zumbo et al., 2002). Coefficient Alpha involves a mathematical computation of all the possible split-half estimates. It compares the variance of all the items together to the variance of the individual items (Christensen, 2001; Zumbo et al., 2002). An alpha level of at least .70 is considered acceptable (DeVellis, 1991). Low internal consistency can be a result of poor defining and operationalizing of the construct, having a multi-dimensional construct instead of a unidimensional construct, or consisting of poor quality items (DeVellis, 1991; Zeidner & Most, 1992).

In relation to this study, a measure of the stability of the results and the internal consistency of the test was calculated because both of these aspects are important for good and psychometrically sound tests. Stability was measured using test retest reliability to determine the extent to which leadership remains stable in youth over time. Coefficient Alpha was used for calculating the internal consistency because it is one of the more common and widely accepted methods used in test construction.

Summary

The construction and validation of the YLQ was an ongoing process that was based upon both theoretical knowledge and empirical data. This process consisted of a literature review,

operationalization/conceptualization of the construct of leadership, a pilot study, statistical analyses (factor analysis, correlations, and Coefficient Alpha) on the items, and item revisions (re-wording, additions, and deletions). The version of the YLQ that resulted from this process was evaluated in terms of its reliability and validity.

Whether a test is being used for research, clinical, educational or practical purposes, the validity and reliability of it are important (McIntire & Miller, 2007; Messick, 1980; 1989). Just as we would not use a reading test to assess math skills, we would not want to use the YLQ if it was not accurately and consistently measuring youth leadership. With the awareness of the importance of psychometrically sound instruments, key properties focusing on consistency and accuracy (reliability and validity) were evaluated using data from the Rapport Teen Leadership Breakthrough (TLB) program. As validation is a continual process that should occur over time (Gregory, 2004; Kane, 2006; Murphy & Davidshofer, 2005; Zumbo, 2007), preliminary insight into the psychometric properties was obtained through factor analyses, reliability estimates, and correlations. Since reliability is necessary, but not sufficient for validity it was important to assess both of these properties of the YLQ (McIntire & Miller, 2007; Messick, 1980; 1989). Construct validity was of primary interest as it is deemed by many to be the overarching framework for establishing validity (Murphy & Davidshofer, 2005; Shepard, 1993; Sireci, 2007).

Chapter Three- Methods

Objectives of the Current Study

This study primarily explored the skills aspect of the previously presented model because these competencies can be measured, taught, and learned, and they are fundamental to leadership and performance. Specifically, the study focused on emotional intelligence, self-efficacy, and self-concept in youth to deepen understanding of leadership development. These concepts are reflective of the competencies and skills indicated in the previous conceptual model. Self-efficacy, emotional intelligence, and self-concept were chosen for the Youth Leadership Questionnaire (YLQ) because these domains have been repeatedly shown to be associated with leadership and encompass many of the behaviours and abilities that are being taught in current leadership programs (Bandura, 1997; Maurer, 2001; Northouse, 2007). In addition, results from a previous small scale leadership study demonstrated that emotional intelligence and self-concept were empirically related to leadership (Hindes et al., 2008). The other areas of the model (individual attributes, developmental, and environmental factors) were viewed as precursors to the development of leadership skills and leadership performance. Thus, these factors also were taken into consideration in the construction of the YLQ.

Data from the Rapport Teen Leadership Breakthrough (TLB) program was gathered to assess the psychometric properties of the YLQ. Having a well-developed and psychometrically sound test is important because it provides more confidence and assurance that the results and interpretations are consistent and accurate (McIntire & Miller, 2007). In this research the reliability and degree of validity, particularly construct validity, of the YLQ were investigated.

Methods

Participants. Six hundred and ninety four Calgary Board of Education (CBE) students (286 males, 408 females, mean age of 16.06 years, $SD= 1.04$) from each quadrant of the city, enrolled in grades ten ($N= 301$), eleven ($N= 186$), and twelve ($N= 207$), participated in this study. It is important to note that the exact manner in which the schools and participants were chosen remains unknown because the schools that were involved in the study were those schools that had been previously selected by the CBE to participate in this project and in the Rapport TLB leadership training. In addition, the students were chosen by various school personnel and not the primary researcher.

Four hundred sixty-six of these students participated in the TLB program (188 males, 228 females, mean age of 15.94 years, $SD= 1.02$). Another 228 students, who did not receive the TLB training, served as a control group (98 males, 130 females, mean age of 16.31 years, $SD= 1.04$). The first sixty seven participants were used as part of a pilot study for the development of the YLQ. Two other participants were further excluded as a result of completing the survey multiple times. Overall, 625 students participated in the study.

In addition, a separate sample of 73 University of Calgary (UofC) students (18 males and 55 females with a mean age of 21.44 years and $SD= 4.77$) who had a mean of 2.71 years of post-secondary education ($SD= 1.67$) were also used in this study for test validation purposes. These participants were recruited from a convenient sample of students in undergraduate level courses at the UofC. Seventy-one participants were fulltime students and two participants were part-time students from various undergraduate programs (See Table 1). None of these participants participated in the TLB program, but they completed the YLQ and additional questionnaires (Satisfaction With Life Scale and the Resiliency Scale for Children and Adolescents- Adult

Revision) to evaluate the degree of construct validity of the YLQ. These participants were included because the aforementioned CBE participants were unable to complete the additional questionnaires that were part of the validation study for evaluating the YLQ's degree of construct validity.

Insert Table 1 Here

All participants completed an informed consent form, and for those who were under the age of 18 years, parental consent was required. Participation by students was completely voluntary. Refusing to participate or withdrawing from the study did not, in any way, affect the student or the services he/she received from the CBE or the UofC, nor did it bring with it any penalty. Participants were not required to put their names on the online measure and therefore, their privacy and anonymous identity was ensured. Participant numbers, which were randomly generated, were used instead of names. The master list, linking student names to identity numbers, was stored securely at the CBE and the UofC, and was accessible only to authorized CBE personnel and the primary UofC researcher. Once all information was gathered, the master list linking names to identity numbers was destroyed. All information gathered from students, now identified by identification number only, was kept in a locked file cabinet in a secure office at the UofC. The computer files were kept only on the primary researcher's personal computer, which was password protected. All data will be destroyed about five years after completion of the study (by September 30, 2015).

Measures.

The Youth Leadership Questionnaire (YLQ). The YLQ is the youth leadership questionnaire that was developed for the purpose of this dissertation. It yields a total composite score, as well as three domain scores associated with leadership: a) Self-Concept (SC) b) Emotional Intelligence (EI) and c) Self-Efficacy (SE). It also consists of demographic information (e.g., grade, age, gender, and school attended) to provide insight into the population being used for validating the test. The YLQ takes about 10-15 minutes to complete, and respondents' answers are based on a seven point Likert scale (Never, Very rarely, Seldom, Sometimes, Frequently, Most Often, Always). Each of the items is rated on the seven point Likert scale, where 1 = Never, 2 = Very rarely, 3 = Seldom, 4 = Sometimes, 5 = Frequently, 6 = Most often, and 7 = Always. To calculate the total leadership score, the ratings across all items are added together for each participant. To calculate the scores for each of the other domains/factors, the ratings for the items that comprise each of the domains are added together. The reliability and validity of the YLQ were determined as part of the study and are discussed below in the results section. Appendix A includes the items of the YLQ.

The Resiliency Scale for Children and Adolescents-Adult Revision (RSCA-AR). The RSCA-AR was chosen as a validity measure because resiliency models have been used to predict successful youth transitions into adulthood, as well as have been associated with problem-solving abilities, optimism, and social support, which are all factors that can be linked with leadership (Chang, 2001; Masten, Burt, Roisman, Obradovic, Long, & Tellegen, 2004; Roth & Brooks-Gunn, 1998; Tusaie, Puskar, & Sereika, 2007). It was hypothesized that scores on the

YLQ would be positively correlated with the RSCA-AR, providing evidence of construct validity.

The RSCA-AR represents a scientifically grounded and reasoned self-report approach to evaluating resiliency. It is a 64 item questionnaire that taps three major resiliency dimensions (Sense of Mastery, Sense of Relatedness, and Emotional Reactivity). The Sense of Mastery scale provides information on an individual's optimism about life and one's own competence, self-efficacy related to developing problem-solving attitudes and strategies, and adaptability. The Sense of Relatedness scale is included to assess an individual's degree of trust in others, support from others, comfort with others, and tolerance with others. The Emotional Reactivity scale focuses on sensitivity, recovery, and impairment. These scales provide a framework for understanding processes and factors that offset the negative effects of adversity and increase the probability of positive development. Respondents' answers to the questions are based on a five point Likert scale (Never, Rarely, Sometimes, Often, Almost Always) (Prince-Embury, 2006).

Overall, the RSCA-AR is designed to identify areas of perceived strength and/or vulnerability. It is psychometrically sound with Coefficient Alpha coefficients ranging from .93 to .95, indicating exceptional internal consistency. The standard error of measurement (ranging from .90 to 2.45) and test-retest reliability (ranging from .70 to .92) are good to excellent, suggesting acceptable reliability. Validity correlations for the internal structure indicate that the scales are significantly related with each other, but they are also distinct and represent different components/subscales of resiliency (Masten, 2001; Prince-Embury, 2006; Rutter, 1987).

The Satisfaction With Life Scale (SWLS). The SWLS was chosen as a second validity measure because previous research has shown that leadership is important for youth's

satisfaction (Scheer, 1997). It was hypothesized that life satisfaction would be positively correlated with leadership skills.

The SWLS, which is a five-item self-report instrument, was developed to assess an individual's satisfaction with life as a whole. The participants are required to integrate and weigh life domains, such as health or finances, in whatever way they choose. The SWLS employs a seven-point Likert rating system, with responses ranging from "Strongly Agree" to "Strongly Disagree". Scores on the SWLS may be interpreted in terms of absolute as well as relative life satisfaction. The SWLS has favourable psychometric properties with a test-retest correlation coefficient of .82 and a Coefficient Alpha of .87 (Pavot & Diener, 1993).

Procedure.

Stage one: Test construction. The study began with the identification and operationalization of constructs and related questions associated with youth leadership. Findings from a previous small scale leadership study (Hindes et al., 2008) were used to provide information on the areas/skills that were empirically associated with leadership. Factor analyses and empirical data from the previous leadership study were used to determine which domains from various questionnaires (The Behaviour Assessment System for Children, The Bar-On Emotional Quotient Inventory, The Resiliency Scales for Adolescents, The Family Assessment Scale, and The Multidimensional Self-Concept Scale) were empirically related to the construct of youth leadership. Specifically, the previous study found that emotional intelligence and self-concept could be improved in youth who attended a leadership training program (TLB program), suggesting that these two areas are empirically related to leadership.

In addition, a literature review of the conceptualization of leadership, as well as a thematic review of leadership programs, was conducted to provide information on those competencies and skills that are theoretically important to youth leadership. This theoretical and conceptual information defined the content domain of leadership and was used to help narrow down the initial item pool to encompass only those domains that have been shown to be theoretically important to youth leadership development.

Items were developed based on data from the previous study and the theoretical background that was discovered via the literature review. The information and empirical data gained from the previous study was used to help define the domains/factors for item creation, guide the writing of the items, and reduce the number of possible test items. Theory and research has shown that self-efficacy, self-concept, and emotional intelligence are associated with leadership (Hindes et al., 2008; Maurer, 2001). Thus, items that in principle measured the domains of interest were re-written and included in the initial item pool. The items that had high factor loadings and were correlated with one another, but were for the most part independent of the other items, were combined into different factors. Those items with low factor loadings were eliminated. The items had to be correlated and theoretically related to be included in the scale. Therefore, the theoretical domains remained intact during the empirical evaluation.

Overall, the information gathered from the statistical analyses and the literature review were used to select the number of possible domains that were central to leadership and which test items were to be included in the YLQ. In addition, those items that were not empirically and theoretically related to leadership were eliminated from the item pool. Statistical analyses and the elimination process were continued until a specified number of items (25) remained.

Stage two: Pilot study. After the final items were selected, a pilot study was completed to assist in finalizing the item selection. The first 67 students who participated in the study served as the participants for the pilot study. Basic descriptive analyses, factor analyses, correlations, and visual inspection were conducted on the pilot data to determine whether the items were continuing to show empirical and theoretical significance to youth leadership. Exploratory factor analysis was used on the pilot data to demonstrate that the items in each domain grouped together (intercorrelate) in the way that theory and research would predict. Following analysis of this pilot data, revisions, additions, and deletions were made to the items of the questionnaire (see Appendix A for the initial and revised scale). In total, twelve items were retained or reworded to improve the meaning and clarity of these items. Thirteen items were deleted because they were viewed as being poorly worded, potentially problematic, or not as significantly related to the theory and research on leadership. Thirteen new items that were believed to be of better quality and more associated and representative of leadership were added.

Stage three: Test validation. Students from the CBE completed the YLQ, which was conducted online. Students accessed the questionnaire via a link on the internet where they then were required to enter a password. The questionnaire was password protected to ensure that only those students participating in the study could access and complete it. The questionnaire required approximately 10 minutes for completion, and participants completed it at school during a time arranged by school staff. The questionnaire was completed at four points in time: 1) before participation in the TLB program, 2) immediately after the TLB program, 3) six-weeks after the program, and 4) six-months after participation in the TLB program. The students in the control group also completed the questionnaire at four points in time consistent with those of the

treatment group. For the purpose of validating the questionnaire, only data from the pre-test and first and second post-test were utilized to examine the reliability and degree of validity of the questionnaire.

To provide both exploratory and confirmatory evidence of the factor structure, the treatment participants who received the TLB training and had completed the YLQ at the first post-test time period were randomly divided into two groups (one for the exploratory analysis (EFA group) and one for the confirmatory analysis (CFA group)). The groups were similar to one another in relation to sample size (see Table 2). The statistical analyses described below were computed on both of these groups. The post-test 1 data (the second data collection period) was chosen because it was postulated that the factor structure of the scale on leadership could best be described on youth who had some specific training in leadership, plus this time point yielded the largest sample size compared to the other post program data collections. In addition, the factor structure of the YLQ was also examined using the control group's post-test 1 data (N= 169; 93 females, 76 males, 48 grade 10, 49 grade 11, 72 grade 12, mean age = 16.39 years, SD = 1.05) to determine if the factor structure held true for youth who had not received any leadership training. Data were analyzed in aggregate form and therefore, no individual analyses were warranted and no individually-identifying information was included in the results.

 Insert Table 2 Here

In examining the reliability of the YLQ, all of the participants' scores on the YLQ at the pre-test (N= 625), and post-tests 1 (N= 524) and 2 (N= 221) were examined, and the statistical

analyses described below were computed. To examine the YLQ's degree of construct validity, results from the 73 UofC participants who had completed the YLQ, SWLS, and the RSCA-AR online were examined and the statistical analyses described below were computed.

Factor analysis. The YLQ was constructed with a hierarchical factor structure consisting of a total leadership score and three domain/factor scores that contribute to the higher order total score. Based on theoretical conceptualization and empirical analyses items 1, 3, 4, 5, 7, and 21 were modeled onto factor 1 (Self-Concept), items 2, 9, 10, 14, 23, and 24 onto factor 2 (Emotional Intelligence), and items 13, 15, 16, 17, 20, and 22 onto factor 3 (Self-Efficacy). Table 3 provides a summary of the items that contributed to each of the factors. Item 8 was removed prior to the factor analyses as review of this item suggested that it was a poorly written item and potentially problematic because the item-total correlation with its factor was about .008. Appendix A provides a list of the items that make up the YLQ.

Insert Table 3 Here

Factor analyses were computed using LISREL. The first stage of this analysis was to conduct an exploratory factor analysis on the first group's set of data to determine whether the theoretical factor structure of the YLQ was reasonable. To provide evidence for this model, a confirmatory factor analysis was used on the second group's set of data to confirm that the factor structure was consistent across multiple samples. This approach to evaluating the factor structure of the YLQ is known as cross-validation (Gregory, 2004).

Exploratory factor analysis (EFA) was used on the first group of participants to help identify and gain an initial understanding of the factor structure of the YLQ. A confirmatory factor analysis framework in LISREL was used to conduct the EFA. Initially, all of the items of the YLQ, except for item 8 (My fears limit me), were grouped into the domain (Self-Concept, Self-Efficacy or Emotional Intelligence) that each item was believed to be theoretically most consistent with and then an EFA was conducted. Based on the results of the EFA, items that had weak factor loadings (less than $|\lambda| \geq .30$) or correlations (less than $|\lambda| \geq .40$) and large standardized residuals (greater than $|\lambda| \geq 3.0$) were removed from the YLQ (Jöreskog & Sörbom, 1993; Tabachnick & Fidell, 2001; Velicer & Fava, 1998). This analysis resulted in a revised version of the YLQ that would be used in the remainder of the analyses. After these items were removed, a second EFA using the first group's data was computed to explore and evaluate the factor structure of the revised scale. Similar to the previous EFA, a confirmatory factor analysis framework in LISREL was used for the EFA.

Confirmatory factor analysis (CFA) was used on the second group of participants to confirm the factor structure and demonstrate that the domains and the items in each domain cluster together (covary) in the way that research and theory would predict (Tabachnick, & Fidell, 2007). The reason for using a CFA as opposed to an EFA on the second group was that the items of the test were chosen based upon theory and empirical evidence and an EFA was already computed. The CFA was performed to determine how well the original three factor model fits the data. The three first-order factors included: Self-Concept, Self-Efficacy, and Emotional Intelligence. In general, CFA is a more stringent and strict model for testing factor

structure than EFA because it provides for test significance and it does not allow for items to cross-load onto multiple factors (Tabachnick & Fidell, 2007).

Reliability. Internal consistency, using Coefficient Alpha, and mean inter-item correlations for the total and subscales of the YLQ were used to evaluate the YLQ's reliability. These statistical analyses helped to maximize the YLQ's internal consistency (the degree to which the items are measuring the same trait) (Christensen, 2001). Temporal consistency (test-retest using post-test 1 and post- test 2 data) was calculated on the participants' total score on the YLQ. Parametric correlations were computed to determine the test-retest reliability of the YLQ at two different points in time. The correlation coefficient between the sets of responses was used to measure the test-retest reliability. The higher the correlation between the responses indicated the greater test-retest reliability of the YLQ (DeVellis, 1991). A shortcoming of the test-retest reliability is that there may be a test/practice effect, meaning that respondents "learn" to answer the same questions in the first testing period, and this affects their responses the next time they take the test (Christensen, 2001).

Construct validity. From a theoretical and conceptual framework, one would expect that leadership would be related to resiliency and satisfaction with life, but also be distinct. Previous research has shown that youth development/leadership programs can result in positive outcomes in the areas of personal and identity development, coping, problem-solving, resiliency, motivation, goal attainment, satisfaction, and academic and vocational achievement/performance (Benson & Saito, 2001; Dolbier, Soderstrom, & Steinhardt, 2001; Kendall, Hrycaiko, Martin, & Kendall, 1990; National Research Council and Institute of Medicine, 2002; Neck, Neck, Manz, & Godwin, 1999; Scales & Leffert, 1999). To assess convergent validity, the SWLS and the

RSCA-AR were used. Parametric correlational analysis from university participants' data from the YLQ and the other scales were conducted to evaluate the degree of construct validity of the YLQ. The correlation coefficients were reviewed to determine if the YLQ was more related to (has high correlations with) the other scales. In addition, the CFA provided further evidence for the construct validity of the YLQ.

Summary

The objectives of this study were to develop a measure of youth leadership and evaluate its psychometric properties. The development of the items of the YLQ was primarily an evidence/empirically-based process stemming from data from a previous research study on youth leadership; however, a literature and theoretical review also provided insight into the conceptualization of the construct and the domains/skills that are important to leadership. Stemming from the empirical and theoretical data, self-concept, emotional intelligence, and self-efficacy were chosen as the domains for the YLQ.

Pre-test and post-test one and two data from the group of students who participated in the TLB program as well as similar data from a control group were used for evaluating the reliability and degree of construct validity of the YLQ. Reliability was measured using Coefficient Alpha and test-retest techniques. EFA and CFA were computed to assess the fit of the model and assess the degree of construct validity of the YLQ. Construct validity was further investigated through correlations of undergraduate university students' scores on the YLQ, SWLS, and RSCA-AR.

Chapter Four: Results

Results

Exploratory factor analysis (EFA). A proposed factor structure based on theory was used for the EFA. It consisted of three factors (Self-Concept, Emotional Intelligence, and Self-Efficacy). The initial EFA in LISREL 8.80 with all the items of the YLQ, except for item 8, used raw data to estimate the parameters. The root mean square error of approximation (RMSEA = .08) suggested that the fit of the model was not adequate. The comparative fit index (CFI = .97), the Tucker-Lewis Index/Non-normed fit index (NNFI = .97), and the goodness of fit index (GFI = .81) suggested that the model fit might be adequate. Exploration of the regression equations (i.e., the factor loadings, R-squares and the standardized residuals) revealed that some of the items did not load on any of the factors (See Table 4 for the regression equations). Large standardized residuals indicated fit problems (See Appendix B for standardized residuals). Table 5 provides the factor loadings for each of the items on its factor. The items with poor factor loadings or large standardized residuals were eliminated from the YLQ and the remaining analyses. In total, seven items were removed, resulting in a total of 18 items to be analyzed (See Appendix A for the revised scale).

Insert Table 4 Here

Insert Table 5 Here

A second EFA in LISREL 8.80 involving the same factor structure and the same sample used the revised YLQ items' raw data to estimate the parameters. The root mean square error of approximation (RMSEA = .067), the comparative fit index (CFI = .98), the Tucker-Lewis Index/Non-normed fit index (NNFI = .98), and the goodness of fit index (GFI = .87) suggested that the fit of the model was adequate. Table 6 provides the factor loadings of each item on its factor, Table 7 provides the regression equations, and Appendix B provides the standardized residuals for the model.

Insert Table 6 Here

Insert Table 7 Here

Confirmatory factor analysis (CFA). The three factor model structure that was specified for the EFA was used for the CFA in order to confirm the factor structure of the exploratory model. The CFA in LISREL 8.80 used the raw data to estimate the parameters. The root mean square error of approximation (RMSEA = .07), the comparative fit index (CFI = .97), the Tucker-Lewis Index/Non-normed fit index (NNFI = .97), and the goodness of fit index (GFI

= .86) suggested that the fit of the model was adequate. Table 8 provides the factor loadings of each item on its factor, Table 9 provides the regression equations, and Appendix B provides the standardized residuals for the model.

Insert Table 8 Here

Insert Table 9 Here

Control group factor analysis. The model structure that was specified for the EFA and CFA was used for the factor analysis of the control group data in order to determine if the model structure is a good fit for individuals who have not received formal leadership training. The factor analysis computed in LISREL 8.80 was based on the raw data. The root mean square error of approximation (RMSEA = .10) suggested that the model was not a good fit for individuals who have not received leadership training. The comparative fit index (CFI = .95), the Tucker-Lewis Index/Non-normed fit index (NNFI = .94), and the goodness of fit index (GFI = .81) suggested that the fit of the model might be adequate. In addition, the standardized residual matrix for the control group had larger residuals than the groups used in the EFA and CFA (See Appendix B for the standardized residuals). Large residual values indicated fit problems. Table 10 provides the factor loadings for each item on its factor and Table 11 provides the regression equations.

Insert Table 10 Here

Insert Table 11 Here

Reliability. The YLQ as a whole had an exceptional level of internal consistency, as evaluated by a Coefficient Alpha of .95, with mean inter-item correlations ranging from $r = .31$ to $.67$. These results suggested that all items of the YLQ correlated with the full scale in an acceptable manner. The subscales of the YLQ also demonstrated acceptable reliability. Reliability coefficients, inter-item correlations, and item-total correlations for the YLQ are provided in Tables 12, 13, and 14. The performance of students on the YLQ was examined twice to evaluate the test-retest reliability. The relationship between the two testing periods showed an adequate and significant temporal consistency for the YLQ, as measured by the Pearson correlation coefficient ($.71, p < .01$).

Insert Table 12 Here

Insert Table 13 Here

Insert Table 14 Here

Construct validity. The construct validity analysis of the YLQ was assessed through the CFA described above, which confirmed the factor structure of the YLQ. Construct validity was further evaluated through correlations with the Satisfaction With Life Scale (SWLS) and the Resiliency Scale for Children and Adolescents- Adult Revision (RSCA-AR). The YLQ was significantly correlated with the SWLS and the RSCA-AR. Parametric correlation coefficients are provided in Table 15. These results indicated low to moderate correlation coefficients.

Insert Table 15 Here

Summary

Overall, the preliminary analyses of some of the psychometric properties of the YLQ were promising as the results indicated exceptional reliability and an acceptable degree of validity. Reliability analyses suggested that the YLQ was consistent, both internally and temporally. The results from the CFA confirmed the hierarchical three factor structure that was proposed for the EFA. This structure consisted of the domains of Self-Concept, Emotional

intelligence, and Self-Efficacy, which were drawn from the previously mentioned theoretical framework, and contributed to an overall/total leadership score on the YLQ. These findings suggested that the fit of the model was adequate based on data from the revised version of the YLQ. In addition, the factor analysis on the control group provided additional support for the proposed model. Results from these factor analyses as well as positive correlations/relationships with the SWLS and RSCA-AR provided support for the construct validity of the YLQ.

Chapter Five: Discussion

Although some measures of leadership exist for post-secondary students and young adults (for example: The Student Leadership Practices Inventory by Kouzes & Posner, 1998; The Leadership Attitudes and Beliefs Scale Third Edition by Wielkiewicz, 2000) and they may be appropriate for youth, currently, there are no known measures that specifically assess youth's leadership skills and performance. Consequently, there is limited research into the efficacy of youth leadership processes/programs and researchers' understanding of why or how leadership programs work (Allio, 2005). Therefore, the need for the development of leadership measures that provide insight into the perceived impact and enhancement of youths' leadership skills exists. This study is a first step towards establishing a metric for assessing youth leadership.

The goal of this research was to develop a reliable and valid instrument that could assess youth leadership. In Canada, the majority of adolescents do not receive training in leadership and many are unaware of their performance and capabilities on a variety of leadership skills (Astin & Astin, 2000; van Linden & Fertman, 1998). As a result, they have limited understanding of how leadership effects their academic and daily life functioning. Therefore, the need exists to further educate and evaluate youth on the skills associated with leadership. Youth require opportunities and experiences to learn and develop leadership skills, as well as they need to be supported in the process. By providing opportunities for youth to develop leadership skills, youth will gain an increased awareness of the importance of leadership and an understanding of why and how leadership can improve their own as well as others' lives (Schwean et al, 2009).

In this study, the Youth Leadership Questionnaire (YLQ) was described and evaluated. The YLQ is a brief self-report measure that quickly provides information about youth's leadership skills. Overall, the preliminary results were very positive, as the YLQ had acceptable

psychometric properties and accuracy as a screening instrument to detect leadership skills and competencies among youth.

In statistical terms, reliability is the ability of an instrument to measure something consistently and repeatedly. It is the stability of that measure to produce the same results when measuring a construct (Christensen, 2001). Reliability is an essential and necessary property of tests (McIntire & Miller, 2007). The reliability coefficients of the YLQ were at least .7 or greater, which is the widely accepted value (Zeidner & Most, 1992). The scale showed exceptional internal consistency, suggesting that the items are measuring a coherent leadership construct. The high test re-test coefficients also suggest that youth leadership, as measured by the YLQ, is relatively stable over time.

Construct validity refers to how theoretically sound the instrument is (Kane, 2006). The YLQ scores were positively related to other measures (Satisfaction With Life Scale (SWLS) and Resiliency Scale for Children and Adolescents- Adult Revision (RSCA-AR)) of domains/areas that are associated with social and emotional functioning and leadership (Roth & Brooks-Gunn, 1998; Scheer, 1997; Tusaie et al., 2007). It, therefore, appears that the YLQ may be useful as an indicator of not only leadership, but also to some degree of resiliency and life satisfaction. Overall, these findings provide insight into a few of the behaviours and other constructs that are both theoretically and empirically related to youth leadership. Although there are other behaviours and constructs that, in subsequent research, should be examined in relation to leadership, this study provides preliminary evidence of the construct validity of the YLQ relative to the SWLS and RSCA-AR.

Factor analysis is a statistical process that is used to establish how individual items cluster around a given dimension (Tabachnick, & Fidell, 2007). It can also be used as supporting

evidence for the construct validity of the measure under construction (Gregory, 2004; Zumbo, 2007). In the current study, the chi-square values of the exploratory (EFA) and confirmatory factor analyses (CFA) for the YLQ were not reported because researchers suggest that you should not rely on the chi-square statistic as it is sensitive to large sample size (Stevens, 2002). A significant chi-square result may, at least, be in part to the large sample size and less to any serious misspecifications of the model. The use of several indexes and those that are less influenced by sample size (example, Root Mean Square Error of Approximation) is encouraged (Stevens, 2002). Overall, results of the EFA and CFA (based on the Root Mean Square Error of Approximation, Comparative Fit Index, Tucker-Lewis Index/Non-Normed Fit Index, and Goodness of Fit Index) indicated that the YLQ three factor model (Self-Concept, Emotional Intelligence, and Self-Efficacy) provided a good fit of the data. Moreover, the better fit of the model for the group that received the Rapport Teen Leadership Breakthrough (TLB) program compared to the control group who did not receive training provided further support that the YLQ is measuring youth leadership.

Overall, based on the data obtained in this study regarding its factorial structure, internal and temporal consistencies, and degree of construct validity, the YLQ appears to display an adequate degree of construct validity and is a reliable test for use in screening and measuring youth leadership. The EFA and CFA provided support for the proposed hierarchical factor model, with total leadership as the general factor and self-concept, self-efficacy, and emotional intelligence as the domains/factors. Thus, the way youth leadership was theoretically conceptualized was resembled empirically by the data. The model was further supported and strengthened by correlations that showed that the scores on the YLQ had significant associations with satisfaction with life and resiliency.

Significance/Implications

The first evaluation of the YLQ is promising. The YLQ may be a helpful measure for researchers who want to evaluate the outcomes and efficacy of leadership programs. It would also be helpful for educators or practitioners who want to assess youths' perceptions of their leadership skills. One of the YLQ's chief advantages is its simplicity; it is brief and easy to administer and score. These advantages make it potentially useful for quickly gauging a youth's level of leadership on particular skills.

The YLQ has potential utility within an educational perspective as it was constructed and evaluated within a school context. The findings from this study may assist the Calgary Board of Education and other school boards or organizations in their continuing efforts and commitment to offer quality programs to students, in that a measure to determine program efficacy is available. Given that youth spend a significant amount of their time in schools and school personnel play a strong role in nurturing youth, educational programs are ideal for providing interventions and prevention programs that promote and enhance youth leadership skills as identified on the YLQ (Schwean et al., 2009; Webber & Mulford, 2007). Schools can provide youth with the resources to develop leadership skills as well as they can provide opportunities to model and practice the skills. Through the development of interventions and programs that promote or enhance mental health and leadership, it is anticipated that youth will experience more positive outcomes, including increased school and life satisfaction (Gilman & Huebner, 2003), greater school engagement (Furlong et al., 2003), and a better quality of life (Huebner & Gilman, 2004). Further, it has considerable importance for determining the efficacy of other youth leadership development programs.

Given the relation of leadership to a number of educational, vocational and developmental aspects- academic achievement, resiliency, emotional intelligence, and interpersonal skills (Bandura, 1997; Edward & Warelow, 2005; Hindes et al., 2008; Stajkovic & Luthans, 1998) - research on the evaluation of leadership skills in youth should be a priority. Improved evaluation of youth leadership with reliable and valid scales will facilitate an increased understanding of youth leadership and its relationship to emotional well-being, interpersonal relationships, and social competence (Gardner, 1993; Pekrun, 1992). Research has shown that youth who are involved in youth development and leadership programs have positive personal, social, identity, moral, and life skill development, better mental health, lower rates of school drop-out, and increased social acceptance, self-efficacy, self-determinism, problem-solving skills, and academic achievement/outcomes (Benson & Saito, 2001; Dolbier et al, 2001; Eccles & Gootman, 2002; National Research Council and Institute of Medicine, 2002; Pittman, Irby, Tolman, Yohalem & Ferber, 2003; Sagawa, 2003; Scales & Leffert, 1999). Positive relationships have also been found between leadership and work-related outcomes (Neck et al, 1999) and an active coping style (Dolbier et al., 2001). Therefore, development and measurement of youth leadership skills can be beneficial to youth in multiple domains of their life.

A well-developed test of youth leadership was sought out and is important because it can provide researchers, educators, and clinicians with quality data to answer various research and practical questions. The development of a scale, such as the YLQ, can be a catalyst for research in many areas, including exploring program efficacy, complex modeling of leadership, and examining protective factors. Having a psychometrically sound scale for youth leadership is important because the ability to answer these types of research and clinical questions is only as good as the instrument used (McIntire & Miller, 2007; Messick, 1980; 1989).

Limitations

There are a number of issues that require continuing investigation. Although the YLQ was designed for assessing youth in a school setting, data regarding its effectiveness to evaluate youth's leadership skills in other contexts and situations is unknown. Youth leadership is not confined to schools, but instead exists across diverse settings (community, home, school, society), with each context influencing one another. One needs to consider the context within which the youth is leading to decide which skills to evaluate as they may differ in different environments (Northouse, 2007; van Linden & Fertman, 1998). Reliability and validity in other community and familial contexts will need to be assessed, as well as the utility of the YLQ with more diverse youth populations (for example, youth from differing cultures, at-risk populations and those with social-emotional difficulties).

In addition, there is a lack of a universally accepted definition of leadership for which the conceptualization of the construct can be based (Burger, 2007; Wallace et al., 2007). Leadership is complex and multifaceted, as there is no single trait or skill that singularly determines performance as a leader (Schwean et al., 2009). Not all leaders have the same level of skills/competencies in all the domains of leadership, and therefore, leadership cannot be measured on a single dimension. For this study, leadership was viewed within a narrow framework comprising of only three domains (Self-Concept, Emotional Intelligence, and Self-Concept). However, it may be that other skills/domains are associated with youth leadership and should be included within the YLQ.

Another limitation of the study is that it was unknown specifically how the schools and youth were recruited and chosen for participation in the study because the sample was comprised of schools and students who were chosen specifically by school personnel to participate in either

the control or the treatment groups. Therefore, random selection and random assignment were not used, which impacts the internal validity and generalizability of the results (Christensen, 2001). In addition, it remains unknown if the youth selected were those who school personnel felt lacked some leadership skills or if they were youth who were already in leadership roles or perceived as having the potential to be a leader. These cautions must be kept in mind when interpreting the findings.

Lastly, the study utilized two different populations (high school students and undergraduate university students) for validating the YLQ because, at the time when evidence was being collected for the degree of construct validity, the university students were readily available and the high school students were unable to complete the additional measures. It is possible that university students' responses were different from those of high school students. However, given that the domains chosen for the YLQ are robust, it is likely that they extend and can be applied to young adults, and can also be used to represent/evaluate leadership in this populations/age group. Research has shown that youth and young adults in post-secondary education are developing and increasing their skills associated with leadership (for example, self-awareness, self-concept and emotional intelligence), suggesting that leadership may not vary significantly between these two populations (Pascarella & Terenzini, 2005). However, it will be important that in future research on the YLQ, the sample be as representative as possible of the population for which the YLQ is intended (Zeider & Most, 1992).

Future Directions

The results presented herein represent a preliminary step in the development of the YLQ. One of the purposes of this study was to test the items of the YLQ so they could be evaluated and if needed revised, eliminated or replaced. It is common during the process of test development to

go through these steps to possibly increase the reliability and degree of validity of the scores and interpretations of the YLQ (Clark & Watson, 1995). Therefore, the next step in the development of the YLQ is to review and evaluate the current items and structure of the test. This next step should include a re-assessment of the content validity of the YLQ by having a panel of experts review the domain specifications and judge whether the items are appropriate and adequate. In addition, a more in-depth item and domain analysis of leadership and the YLQ is required to find those items and domains that best form an internally consistent scale. This analysis will also help determine if the construct of leadership used in the YLQ was too broad or too vaguely defined. Therefore, items representing other aspects of leadership could possibly be included in the scale. Future work should also focus on whether leadership is best represented by a multidimensional/multifactorial scale or would a unidimensional model better fit the data.

The subsequent step would be to collect new data from a second sample that is similar to the original one using the new revised YLQ scale. Construct validity requires a variety of research/evidence to be collected over time from numerous sources (Gregory, 2004; Kane, 2006; Murphy & Davidshofer, 2005; Zumbo, 2007). Currently, there is only one small piece of research supporting the degree of construct validity of the YLQ. Therefore, further research into the kinds of relationships the YLQ has with other similar and dissimilar tests is required. The purpose of collecting new and additional data is to repeat the analysis procedures of this original study to provide support for the revised YLQ and determine if the measure can be finalized.

Further analyses are also needed to determine the YLQ's concurrent and predictive validity for areas such as academic achievement, graduation, and post-secondary or vocational success. Therefore, future directions for research with the YLQ should include a cross-validation study. Cross-validation studies involve using a regression equation obtained from a previous

study on the YLQ in a new sample to determine whether the test predicts the criterion as well as it did in the previous study. Cross-validation studies are recommended to provide support that the findings did not occur simply by chance (Ghiselli, Campbell, & Zedeck, 1981; Gregory, 2004).

In conclusion, the results of the study are preliminary, but positive. They are the initial rather than the final steps in the development of the YLQ as test construction requires time and multiple procedures. Given the initial findings of the YLQ and some of the proposed next steps in research/development, the construct of leadership will be more clearly conceptualized and measured, and the practical utility of the YLQ will become more defined.

References

- Aiken, L. (1983). Number of response categories and statistics on a teacher rating scale. *Educational and Psychological Measurement*, 43, 397-401.
- Allio, R.J. (2005). Leadership development: teaching versus learning. *Management Decision*, 43, 1071-1077.
- Anastasi, A. (1986). Evolving concepts of test validation. *Annual Reviews in Psychology*, 37, 1-15.
- Angoff, W.H. (1988). Validity: An evolving concept. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp. 19-32). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Astin, A. & Astin, H. (2000). *Leadership reconsidered: Engaging higher education in social change*. Battle Creek, MI: W. K. Kellogg Foundation.
- Avolio, B. (1999). *Full leadership development: Building the vital forces in organizations*. Thousand Oaks, CA: Sage.
- Avolio, B., & Bass, B. (1995). Individual consideration viewed at multiple levels of analysis: A multi-level framework for examining the diffusion of transformational leadership. *Leadership Quarterly*, 6, 199-218.
- Avolio, B., Bass, B., & Jung, D. (1999). Re-examining the components of transformational and transactional leadership using the multifactor leadership questionnaire. *Journal of Occupational and Organizational Psychology*, 72, 441-462.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: H.W. Freeman.

- Bass, B. (1985). *Leadership and performance beyond expectations*. New York: Free Press.
- Bass, B. (1990). From transactional to transformation leadership: Learning to share the vision. *Organizational Dynamics*, 18, 19-31.
- Bass, B., Avolio, B., Jung, D., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology*, 88, 207-218.
- Benson, P., & Saito, R. (2000). *The scientific foundations of youth development*. Minneapolis: Search Institute.
- Berg, C., & Sternberg, R.J. (2003). Multiple perspectives on the development of adult intelligence. In J. Demick & C. Andreoletti (Eds.), *Handbook of adult development* (pp. 103-119). New York: Kluwer Academic.
- Burger, J. (2007). In search of the elusive general theory. In J.M. Burger, C.F. Webber, & P. Klinck (Eds.), *Intelligent leadership* (pp. 1-12). Dordrecht, The Netherlands: Springer.
- Chan, D. W. (2007). Leadership competencies among Chinese gifted students in Hong Kong: The connection with emotional intelligence and successful intelligence. *Roeper Review*, 29(3), 183-189.
- Chang, E. (2001). *Optimism & pessimism: Implications for theory, research, and practice*. Washington, DC: American Psychological Association.
- Chemers, M.M., Watson, C.B., & May, S.T. (2000). Dispositional affect and leadership effectiveness: A comparison of self-esteem, optimism, and efficacy. *Personality and Social Psychology Bulletin*, 26, 267-277.

- Christensen, L.B. (2001). *Experimental methodology*. Boston: Allyn & Bacon.
- Churchill, G.A. & Peter, P.A. (1984). Research design effects on the reliability of rating scales: A meta analysis. *Journal of Marketing Research*, 21(4), 360-375.
- Ciarrochi, J. V., Chan, A. YC, & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences*, 28, 539-561.
- Clark, L.A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309-319.
- Conner, J., & Strobel, K. (2007). Leadership development. *Journal of Adolescent Research*, 22, 275-297.
- Cox, E.P. (1980). The optimal number of response alternatives for a scale. *A Review Journal of Marketing Research*, 17(4), 407-422.
- Cronbach, L.J. (1980). Validity on parole: How can we go straight? *New Directions for Testing and Measurement*, 5, 99-108.
- Dachler, H.P., & Hosking, D.M. (1995). The primacy of relations in socially constructing organisational reality. In Hosking D.M., Dachler, H., & Gergen, K.J. (Eds.), *Management and organisation: Relational alternatives to individualism*. Avebury, Aldershot.
- Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences*, 1021, 1-22.
- Dawis, R.V. (1987). Scale construction. *Journal of Counseling Psychology*, 34, 481-489.

- Day, D. V. (2000). Leadership development: A review in context. *Leadership Quarterly*, 11, 581-613.
- Day, D. V., & O'Connor, P. M. G. (2003). Leadership development: Understanding the process, in S.E. Murphy & R. E. Riggio (Eds), *The future of leadership development* (pp. 11-28). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Dempster, N., & Lizzio, A. (2007). Student leadership: Necessary research. *The Australian Journal of Education*, 51, 276-285.
- Des Marais, J., Yang, Y., & Farzanehkia, F. (2000). Service-learning leadership development for youths. *Phi Delta Kappan*, 81, 678 – 680.
- DeVellis, R.F. (1991). *Scale development*. Newbury Park, NJ: Sage Publications.
- Dickmann, M.H., & Stanford-Blair, N. (2009). *Mindful leadership: A brain-based framework, Second edition*. Thousand Oaks, CA: Corwin Press.
- Dolbier, C. L., Soderstrom, M., & Steinhardt, M. (2001). The relationships between self-leadership and enhanced psychological, health, and work outcomes. *The Journal of Psychology*, 135(5), 469-485.
- Eccles, J., & Gootman, J.A. (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.
- Edward, K. L., & Warelow, P. (2005). Resiliency: When coping is emotionally intelligent. *Journal of the American Psychiatric Nurses Association*, 11, 101-103.
- Friedman, H. & Friedman, L.W. (1986). On the danger of using too few points in a rating scale: A test of validity. *Journal of Data Collection*, 26(2), 60-63.

- Friedman, H. H., Wilamowsky, Y., & Friedman, L. W. (1981). A comparison of balanced and unbalanced rating scales. *The Mid-Atlantic Journal of Business*, 19, 2, 1-7.
- Frisen, A., Jonsson, A., & Persson, C. (2007). Adolescents' perception of bullying: Who is the victim? Who is the bully? What can be done to stop bullying? *Adolescence*, 42, 749-761.
- Furlong, M. J., Whipple, A. D., St. Jean, G., Simental, J., Soliz, A., & Punthuna, S. (2003). Multiple contexts of school engagement: Moving toward a unifying framework for educational research and practice. *The California School Psychologist*, 8, 99 – 113.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin*, 2, 66-70.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W.H. Freeman.
- Gilman, R., & Huebner, S. (2003). A review of life satisfaction research with children and adolescents. *School Psychology Quarterly*, 18, 192 – 205.
- Gregory, R.J. (2004). *Psychological testing: History, principles, and applications (4th ed.)*. New York: Pearson.
- Halloran, J., & Benton, D. (1987). *Applied human relations*. Englewood Cliffs, N.J.: Prentice Hall.
- Hindes, Y.H., Thorne, K.J., Schwean, V.L. & McKeough, A.M. (2008). Promoting Intrapersonal Qualities in Adolescents: Evaluation of Rapport's Teen Leadership Breakthrough Program. *Canadian Journal of School Psychology*, 23, 206-222.

- Hoare, C. (Ed.). (2006). *Handbook of adult development and learning*. Oxford: Oxford University Press.
- Holander, E.P. (1986). On the central role of the leadership processes. *International Review of Applied Psychology*, 35, 39-52.
- Hollander, E.E. (1978). *Leadership dynamics: A practical guide to effective relationships*. New York: Free Press/Macmillan.
- Hotton, T., & Haans, D. (2004). Alcohol and drug use in early adolescence. *Health Reports*, 15(3), 9-19.
- Howell, J. & Avolio, B. (1993). Transformational leadership, transactional leadership, locus of control, and support for innovation: Key predictors of consolidated business unit performance. *Journal of Applied Psychology*, 7, 891-902.
- Huebner, S. E., & Gilman, R. (2004). Perceived quality of life: A neglected component of assessments and intervention plans for students in school settings. *The California School Psychologist*, 9, 127 – 134.
- Izard, J. (2005). Trial testing and item analysis in test construction. In K.N. Ross' (Ed.), *Quantitative research methods in educational planning*. Paris, France: UNESCO.
- Jöreskog, K. G. & Sörbom, D. (1993). *LISREL 8 user's reference guide*. Chicago: Scientific Software International.
- Judge, T.A., Bono, J.E., Ilies, R., & Gerhardt, M.W. (2002). Personality and Leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87, 765-780.

- Kane, M.T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38, 319-342.
- Kane, M.T. (2006). Validation. In R.L. Brennan (Ed.), *Educational measurement fourth edition* (pp. 17-64). Westport, CT: American Council on Education and Praeger.
- Kegan, R. (1980). Making meaning: The constructive-developmental approach to persons and practice. *The Personnel and Guidance Journal*, 58, 373-380.
- Kendall, G., Hrycaiko, D., Martin, G., & Kendall, T. (1990). The effects of an imagery rehearsal, relaxation and self-talk package on basketball game performance. *Journal of Counseling Psychology*, 32, 263-271.
- Kerlinger, F.N., & Lee, H.B. (2000). *Foundations of behavioral research fourth edition*. Toronto: Wadsworth Thomson Learning.
- Kirkpatrick, S.A., & Locke, E.A. (1991). Leadership: Do traits matter? *Academy of Management Executive*, 5, 48-60.
- Kouzes, J., & Posner, B. (1998). *Student leadership practices inventory*. San Francisco: Jossey-Bass.
- Williams, E. (2002).
- Lehmann, D.R. & Hulbert, J. (1972). Are three-point scales always good enough? *Journal of Marketing Research*, 9(4), 444-446.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694.
- Lopes, P. N., Salovey, P., Cote, S., & Beers, M. (2005). Emotion regulation abilities and the quality of social interaction. *Emotion*, 5(1), 113-118.

- Lord, R.G., Brown, D.J., & Freiberg, S.J. (1999). Understanding the dynamics of leadership: The role of the follower self-concepts in the leader/follower relationship. *Organizational Behavior and Human Decision Processes*, 78, 167-203.
- Lord, R.G., De Vader, C.L., & Alliger, G.M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, 71, 402-410.
- Lussier, R., & Achua, C. (2004). *Leadership: Theory, application & skill development second edition*. Eagan, MA: Thomson South-Western.
- Masten, A. S. (2001). Ordinary magic: Resiliency processes in development. *American Psychologist*, 56, 227-238. Reprinted in M. E., Hertzog & E. A. Farber (Eds.) (2005), *Annual progress in child psychiatry and child development: 2002*. New York: Routledge.
- Masten, A., Burt, K., Roisman, G., Obradovic, J., Long, J., & Tellegen, A. (2004). Resources and resiliency in the transition to adulthood: Continuity and change. *Development & Psychopathology*, 16, 1071-1094.
- Matell, M.S. & Jacoby, J. (1972). Is there an optimal number of alternatives for Likert scale items? Effects of testing time and scale properties. *Journal of Applied Psychology*, 56, 506-509.
- Maurer, T.J. (2001). Career-relevant learning and development, worker age, and beliefs about self-efficacy for development. *Journal of Management*, 27, 123-140.

- McCauley, C.D., Drath, W.H., Palus, C.J., O'Connor, P.M.G., & Baker, B.A. (2006). The use of constructive-developmental theory to advance the understanding of leadership. *The Leadership Quarterly*, 17, 634-653.
- McIntire, S. A. & Miller, L. A. (2007). *Foundations of psychological testing: a practical approach*. 2nd ed. Thousand Oaks, CA: Sage Publications.
- McLaughlin, M. (2000). *Community counts: How youth organizations matter for youth development*. Washington, DC: Public Education Network.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement third edition* (pp. 13-103). New York: American Council on Education and Macmillan.
- Mohamed, I., & Wheeler, W. (Eds.) (2001). *Broadening the bounds of youth development: Youth as engaged citizens*. Chevy Chase, MD: The Innovation Center for Community and Youth Development and The Ford Foundation, and New York: Ford Foundation.
- Mumford, M.D., Zaccaro, S.J., Harding, F.D., Jacobs, T.O., & Fleishman, E.A. (2000). Leadership skills for a changing world: Solving complex social problems. *The Leadership Quarterly*, 11, 11-35.
- Murphy, K.R., & Davidshofer, C.O. (2005). *Psychological testing: Principles and applications sixth edition*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Nagle, A. (2003). *Power and possibilities*. New York: Ms. Foundation for Women.

- National Research Council and Institute of Medicine (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.
- Neck, C. P., Neck, H. M., Manz, C. C., & Godwin, J. (1999). I think I can, I think I can": A self-leadership perspective on enhancing entrepreneur thought patterns, self-efficacy, and performance. *Journal of Managerial Psychology*, 14(6), 477-501.
- Northouse, P.G. (2007). *Leadership: Theory and practice fourth edition*. Thousand Oaks, CA: Sage Publications.
- Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students: A third decade of research*. San Francisco: Jossey-Bass.
- Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, 5(2), 164-172.
- Pekrun, R. (1992). The impact of emotions on learning and achievement: towards a theory of cognitive/motivational mediators. *Applied Psychology: An international review*, 41(4), 359-376.
- Pittman, K.J., Irby, M., Tolman, J., Yohalem, N., & Ferber, T. (2003, March). *Preventing problems, promoting development, encouraging engagement: Competing priorities or inseparable goals?* Washington, DC: The Forum for Youth Investment, Impact Strategies, Inc.
- Prince-Embury, S. (2006). *Resiliency scales for children and adolescents: A profile of personal strengths*. San Antonio: Harcourt.

Rapport Leadership International. (N.D.). *Teen leadership breakthrough*. Retrieved October 18, 2009 from: <http://rapporteducation.com/courses-for-students/teen-leadership-breakthrough.html>.

Roach, A. A. (1999). Leadership giftedness: Models revisited. *Gifted Child Quarterly*, 43, 13-24.

Roth, J., & Brooks-Gunn, J. (1998). Promoting healthy adolescents: synthesis of youth development programmes. *Journal of Research on Adolescence*, 8, 423–459.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry*, 57, 316-331.

Sagawa, S. (2003). Service as a strategy for youth development. In A. Lewis (Ed.) *Shaping the future of American youth: Youth policy in the 21st century*. Washington, DC: American Youth Policy Forum.

Scales, P. & Leffert, N. (1999). *Developmental assets: A synthesis of the scientific research on adolescent development*. Minneapolis: Search Institute.

Scheer, S. D. (1997). Youth Leadership and Community Service: A perfect combination. *Leadership Link*, 1-2.

Schneider, B., Ehrhart, K. H., & Ehrhart, M. G. (2002). Understanding high school student leaders II: Peer nominations of leaders and their correlates. *The Leadership Quarterly*, 13, 275-299.

Schwean, V., Bishop, P., & McKeough, A. (2009). *Executive summary: Centre for child & youth leadership. "Thriving learners- Leaders for social change"*. (Unpublished manuscript).

- Shepard, L.A. (1993). Evaluating test validity. In L. Darling-Hammon (Ed.), *Review of research in education* (pp. 405-450). Washington, DC: AERA.
- Sireci, S.G. (1998). Gathering and analyzing content validity data. *Educational Assessment*, 5, 299-321.
- Sireci, S.G. (2007). On validity theory and test validation. *Educational Researcher*, 36, 477-481.
- Stajkovic, A.D. & Luthans, F. (1998a). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124, 240-261.
- Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences- fourth edition*. Mahwah, NJ: Lawrence Erlbaum Associates
- Stodgill, R.M. (1974). *Handbook of leadership: A survey of theory and research*. New York: the Free Press.
- Suen, H.K. (1990). *Principle of test theories*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.
- Tabachnick, B., & Fiddell, L. (2007). *Using multivariate statistics- fifth edition*. Boston: Pearson Education.
- Thomas, R.M. (2005). *Comparing theories of child development sixth edition*. Belmont, CA: Thomson Wadsworth.
- Tull, D.S., & Hawkins, D. I. (1993). *Marketing research: Measurement and method*. New York: Macmillan Publishing.

- Tusaie, K., Puskar, K., & Sereika, S.M. (2007). A predictive and moderating model of psychosocial resilience in adolescents. *Journal of Nursing Scholarship*, 39, 54-60.
- Uhl-Bien, M. (2006). Relational leadership theory: Exploring the social processes of leadership and organizing. *The Leadership Quarterly*, 17, 654-676.
- van Linden, J. A., & Fertman, C. I. (1998). *Youth leadership: A guide to understanding leadership development in adolescents*. San Francisco: Jossey-Bass.
- Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3(2), 231-251.
- Wallace, J., Foster, R., & DaCosta, J. (2007). A survey of existing school leadership preparation and development programs. In J.M. Burger, C.F. Webber, & P. Klinck (Eds.), *Intelligent leadership* (pp. 183-210), Dordrecht, The Netherlands: Springer
- Webber, C.F. & Mulford, B. (2007). School and community. In J.M. Burger, C.F. Webber, & P. Klinck (Eds.), *Intelligent leadership* (pp. 119-142), Dordrecht, The Netherlands: Springer.
- Wielkiewicz, R. M. (2000). The Leadership Attitudes and Beliefs Scale: An instrument for evaluating college students' thinking about leadership and organizations. *Journal of College Student Development*, 41, 335-347.
- Worcester, R.M. & Burns, T.R. (1975). A statistical examination of the relative precision of verbal scales. *Journal of Market Research Society*, 17(3), 181-197.
- Woyach, R. (1996, Spring). Five principles for effective youth leadership development programs. *Leadership Link*, 1-2.

Zaccaro, S.J. (2007). Trait based perspectives of leadership. *American Psychologist*, 62, 6-16.

Zaccaro, S.J., Gilbert, J.A., Thor, K.K. & Mumford, M.D. (1991). Leadership and social intelligence: Linking social perceptiveness to behavioral flexibility. *Leadership Quarterly*, 2, 317-347.

Zaccaro, S.J., Kemp, C., & Bader, P. (2004). Leader traits and attributes. In J. Antonakis, A.T. Cianciolo, & R.J. Sternberg (Eds.), *The nature of leadership* (pp. 101-124). Thousand Oaks, CA: Sage.

Zeidner, M., & Most, R. (Eds.). (1992). *Psychological testing: An inside view*. Mountain View, CA: Consulting Psychologists Press.

Zumbo, B.D. (2007). Validity: Foundational issues and statistical methodology. *Handbook of Statistics*, 26, 45-79.

Zumbo, B.D., Gelin, M.N., & Hubley, A.M. (2002). *The construction and use of psychological tests and measures. In the psychology theme of the encyclopedia of life support systems*. Oxford, UK: Eolss Publishers.

Table 1.

University Participants' Area of Study/Degree Programs

Area of Study	N
Bachelor of Arts	37
Bachelor of Science	23
Bachelor of Communication	2
Bachelor of Arts & Science	2
Bachelor of Education	2
Bachelor of Fine Arts	1
Bachelor of Health Science	3
Business	1
Open Studies	2

Table 2.

Descriptive Statistics of the Groups

	N	Mean Age	Males	Females	Grade 10	11	12
EFA Group	180	16.05 (SD= .94)	72	108	79	55	46
CFA Group	175	15.87 (SD= 1.08)	77	98	96	42	37

Table 3.

YLQ Items and Their Associated Factors

	Self-Concept	Emotional Intelligence	Self-Efficacy
Item 1. I will keep on trying, even if I am not successful the first time.	x		
Item 2. I always make an effort to make new people feel welcome.		x	
Item 3. I feel comfortable speaking in public (in front of an audience).	x		
Item 4. I enjoy new challenges.	x		
Item 5. I am willing to stand up for myself.	x		
Item 6. I try to be sensitive to others' values, beliefs, and/or opinions.	x		
Item 7. I am aware that my decisions have consequences.	x		
Item 8. My fears limit me.	Poor item		
Item 9. I can effectively use others' feedback to improve myself.		x	

	Self-Concept	Emotional Intelligence	Self-Efficacy
Item 10. I make an effort to effectively communicate with others.		x	
Item 11. I can easily tell when other people are upset or have their feelings hurt, even when they say nothing.		x	
Item 12. I have my share of responsibilities in my family.			x
Item 13. I can generally avoid negative situations.			x
Item 14. It bothers me when others have their feelings hurt.		x	
Item 15. When I try something, I usually succeed.			x
Item 16. I am good at problem solving.			x
Item 17. I am generally good at making decisions.			x
Item 18. I always make an attempt to be punctual (on time).			x
Item 19. I am reliable.	x		
Item 20. I can do things well.			x

	Self-Concept	Emotional Intelligence	Self-Efficacy
Item 21. I am content with who I am.	x		
Item 22. When I need something, I can usually get it.			x
Item 23. I feel a close bond with my friends and family.		x	
Item 24. I enjoy working with others.		x	
Item 25. I believe that people are generally trustworthy.		x	

Table 4.

Regression Equations for the Initial EFA

YLQ	Factor Loading * Factor	Error Variance	Ry ²
Item1	.66*F1	.18	.50
Item2	.68*F2	.23	.45
Item3	.74*F1	.45	.30
Item4	.62*F1	.25	.36
Item5	.52*F1	.29	.23
Item6	.51*F2	.27	.24
Item7	.54*F1	.51	.13
Item9	.65*F2	.45	.23
Item10	.67*F2	.27	.38
Item11	.36*F2	.44	.05
Item12	.53*F2	.39	.18
Item13	.50*F3	.78	.06
Item14	.50*F2	.63	.08
Item15	.56*F3	.47	.16
Item16	.66*F3	.32	.32
Item17	.74*F3	.21	.52
Item18	.62*F3	.51	.18
Item19	.39*F1	.24	.15
Item20	.43*F3	.26	.18
Item21	.51*F1	.25	.26
Item22	.54*F3	.42	.18
Item23	.59*F2	.29	.30
Item24	.53*F2	.45	.15
Item25	.51*F2	.52	.12

Table 5.

Factor Loadings for the Initial EFA

YLQ	Factor 1	Factor 2	Factor 3
Item1	.66		
Item2		.68	
Item3	.74		
Item4	.62		
Item5	.52		
Item6		.51	
Item7	.54		
Item9		.65	
Item10		.67	
Item11		.36	
Item12		.53	
Item13			.50
Item14		.50	
Item15			.56
Item16			.66
Item17			.74
Item18			.62
Item19	.39		
Item20			.43
Item21	.51		
Item22			.54
Item23		.59	
Item24		.53	
Item25		.51	

Table 6.

Factor Loadings for the EFA on the Revised YLQ

YLQ	Factor 1	Factor 2	Factor 3
Item1	.66		
Item2		.68	
Item3	.74		
Item4	.62		
Item5	.52		
Item7	.53		
Item9		.65	
Item10		.68	
Item13			.52
Item14		.48	
Item15			.59
Item16			.68
Item17			.72
Item20			.45
Item21	.52		
Item22			.56
Item23		.58	
Item24		.53	

Table 7.

Regression Equations for the EFA on the Revised YLQ

YLQ	Factor Loading * Factor	Error Variance	Ry ²
Item1	.66*F1	.18	.50
Item2	.68*F2	.22	.46
Item3	.74*F1	.44	.31
Item4	.62*F1	.24	.37
Item5	.52*F1	.29	.24
Item7	.53*F1	.51	.13
Item9	.65*F2	.45	.24
Item10	.68*F2	.26	.41
Item13	.52*F3	.75	.07
Item14	.48*F2	.65	.07
Item15	.59*F3	.44	.19
Item16	.68*F3	.29	.37
Item17	.72*F3	.24	.46
Item20	.45*F3	.25	.19
Item21	.52*F1	.24	.28
Item22	.56*F3	.40	.19
Item23	.58*F2	.30	.28
Item24	.53*F2	.46	.14

Table 8.

Factor Loadings for the CFA

YLQ	Factor 1	Factor 2	Factor 3
Item1	.47		
Item2		.52	
Item3	.49		
Item4	.56		
Item5	.47		
Item7	.26		
Item9		.48	
Item10		.58	
Item13			.52
Item14		.32	
Item15			.77
Item16			.73
Item17			.75
Item20			.55
Item21	.51		
Item22			.62
Item23		.51	
Item24		.54	

Table 9.

Regression Equations for the CFA

YLQ	Factor Loading * Factor	Error Variance	Ry ²
Item1	.47*F1	.20	.28
Item2	.54*F2	.27	.27
Item3	.49*F1	.39	.14
Item4	.56*F1	.49	.16
Item5	.47*F1	.24	.23
Item7	.26*F1	.31	.03
Item9	.48*F2	.53	.10
Item10	.58*F2	.29	.29
Item13	.52*F3	.66	.08
Item14	.32*F2	.46	.04
Item15	.77*F3	.36	.38
Item16	.73*F3	.23	.48
Item17	.75*F3	.20	.53
Item20	.55*F3	.20	.35
Item21	.51*F1	.17	.36
Item22	.62*F3	.31	.30
Item23	.51*F2	.23	.28
Item24	.54*F2	.39	.18

Table 10.

Factor Loadings for the Control Group

YLQ	Factor 1	Factor 2	Factor 3
Item1	.88		
Item2		.92	
Item3	.88		
Item4	.95		
Item5	.93		
Item7	.57		
Item9		.59	
Item10		.95	
Item13			.65
Item14		.65	
Item15			.93
Item16			.87
Item17			.96
Item20			.85
Item21	.90		
Item22			.79
Item23		.68	
Item24		.84	

Table 11.

Regression Equations for the Control Group

YLQ	Factor Loading * Factor	Error Variance	Ry ²
Item1	.88*F1	.41	.42
Item2	.92*F2	.68	.31
Item3	.88*F1	2.05	.07
Item4	.95*F1	.74	.30
Item5	.93*F1	.63	.34
Item7	.57*F1	.93	.07
Item9	.59*F2	1.22	.05
Item10	.95*F2	.49	.42
Item13	.65*F3	1.04	.08
Item14	.65*F2	1.88	.04
Item15	.93*F3	.67	.31
Item16	.87*F3	.68	.28
Item17	.96*F3	.56	.38
Item20	.85*F3	.65	.28
Item21	.90*F1	1.00	.20
Item22	.79*F3	.67	.23
Item23	.68*F2	1.00	.10
Item24	.84*F2	.96	.18

Table 12.

Coefficient Alpha for the YLQ

YLQ	Total	Self-Concept	Self-Efficacy	Emotional intelligence
Coefficient Alpha	.95	.87	.89	.86

Table 13.

Inter-Item Correlations for the YLQ

YLQ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Item1	1.0																	
Item2	.63	1.0																
Item3	.60	.56	1.0															
Item4	.67	.56	.66	1.0														
Item5	.66	.53	.50	.56	1.0													
Item6	.56	.43	.42	.48	.45	1.0												
Item7	.59	.52	.50	.55	.43	.52	1.0											
Item8	.62	.70	.57	.61	.60	.47	.58	1.0										
Item9	.46	.34	.40	.40	.39	.34	.40	.39	1.0									
Item10	.39	.52	.37	.42	.32	.43	.38	.44	.35	1.0								
Item11	.56	.43	.51	.58	.48	.40	.39	.51	.48	.35	1.0							
Item12	.57	.45	.58	.61	.53	.45	.54	.53	.53	.38	.67	1.0						
Item13	.63	.51	.54	.62	.57	.51	.53	.57	.57	.43	.63	.76	1.0					
Item14	.60	.51	.54	.59	.60	.42	.47	.60	.40	.35	.61	.58	.63	1.0				
Item15	.63	.52	.54	.56	.65	.38	.42	.59	.36	.36	.50	.49	.56	.63	1.0			
Item16	.61	.52	.44	.48	.54	.42	.46	.52	.33	.36	.39	.40	.48	.51	.60	1.0		
Item17	.53	.61	.44	.50	.43	.32	.38	.57	.33	.45	.40	.41	.39	.49	.56	.57	1.0	
Item18	.59	.46	.52	.55	.56	.49	.46	.53	.42	.31	.62	.59	.62	.63	.60	.51	.42	1.0

Table 14.

Item Total Correlations for the YLQ

YLQ	Item Total Correlation
Item1	.80
Item2	.71
Item3	.70
Item4	.76
Item5	.70
Item6	.60
Item7	.65
Item8	.76
Item9	.55
Item10	.53
Item11	.69
Item12	.74
Item13	.77
Item14	.74
Item15	.72
Item16	.65
Item17	.62
Item18	.72

Table 15.

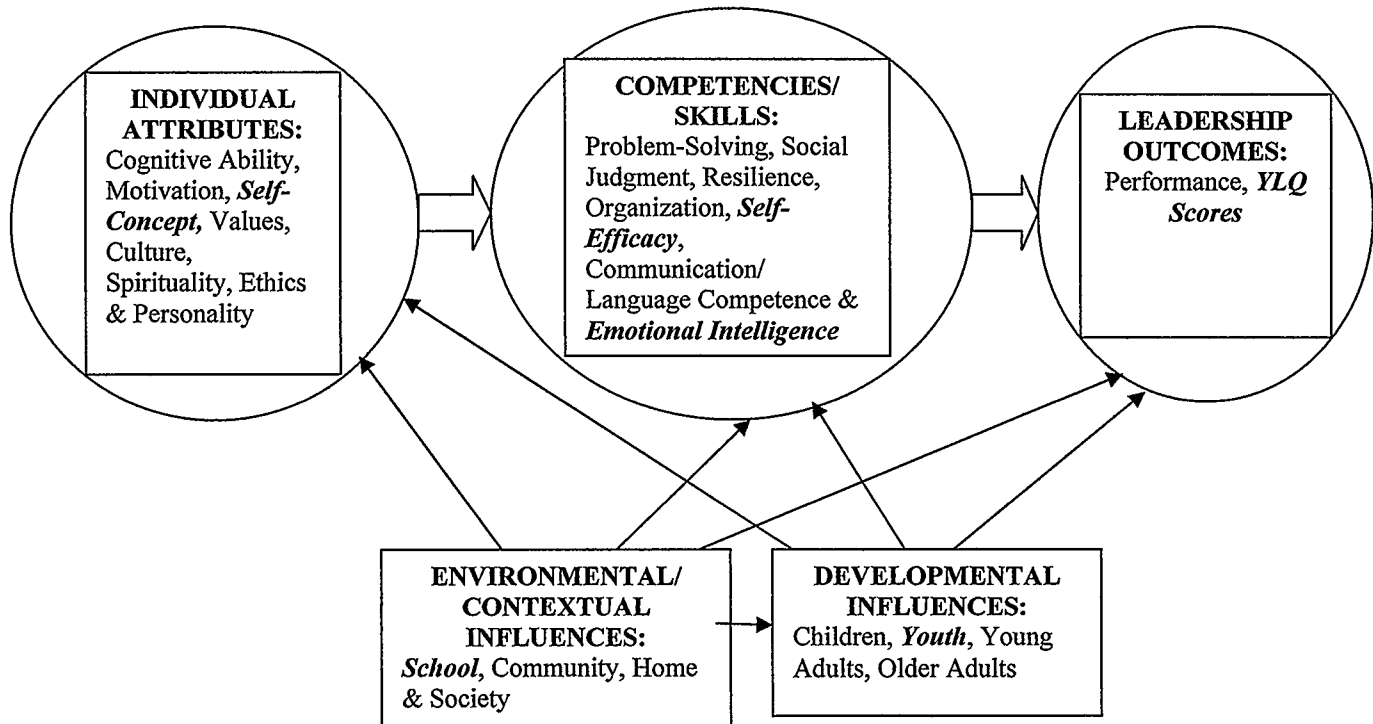
Construct Validity Correlations of the YLQ With Other Measures

	Pearson Correlation		
	YLQ	SWLS	RSA
YLQ	1.0		
SWLS	.53**	1.0	
RSA	.49**	.44**	1.0

** Significant at .01 level

Figure 1.

Conceptual Model of Youth Leadership



(Adapted from “Leadership skills for a changing world: Solving complex social problems,” by M.D. Mumford, S.J. Zaccaro, F.D. Harding, T.O. Jacobs, and E.A. Fleishman, 2000, *Leadership Quarterly*, 11, 23).

* The items in bold and italics represent those areas used in this study.

Appendix A: The YLQ

Initial Youth Leadership Questionnaire Prior to Pilot Testing

1. People think that I am fun to be with.
2. I cannot easily use different ways of solving a problem.
3. I really care about my family.
4. I feel that most people do not respect me.
5. I can make good things happen.
6. I am good at things.
7. I know when other people are upset, even when they say nothing.
8. I know that I can count on the rest of my family.
9. I am successful at most things.
10. I can get the things that I need.
11. I am not good at making decisions.
12. I like doing things for others.
13. I do my share of duties in the family.
14. I enjoy life.
15. If at first I don't succeed, I will keep on trying.
16. I am liked by others.
17. I feel bad when other people have their feelings hurt.
18. My family does not know what I mean when I say something.
19. I am proud of myself.
20. Good things will not happen to me.
21. I am someone you can rely on.
22. I can tell when one of my close friends is unhappy.
23. What I expect from my family is fair.
24. I am happy with myself just the way that I am.
25. My life will be happy.

The Youth Leadership Questionnaire (Revised After Pilot Testing)

1. I will keep on trying, even if I am not successful the first time.
2. I always make an effort to make new people feel welcome.
3. I feel comfortable speaking in public (in front of an audience).
4. I enjoy new challenges.
5. I am willing to stand up for myself.
6. I try to be sensitive to others' values, beliefs and/or opinions.
7. I am aware that my decisions have consequences.
8. My fears limit me.
9. I can effectively use others' feedback to improve myself.
10. I make an effort to effectively communicate with others.
11. I can easily tell when other people are upset or have their feelings hurt, even when they say nothing.
12. I have my share of responsibilities in my family.
13. I can generally avoid negative situations.
14. It bothers me when others have their feelings hurt.
15. When I try something, I usually succeed.
16. I am good at problem solving.
17. I am generally good at making decisions.
18. I always make an attempt to be punctual (on time).
19. I am reliable.
20. I can do things well.
21. I am content with who I am.
22. When I need something, I can usually get it.
23. I feel a close bond with my friends and family.
24. I enjoy working with others.
25. I believe that people are generally trustworthy.

The Youth Leadership Questionnaire- Revised Version Used in Analyses

1. I will keep on trying, even if I am not successful the first time.
2. I always make an effort to make new people feel welcome.
3. I feel comfortable speaking in public (in front of an audience).
4. I enjoy new challenges.
5. I am willing to stand up for myself.
7. I am aware that my decisions have consequences.
9. I can effectively use others' feedback to improve myself.
10. I make an effort to effectively communicate with others.
13. I can generally avoid negative situations.
14. It bothers me when others have their feelings hurt.
15. When I try something, I usually succeed.
16. I am good at problem solving.
17. I am generally good at making decisions.
20. I can do things well.
21. I am content with who I am.
22. When I need something, I can usually get it.
23. I feel a close bond with my friends and family.
24. I enjoy working with others.

Appendix B: Standardized Residuals for Each of the Models

Initial EFA Model Standardized Residuals

YLQ	1	2	3	4	5	6	7	9	10	11	12
Item1	1.0										
Item2	.39	1.0									
Item3	-1.06	2.90	1.0								
Item4	.26	-.30	.28	1.0							
Item5	1.62	-.08	-.18	-.80	1.0						
Item6	-.06	.19	-.57	-.98	-.97	1.0					
Item7	.60	.90	.67	-1.52	-.55	1.39	1.0				
Item9	.70	-1.59	-1.06	.00	.52	1.09	.97	1.0			
Item10	-.27	1.47	-.61	-.91	.54	.03	.70	2.55	1.0		
Item11	-1.81	.80	-.46	-.99	-.21	.01	1.37	.07	1.59	1.0	
Item12	1.42	.29	-.37	.36	1.15	-2.82	-.89	-1.66	-.88	1.46	1.0
Item13	-1.59	-1.60	-.02	-.96	-1.04	-.64	-1.67	.59	-1.43	.35	.76
Item14	-1.59	1.46	-.45	.24	-1.36	3.84	1.77	.79	-1.23	2.24	.32
Item15	-2.72	-2.21	-1.01	-.53	-4.25	-2.86	-.49	-2.20	-1.82	-1.45	1.24
Item16	-3.34	-.76	.89	.34	-2.40	-4.17	-2.63	-.19	-2.17	.86	1.85
Item17	-.56	.41	1.80	1.36	.41	-2.25	-.67	.89	-1.30	.15	3.27
Item18	3.19	4.84	3.52	.76	-.11	1.79	3.14	1.55	2.63	3.06	3.28
Item19	.30	-1.32	-1.61	-.48	-.07	2.35	.33	.38	-1.55	-.65	-.30
Item20	-1.28	-.63	-.36	.97	-.29	-.29	-2.12	.90	.47	-1.08	2.20
Item21	.70	-1.07	.34	.97	.98	-1.51	-.59	.28	-.26	-1.22	-1.77
Item22	.49	.77	.16	1.57	.51	-1.26	1.82	.90	1.67	-.98	1.23
Item23	.58	-1.90	-1.25	-1.24	1.31	.40	.08	-.22	.47	-1.43	1.69
Item24	.70	.45	-1.33	1.85	-1.03	-.84	-.83	-1.46	-.51	-1.14	.25
Item25	.62	-2.20	-.86	-.32	1.64	.61	-.46	-1.21	-3.03	-1.12	.67

Revised EFA Model Standardized Residuals

YLQ	1	2	3	4	5	7	9	10	13	14	15
Item1	1.0										
Item2	.24	1.0									
Item3	-1.18	2.77	1.0								
Item4	.19	-.67	.02	1.0							
Item5	1.69	-.10	-.31	-.96	1.0						
Item7	.81	.91	.70	-1.51	-.48	1.0					
Item9	.60	-2.03	-1.34	-.25	.40	.52	1.0				
Item10	-.76	.91	-1.16	-1.55	.22	.59	2.30	1.0			
Item13	-1.36	-1.41	.02	-.87	-.94	-1.54	.70	-1.32	1.0		
Item14	-1.01	2.10	-.11	.63	-1.02	2.06	1.16	-.87	.84	1.0	
Item15	-2.56	-2.08	-1.04	-.56	-4.20	-.42	-2.12	-1.82	.78	-.61	1.0
Item16	-2.67	-.25	1.17	.68	-2.04	-2.24	.17	-1.87	.61	.21	2.56
Item17	1.40	1.85	2.82	2.58	1.47	.35	1.82	-.04	1.75	2.48	-3.14
Item20	-.77	-.21	-.10	1.22	-.02	-1.79	1.19	.76	-.33	-.77	3.06
Item21	.25	-1.73	-.14	.48	.67	-.76	-.11	-1.02	-1.42	-1.03	-1.41
Item22	.89	1.14	.39	1.79	.75	2.03	1.16	1.94	-1.52	-1.53	1.81
Item23	.99	-1.66	-1.14	-1.10	1.50	.32	-.03	.54	-1.56	-.26	-2.48
Item24	.87	.54	-1.36	1.91	-.99	-.72	-1.46	-.66	-.42	-2.42	-.13
YLQ	16	17	20	21	22	23	24				
Item16	1.0										
Item17	1.24	1.0									
Item20	-.67	-1.98	1.0								
Item21	-.98	-.73	2.18	1.0							
Item22	-1.32	-1.70	.26	2.08	1.0						
Item23	-2.89	-.54	.10	1.68	1.64	1.0					
Item24	1.29	-.02	3.13	2.54	1.77	.91	1.0				

CFA Model Standardized Residuals

YLQ	1	2	3	4	5	7	9	10	13	14	15
Item1	1.0										
Item2	.74	1.0									
Item3	.90	.21	1.0								
Item4	1.19	-.75	3.65	1.0							
Item5	-1.11	-.60	-1.57	-2.23	1.0						
Item7	-.12	1.25	-1.25	-.12	-.37	1.0					
Item9	-.17	-.47	.22	1.53	-.63	1.16	1.0				
Item10	-2.57	4.21	-.02	-.01	.16	.96	3.09	1.0			
Item13	-.01	-.80	.65	-.33	.12	2.41	1.98	1.23	1.0		
Item14	-2.13	1.26	-.45	-.59	-.52	2.58	-1.07	.78	2.29	1.0	
Item15	-1.58	-2.73	.74	.25	-.57	-1.25	1.68	.11	-.22	-1.10	1.0
Item16	-1.93	-3.12	.97	.52	.57	-.65	.56	-.11	-.51	-.07	1.74
Item17	-.99	-4.01	-1.14	-.02	1.53	1.50	.58	.73	.94	.12	-.45
Item20	.18	.54	.00	1.53	1.22	.59	2.31	2.60	-.53	.45	-.79
Item21	2.00	-1.52	-.97	-1.81	1.98	-2.18	-1.42	-1.56	-.26	-.52	-1.82
Item22	.18	-1.41	-1.12	.88	2.19	.83	1.11	1.44	-.77	.58	1.73
Item23	2.66	-1.04	-.19	-1.12	1.66	-1.41	-1.55	-4.68	.69	-1.25	-.83
Item24	-.11	-.29	-.59	-.18	-1.36	.70	-1.63	-.12	1.01	2.10	-1.70
YLQ	16	17	20	21	22	23	24				
Item16	1.0										
Item17	2.75	1.0									
Item20	-1.21	-1.48	1.0								
Item21	-2.86	.24	2.55	1.0							
Item22	-1.63	-1.28	.83	.29	1.0						
Item23	1.04	1.19	3.44	4.02	.51	1.0					
Item24	-1.21	-2.79	1.75	1.12	.93	1.39	1.0				

Control Model Standardized Residuals

YLQ	1	2	3	4	5	7	9	10	13	14	15
Item1	1.0										
Item2	-.55	1.0									
Item3	-.54	-.17	1.0								
Item4	-.37	-.88	3.48	1.0							
Item5	1.24	-1.64	-1.29	-1.26	1.0						
Item7	2.03	-2.13	-2.48	.16	.19	1.0					
Item9	3.53	-.15	.78	2.20	-.84	3.94	1.0				
Item10	.08	.46	1.53	1.44	1.20	-.31	.04	1.0			
Item13	.75	-1.16	-.81	-1.08	-.58	-.08	1.15	-.30	1.0		
Item14	-1.43	2.34	-1.94	-.11	-1.93	1.45	-.47	-1.58	1.03	1.0	
Item15	.36	-1.09	.53	2.00	-.51	-.26	-.57	1.67	.60	.96	1.0
Item16	-1.98	-2.59	.94	.76	-1.05	1.19	3.78	.47	1.97	.02	.95
Item17	-1.73	-.93	-1.11	-.09	-2.25	1.40	2.55	.80	2.73	.86	-1.46
Item20	1.13	.59	.91	.68	2.68	.29	.95	3.74	-3.29	-.10	-1.09
Item21	-1.15	-1.13	-.39	-1.76	3.28	-3.09	-1.62	1.74	-1.78	-1.09	-.34
Item22	-.09	-1.72	.17	-1.54	.70	1.53	.45	.01	-1.06	-1.27	.42
Item23	2.34	-1.66	-1.09	-.75	1.43	.88	.07	-2.53	.01	-.70	-.32
Item24	-1.52	2.63	-1.44	-1.44	-2.44	-3.24	-3.08	-.69	-1.14	2.31	-1.06

YLQ	16	17	20	21	22	23	24
Item16	1.0						
Item17	4.57	1.0					
Item20	-3.18	-.64	1.0				
Item21	-3.12	-.87	3.03	1.0			
Item22	-1.85	-1.15	2.32	2.77	1.0		
Item23	-1.60	.38	2.03	3.40	2.45	1.0	
Item24	-3.03	-3.24	.32	1.49	-2.27	2.86	1.0



UNIVERSITY OF
CALGARY

CERTIFICATION OF INSTITUTIONAL ETHICS REVIEW

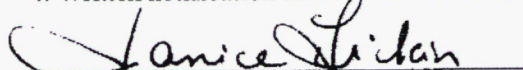
This is to certify that the Conjoint Faculties Research Ethics Board at the University of Calgary has examined the following research proposal and found the proposed research involving human subjects to be in accordance with University of Calgary Guidelines and the Tri-Council Policy Statement on *"Ethical Conduct in Research Using Human Subjects"*. This form and accompanying letter constitute the Certification of Institutional Ethics Review.

File no: **5844**
Applicant(s): **Anne M. McKeough**
David Nordstokke
Yvonne L. Hindes
Vicki L. Schwean
Department: **Applied Psychology, Division of**
Project Title: **Instrument Development/Program Evaluation for the Werklund Foundation**
Sponsor (if applicable): **SSHRC**

Restrictions:

This Certification is subject to the following conditions:

1. Approval is granted only for the project and purposes described in the application.
2. Any modifications to the authorized protocol must be submitted to the Chair, Conjoint Faculties Research Ethics Board for approval.
3. A progress report must be submitted 12 months from the date of this Certification, and should provide the expected completion date for the project.
4. Written notification must be sent to the Board when the project is complete or terminated.


Janice Dickin, Ph.D, LLB,
Chair
Conjoint Faculties Research Ethics Board

5 November 2008
Date:

Distribution: (1) Applicant, (2) Supervisor (if applicable), (3) Chair, Department/Faculty Research Ethics Committee, (4) Sponsor, (5) Conjoint Faculties Research Ethics Board (6) Research Services.