THE MEASUREMENT OF QUALITY OF LIFE AND ITS RELATIONSHIP WITH PERCEIVED HEALTH STATUS IN ADOLESCENTS

by

Richard Sawatzky

MSN, University of British Columbia, 2002
BScN, Trinity Western University, 1999

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Nursing)

August, 2007

ABSTRACT

Several assumptions of the indirect reflective model of the Multidimensional Students' Life Satisfaction Scale (MSLSS) were tested to assess its validity as a measure of adolescents' satisfaction with life generally and with five important life domains (family, friends, living environment, school, and self perception). We also examined whether adolescents' perceived mental and physical health status significantly explained their global quality of life (QOL) and whether these relationships were mediated by their satisfaction with the five life domains.

The data were taken from a cross-sectional health survey of 8,225 adolescents in 49 schools in British Columbia, Canada. Global QOL was measured using Cantril's ladder and a single-item rating of the adolescents' satisfaction with their QOL. Confirmatory factor and factor mixture analyses were used to examine the measurement assumptions of the MSLSS, and structural equation modeling was applied to test the hypothesized mediation model. The Pratt index (*d*) was used to evaluate variable importance.

The adolescents did not respond to all MSLSS items in a consistent manner. An abridged 18-item version of the MSLSS was therefore developed by selecting items that were most invariant in the sample. Good model fit was obtained when the abridged MSLSS was used to test the hypothesized mediation model, which explained 76% of the variance in global QOL. Relatively poorer mental health and physical health were significantly associated with lower satisfaction in each of the life domains. Global QOL was predominantly explained by the adolescents' mental health status (d = 30%) and by their satisfaction with self (d = 42%) and family (d = 20%). Self and family satisfactionwere the

predominant mediating variables of the relationships between mental health (45% total mediation) and physical health (68% total mediation) and global QOL.

Satisfaction with life domains and perceived physical and mental health can be viewed as conditions that potentially contribute to adolescents' global QOL. Questions about adolescents' experiences with important life domains require more attention in population health research so as to target appropriate supportive services for adolescents, particularly those with mental or physical health challenges.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BIC	Bayesian information criterion
CFA	Confirmatory factor analysis
CFI	Comparative fit index
Df	Degrees of freedom
EM	Expectation maximization
FIML	Full information maximum likelihood
FMA	Factor mixture analysis
LL	Log likelihood
MAR	Missing at random
MCAR	Missing completely at random
MeSH	Medical subject heading
MI	Multiple imputation
ML	Maximum likelihood
MLM	Maximum likelihood estimation, mean adjusted
MSLSS	Multidimensional Students' Life Satisfaction Scale
RMSEA	Root mean square error of approximation
SEM	Structural equation modeling
SRMR	Standardized root mean residual
WLF	Worst linear function
WLSMV	Weighted least squares estimation, mean and variance adjusted
WRMR	Weighted root mean residual

ACKNOWLEDGEMENTS

Few things in life are achieved through the efforts of one person alone. The work of this dissertation was accomplished with a great deal of practical and emotional support that I received from my family, friends, colleagues, and mentors. I have often wondered why research on quality of life generally does not engage with what I consider to be the most profound question in life: Do you know that you are loved? The support that I received from Helma, my most intimate friend and companion in life, my parents and parents-in-law, my sister and sister-in-law, and my friends has enabled me to confidently answer this question. I most certainly could not have completed this work without their love and support.

The work for this dissertation was completed in an intellectually stimulating community of inspirational and gifted people from diverse disciplinary backgrounds. The discussions that took place have profoundly influenced my thoughts about nursing, health care, quality of life, and academics. I deeply value the dialogue with my supervisor, Dr. Pamela A. Ratner, who has been exceedingly generous in sharing her expertise and time. And I am grateful for the contributions from my other committee members: Dr. Joy L. Johnson, Dr. Bruno D. Zumbo, and Dr. Jacek Kopec. Together, they created a learning environment that has far exceeded my expectations. I have written the dissertation in third person with the intent to acknowledge our shared contributions to this work.

My thinking about quality of life and health has been deeply influenced by many other people including students in the doctoral program in the School of Nursing at the University of British Columbia, and my colleagues and friends at Trinity Western University. I specifically want to acknowledge Dr. Chris Richardson at UBC who has greatly influenced my thinking and who has consistently encouraged me to complete the analyses, and Dr. Barb

Pesut at TWU, with whom I engaged in many philosophical debates about measurement and quality of life.

I am also very grateful for the significant financial contributions that I received from the Canadian Institutes for Health Research, the Michael Smith Foundation for Health Research, the Canadian Nurses Foundation, NEXUS, the UBC School of Nursing, and the faculty of applied science at Trinity Western University. This practical support offered me the opportunity to network with colleagues through several international conferences without having to secure significant other sources of income, and it enabled me to set aside a significant amount of time to focus on the research that culminated in this dissertation.

1 INTRODUCTION AND BACKGROUND

An understanding of how various aspects of people's lives might be affected by illness, disease, medical and other therapeutic interventions, and, more generally, health policy is central to the practice of nurses and other health-care professionals. Although nursing practice involves interventions that have as a goal to specifically improve the chances of survival while minimizing the adverse physical and psychological manifestations of disease, many nursing interventions are more general in nature and address various other aspects of life. Perusal of nursing classification systems of nursing diagnoses, interventions and outcomes reveals numerous interventions that target broadly-defined goals such as enhanced wellbeing and quality of life (M. Johnson & North American Nursing Diagnosis Association, 2001). Similarly, nursing theorists have extended the focus of nursing care far beyond specific mental and physical manifestations of disease by emphasizing more general considerations of how people are affected in various areas of life by health challenges and illness. Questions about an individual's physical abilities, psychological experiences, social involvements, economic status, environmental conditions, and even spiritual experiences have been included in comprehensive health assessment frameworks as the basis for targeting appropriate supportive services and determining the efficacy of health-care interventions. Comparable questions have been included in population health surveys to inform the development of health policies that address a wide range of issues such as those pertaining to equitable access to health care, supportive services, and health promotion initiatives.

The terms *quality of life* or *health-related quality of life* are frequently used in health research to refer to this assortment of questions for the measurement of conditions and experiences in various areas of life as the basis for examining the consequences of disease and illness, and the effectiveness of health-care practices and interventions (Bowling, 2005; Fayers & Machin, 2007; Ferrans, 2005; Grant & Dean, 2003; Padilla, Frank-Stromborg, & Setsuko, 2004; Padilla & Grant, 1985; Spilker, 1996). Often, measures of specific health outcomes,[1] such as symptoms, functional status, and perceived health status, are combined with measures of life satisfaction, wellbeing, and happiness in instruments for the measurement of quality of life. Factor analysis techniques are generally used for the purpose of clustering these diverse measures so as to obtain summary scores that represent so-called *dimensions* of quality of life. These dimensions have been variously labeled as physical, psychological, social, economic, environmental, and spiritual or existential life domains (Ferrans, 2005; King & Hinds, 2003; Padilla et al., 2004; Spilker, 1996). For example, Ferrell et al. (1995) developed a quality of life model that encompasses physical wellbeing and symptoms, psychological wellbeing, social wellbeing, and spiritual wellbeing, all of which are considered relevant to the experience of living with cancer. Not surprisingly, quality of life is widely referred to as a multidimensional construct (Fayers & Machin, 2007; Rapley, 2003).

Clearly, quality of life has become an important concept in health research. However, the use of the term *quality of life* to refer to the measurement of health outcomes, as described above, might not be obvious to those who are not familiar with this field of

[1] We are using the term *health outcomes* to refer to a wide range of possible variables used to assess the impact of disease and illness, and health-care practices and interventions on various aspects of people's lives. Health outcomes do not refer to any particular concept, but rather encompass many different concepts including symptoms, physical and psychological functioning, and perceived health status. Although quality of life is frequently referred to as a health outcome (Bowling, 2002, 2005; Lipscomb, Gotay, & Snyder, 2005), we prefer to distinguish these terms to avoid unnecessarily conflating health with quality of life. Rather, we see quality of life as a person's appreciation of his or her life as a whole (Veenhoven, 2000), which may be affected by various health outcomes.

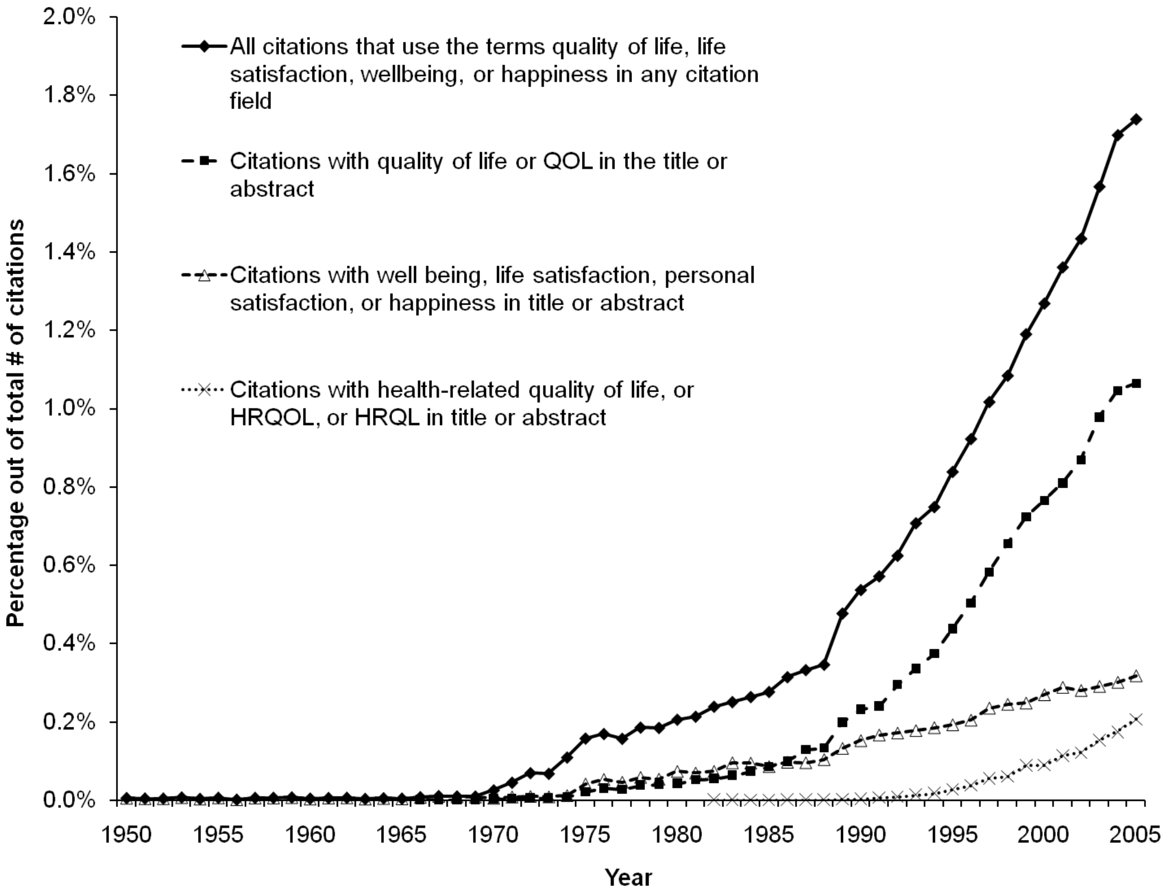
research. In contrast, those who are immersed in this field of research might find it difficult to relate their focus on health outcomes to broader theoretical understandings of quality of life. We therefore provide a brief historical synopsis of the use of the term *quality of life* in health research in the following paragraphs. We then continue with a discussion of several conceptual models pertaining to the relationship between health and quality of life.

1.1 Quality of life in health research

The focus on the measurement of quality of life in health research can be traced back to a historical transition from a predominantly curative focus of medical care in the first half of the 20th century to an emphasis on the side-effects of medical treatments and the impact of disease and illness on physical, social, and emotional wellbeing (Musschenga, 1997). This transition is exemplified in the following well-known definition of health offered by the World Health Organization (WHO) (1946): “Health is a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity” (p. 100). However, although this definition of health is widely cited as providing the impetus for the measurement of quality of life in health research, the term *quality of life* was initially not frequently used for the purposes of health outcomes measurement. Rather, the WHO’s definition was first followed by studies on the relationships between health, wellbeing, life satisfaction and happiness, and, until 1985, these terms were more frequently encountered in citations indexed in PubMed (National Center for Biotechnology Information, 2005) than was the term *quality of life* (see Figure 1).

In PubMed (National Center for Biotechnology Information, 2005), most of the studies pertaining to life satifaction or wellbeing were indexed under the medical subject heading (MeSH) “personal satisfaction,” which was defined in 1970 as “the individual's

Figure 1 Percentage of citations indexed in PubMed from 1950 to 2005 that used the terms quality of life, life satisfaction, wellbeing, or happiness



Notes: Based on the number of citations indexed in Pubmed from 1900 to 2005 as of October 8, 2005 (N = 15,830,354).

experience of a sense of fulfillment of a need or want and the quality or state of being satisfied" (National Library of Medicine, 2005). The term "life satisfaction" was first used in a PubMed citation report pertaining to an influential publication by Neugarten, Havigurst, and Tobin (1961) entitled "The Measurement of Life Satisfaction," which was published in the *Journal of Gerontology*. The authors specifically reported on the development of two instruments for the measurement of life satisfaction.

In contrast to the emergence of life satisfaction as a measurable concept in health research, the publications that referred to quality of life did not initially focus on measurement, but on ethical issues in health care. For example, the first publication in PubMed (National Center for Biotechnology Information, 2005) that included the words *quality of life* in the title was an editorial by Long (1960) entitled "On the Quantity and Quality of Life," and the next citation pertained to a publication by Elkinton (1966) on "Medicine and the Quality of Life" published in the *Annals of Internal Medicine*. Both authors wrote about ethical decisions with respect to life-prolonging medical treatments. Later, in 1975, the term *quality of life* was introduced in PudMed as a medical subject heading with the following description: "[Quality of life is] a generic concept reflecting concern with the modification and enhancement of life attributes, e.g., physical, political, moral and social environment; the overall condition of a human life" (National Library of Medicine, 2005). Since then, the term *quality of life* has increasingly been used to refer to the measurement of broadly defined health outcomes. After 1985, the frequency of the term *quality of life* occurring in the title or abstract of citations in PubMed exceeded that of the terms life satisfaction, wellbeing, or happiness, and the term quality of life has been used in more than 1% of all citations that were indexed in PubMed in 2004 and 2005 (see Figure 1). Thus, the use of the term *quality of life* to refer to the measurement of health outcomes is a fairly recent phenomenon in health research, and the initial focus that arose from the WHO definition of health was on the measurement of life satisfaction and wellbeing, and not, explicitly on the measurement of quality of life. The initial focus on life satisfaction and wellbeing continues to be reflected in the definitions of quality of life and health-related quality of life that are frequently encountered in health research publications. For example, Ferrans (2005) provided an overview of 15 definitions, of which 11 definitions explicitly include the terms wellbeing or life satisfaction.

Even though the term *quality of life* is now widely used, there exists considerable ambiguity about its meaning. In congruence with the previously mentioned description of the medical subject heading for quality of life, the term is now generally used to refer to a wide range of broadly defined health outcomes, such as symptoms, functional status, and perceived health status, as well as global appraisals of life satisfaction, happiness, and wellbeing pertaining to life as a whole (Bowling, 2005; Fayers & Machin, 2007; Ferrans, 2005; Grant & Dean, 2003; Padilla et al., 2004; Padilla & Grant, 1985; Spilker, 1996). The supposed all-embracing nature of quality of life, and the resulting lack of clearly delineated conceptual boundaries, is exemplified in the following quotation, which was taken from the introductory chapter of a book on quality of life research by Fayers and Machin (2007):

> In this book we shall use the now well-established term *quality of life*, and its common abbreviation QoL. By that, we include general questions such as 'How good is your overall quality of life?' or 'How do you rate your overall health?' that represent global assessments; dimensions such as pain or fatigue; symptoms such as headaches or skin irritation; function, such as social and role functioning; issue such as body image or existential beliefs; and so on. (pp. 3 – 4)

This quotation illustrates that quality of life in health research is now meant to refer to almost any aspect of a person's life that may be affected by illness and healthcare interventions. However, the quotation also demonstrates the conceptual ambiguity with respect to the distinction between health outcomes, such as symptoms, functional status, and perceived health status, and global notions of quality of life such as those pertaining to life satisfaction and wellbeing.

1.2 The relationship between quality of life and health

Despite this ambiguity in the conceptualization of quality of life, some researchers have developed conceptual models in an attempt to describe the relationships between quality of life and health (Ferrans, 2005; Vallerand & Payne, 2003). These models generally imply that the presence of disease results in symptoms that affect various so-called dimensions of quality of life, such as physical, psychological, and social functioning, which in turn contribute to overall quality of life (e.g., Burckhardt, 1985; Padilla & Grant, 1985; Patrick & Chiang, 2000; Wilson & Cleary, 1995). Most models also account for the presence of a variety of psychological processes (e.g., coping, adaptation and personality) and social, cultural and environmental factors. For example, Wilson and Cleary (1995) described a model wherein physiological and psychological symptoms affect functional status, which affects general health perceptions and quality of life. Concepts pertaining to characteristics of the individual (e.g., motivation and values) and characteristics of the environment (e.g., social support) were also included in their model.

Other researchers have sought to provide empirical support for similar models of the relationship between health and quality of life. Based on a meta-analysis of studies that used instruments measuring various health status indicators, Smith, Avis, and Assmann (1999) showed that variation in quality of life was explained by variables pertaining to various life domains, which were affected by differences in physiological health status (e.g., the presence of disease) and symptom severity. Examples of life domains in their meta-analysis include variables that reflected psychological, social, or physical functioning. Quality of life was represented by measures of life satisfaction, wellbeing and single-item quality of life indicators. Thus, their “model of the determinants of quality of life” (p. 448) is based on the proposition that the life domains mediate the degree to which quality of life is explained by differences in symptom severity and physiological health status. Although mental health status and physical function were both fairly strongly correlated with life satisfaction (r = 0.68 and r = 0.57, respectively), regression of quality of life on mental health, physical function, and social function revealed that mental health status was by far the most important explanatory variable ($\beta = 0.47$). The authors concluded that health status is conceptually distinct from quality of life.

Similarly, Beckie and Hayduk (2004) used structural equation modeling to provide empirical support for modeling various health outcomes as causal variables that affect quality of life. In doing so, they clearly distinguished global notions of quality of life from particular health outcomes. Health outcomes, such as pain, and physical, social and emotional functioning, were measured using the Short-Form 36-item instrument (SF-36) (Ware, Snow, Kosinski, & Gandek, 1993). Based on a study of 306 people who underwent coronary artery bypass graft surgery, they found that the measured health outcomes explained 67% of the variance in quality of life, and that the effects of general health perceptions ($\beta = 0.47$) and mental health ($\beta = 0.46$) were most substantial. They concluded that “quality of life can be considered as a global personal assessment of a single dimension, which may be causally responsive to a variety of other distinct dimensions including dimensions such as health” (Beckie & Hayduk, 2004, p. 281).

The conceptual models by Smith et al. (1999) and Beckie and Hayduk (2004) are based on the premise that health and quality of life constitute distinct concepts, and that quality of life can be viewed as a unidimensional concept that is to some degree influenced by health. This distinction between health and quality of life was emphatically argued by Michalos (2004) who stated that "there are good reasons for carefully distinguishing ideas of health and quality of life, and for not interpreting SF-36 and SIP [Sickness Illness Profile (Bergner, Bobbitt, Carter, & Gilson, 1981)] scores as measures of the quality of life" (p. 28). This was further substantiated by findings in studies by Michalos and colleagues (Michalos, Hubley, Zumbo, & Hemingway, 2001; Michalos, Thommasen, Read, Anderson, & Zumbo, 2005; Michalos, Zumbo, & Hubley, 2000) who explicitly sought to examine the degree to which health status, and other variables, contributed to quality of life. They used single-item indicators, such as "How satisfied are you with your overall quality of life?" (Michalos et al., 2001, p. 247), for the measurement of quality of life. The results indicated that quality of life was significantly correlated with five of the eight health outcomes measured by the SF-36 in a sample of 687 adults living in a rural district of British Columbia, Canada. Statistically significant correlations with quality of life ranged from 0.11 (95% CI = 0.04 - 0.19) for physical function to 0.36 (95% CI = 0.30 - 0.43) for role performance limited by emotional problems (Michalos et al., 2005). In another study, regression analyses revealed that health status explained between 34% and 37% of the variance in the quality of life of older people living in British Columbia (Michalos et al., 2001). Mental health status explained most of the variance followed by general health status and physical functioning.

1.3 A conceptualization of quality of life

The above conceptual models and empirical findings pertaining to the relationship between health status and quality of life are based on a common sense understanding that there are many aspects of life that may influence a person's quality of life. In accordance with these models, we use the term *global QOL* to mean a person's appraisal of life as a whole, as distinguished from the appraisal of particular life domains. We drew from the influential work by Campbell, Converse, and Rodgers (1976) to define life domains as "the areas of experience which have significance for all or most people and which may be assumed to contribute in some degree to the general quality of life experience" (p. 12). They suggested that life domains, such as the physical, psychological, and social domains of life, can be evaluated in terms of the degree of satisfaction with various conditions in life (they used the word attributes to refer to these conditions) that have the potential to contribute to global QOL. Accordingly, the so-called life domains are often evaluated based on responses to questions about a person's satisfaction with particular conditions in his or her life.

In addition to questions about a person's *satisfaction* with conditions in life, many instruments for the measurement of quality of life also include questions about the *perceived status* of those conditions (Ferrans, 2005). Although it is reasonable to suggest that individuals' satisfaction with conditions in life is influenced by their perceived status of those conditions, their satisfaction with life domains may also be influenced by other factors such as various "personal characteristics" and different "standards of comparison" (Campbell et al., 1976, p. 16). Different standards of comparison may, for example, be associated with differences in the personal values, expectations, aspirations and needs of individuals (Ferrans, 2005). Thus, the degree of satisfaction cannot be synonymous to the perceived status of conditions in life. These two kinds of measures must therefore be represented as distinct variables in conceptual models of the relationships among global QOL, satisfaction with life domains, and health status.

We now turn to the relationships between the various life domains and global QOL. The conceptual models of quality of life that we discussed earlier are based on the premise that global QOL arises from being satisfied with conditions in life that pertain to various life domains. However, Campbell et al. (1976) argued that we need to be careful not to assume that global QOL is synonymous with the sum of degrees of satisfaction with life domains. This is clearly expressed in the following quotation: “It is not unlikely that people evaluate their lives in general terms and it seems very possible that this overall evaluation may not be a simple sum of the domain evaluations” (p. 13). Thus, their proposition holds that, although global QOL can be viewed as something that arises from being satisfied with various life domains, it is unlikely that individuals’ global QOL would be fully captured by their satisfaction with all the life domains that are generally considered to be important.

The notion of conditions in life that may contribute to global QOL is represented in several other theories about quality of life. For example, in his theoretical scheme of “the four qualities of life,” Veenhoven (2000, p. 5) used the terms “liveability of the environment” (p. 6) to refer to conditions for living well that are external to the person (e.g., social and environmental resources) and “life-ability of the person” (p. 6) to refer to conditions that are internal to the person (e.g., physical and mental health status). He argued that both external and internal conditions contribute to what he referred to as “the subjective appreciation of life,” which he defined as “the quality [of life] in the eye of the beholder” (p. 7). Nordenfelt (1993) similarly distinguished between external and inner welfare to refer to contributing conditions for quality of life. In his philosophical analysis of the relationships among quality of life, health, and happiness, “external welfare” referred to those “phenomena which surround us and continuously affect us" (p. 35), and "inner welfare" referred to "that combination of inner properties which lead to or positively affect our wellbeing" (p. 37).

Based on these theories of quality of life and the previously discussed conceptual models, we deduce that global QOL is best conceptualized as being distinct from the multitude of contributing conditions that may affect it. Global QOL can be defined as a unidimensional concept that refers to a person's appreciation of his or her life as a whole. Here, we have adopted Veenhoven's (2000) use of the word *appreciation* to refer to the degree to which a person is satisfied or happy with life as a whole. This definition is broadly applicable because it avoids specifying the particular conditions of life that a person might take into account when appraising his or her global QOL. Accordingly, questions such as "How satisfied are you with your quality of life?" or "How satisfied are you with your life in general?" which have been widely used in health and social indicators research, are appropriate for the measurement of global QOL (Bowling, 2005; Campbell et al., 1976; Cantril, 1966; Michalos et al., 2001; Michalos & Zumbo, 2002). Most important, however, the definition does not provide a specific description of quality of life nor does it presume a particular frame of reference (Ferrans, 2005).

1.4 A conceptualization of health

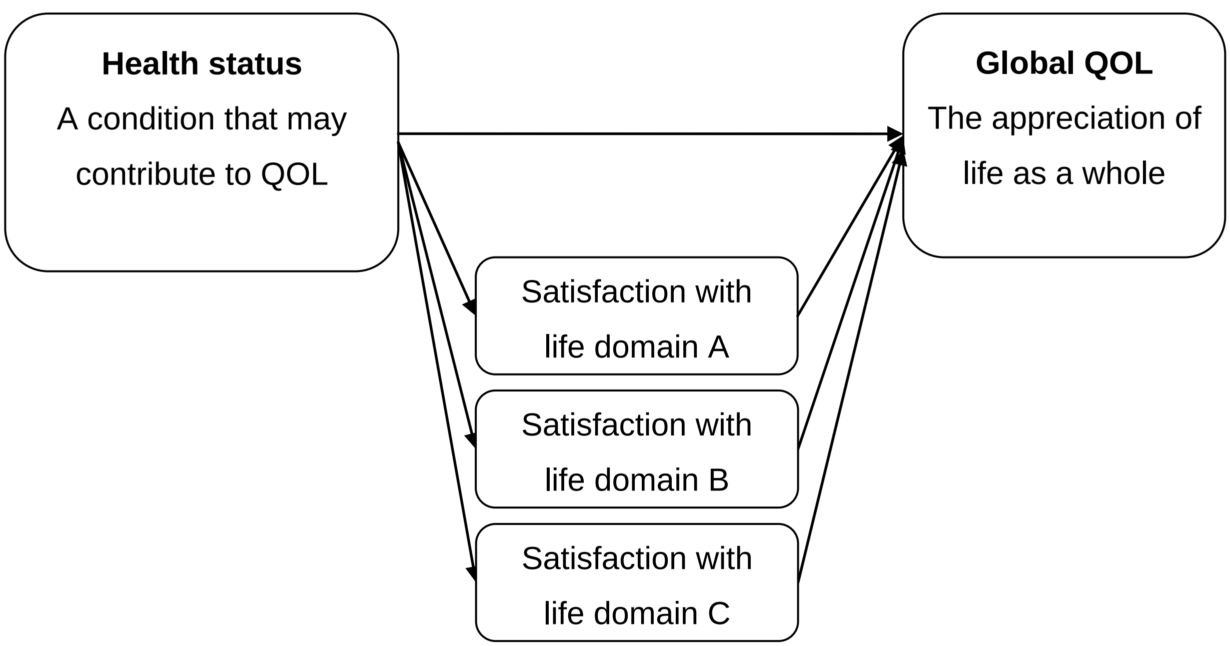
We have mostly discussed conceptual and theoretical considerations pertaining to quality of life. We now turn to a discussion of what we mean by health. In accordance with the conceptualization of quality of life that we put forth, we suggest that health can be viewed as pertaining to certain conditions that may contribute to a person's quality of life. Although further specification of this definition might seem desirable, it is widely acknowledged that the specifics of what constitutes health vary considerably across history and as a result of different cultural perspectives (Bircher, 2005). Any attempt at a universally applicable definition of health could therefore only be stated in non-specific terms. However, our primary concern here is to distinguish health from quality of life. We therefore use the term *health* to refer to various health-related conditions including: (a) the presence or absence of disease, (b) the presence of symptoms of disease or side effects of related treatments, (c) the degree of physical or mental functioning, and (d) perceptions of physical or mental health status. The term *health outcomes* often has been used to refer to this broad range of health-related conditions (Bowling, 2005).

The presence of disease essentially refers to conditions that are defined by alterations of biological functioning. Standardized diagnostic criteria that are based on observable signs and laboratory tests are used to determine the presence of disease. Symptoms can be seen as "a patient's perception of an abnormal physical, emotional, or cognitive state" (Wilson & Cleary, 1995, p. 61). Physical and mental functioning refers to the ability of a person to perform particular activities that require physical action and cognitive processing (Bowling, 2005). Such activities may include daily tasks (i.e., activities of daily living) or social activities to do with employment or other responsibilities as, for example, operationalized by the SF-36 (Ware et al., 1993). And, perceived physical and mental health status can be defined as a person's global perception of his or her physical and mental health. Thus, health status does not reflect whether individuals are satisfied with their physical and mental health, but rather the degree to which individuals perceive themselves to be physically and mentally healthy.

1.5 A conceptual model of health and quality of life

Based on the theories and conceptual models that we have discussed, we conclude that the relationships among global QOL, satisfaction with various life domains, and perceived health status can be summarized as follows: (a) global QOL can be conceptualized as a unidimensional concept that is partially explained by satisfaction with various life domains, (b) health status can be viewed as a condition that has the potential to contribute to global QOL and satisfaction with various life domains, and (c) the relationship between health status and global QOL is mediated by satisfaction with various life domains. These theoretical propositions provide the basis for our conceptual diagram of the relationships among global QOL, satisfaction with dimensions of life, and health status, which are examined in our study (see Figure 2).

Figure 2 A conceptual model of the relationships among global QOL, satisfaction with life domains, and health status



1.6 Issues pertaining to the measurement of quality of life

The theoretical propositions about the relationships among global QOL, health status, and satisfaction with the various life domains have far-reaching implications for the measurement of quality of life. Empirical models of the relationship between health status and quality of life rely on several fundamental ideas about measurement reliability and validity. Reliability can be defined as the degree to which an observed score corresponds to the true score of a particular variable. Or, as stated in a widely-used book on nursing research methods by Polit and Beck (2004), "Reliability is the proportion of true variability to the total obtained variability" (p. 421). Validity can be defined as the degree to which a measure is representative of a particular concept. Thus, whereas reliability can be viewed as referring to the technical accuracy and precision of measurement, validity refers to the relationship between a measure and the intended concept of interest (Viswanathan, 2005). The process of validation can therefore be viewed as referring to the theories and procedures by which inferences about the relationship between a measure and its intended concept is substantiated (Zumbo, 2007). Consequentially, the relationships among concepts are closely intertwined with assumptions about how they are measured (Blalock, 1974). The distinction between concepts and their corresponding measures is of particular importance in health and quality of life research because indicators used for the measurement of health status and satisfaction with various life domains are frequently combined in quality of life instruments.

Before discussing the relationship between concepts and measures it is important to briefly define these terms. Drawing from Chinn and Kramer's (2004) work on knowledge development in nursing, the term *concept* may be defined as a complex mental formulation of perceived phenomena. They distinguished between empirical concepts, which are directly or indirectly observable, and abstract concepts, which involve a larger degree of mental construction. The latter are inferred from related observations. Within a theoretical context, abstract concepts are often referred to as constructs. In other words, constructs can be defined as theoretical representations of phenomena that are not directly observable (Edwards & Bagozzi, 2000). The term *measure* is meant to refer to a scaled score associated with that construct (Edwards & Bagozzi). Measures are obtained by using instruments that, in the context of the non-basic sciences, commonly exist in the form of surveys or questionnaires. It is of fundamental importance to recognize that a measure is not synonymous with the construct being measured. The following discussion of relevant psychometric considerations builds on these foundations by focusing on the relationship between the measurements of health status and QOL.

1.6.1 Psychometric considerations

The nature of the relationship between a construct (e.g., QOL) and a measure (e.g., the combined score of various QOL indicators) can be analytically described in terms of: (a) the direction of the relationship and (b) the proximity of the construct to the measure (Edwards & Bagozzi, 2000). The direction of the relationship relates to the distinction between causal and effect indicators, and proximity relates to the notion of indirect measurement structures. Indicators for the measurement of QOL can be viewed, in terms of their relationship to the unobserved QOL construct, as response variables (also referred to as effect indicators) or as explanatory variables (also referred to as causal indicators) (Bollen & Lennox, 1991; Fayers & Machin, 2007). Edwards and Bagozzi similarly made the distinction between reflective measurement models, where the construct explains the covariances among effect indicators, and formative measurement models, where causal indicators are used to predict the construct.[2] The essential difference between these models relates to the causal mechanisms underlying the covariances among the indicators. Reflective models are based on the assumption of invariant covariances among the effect indicators. This assumption is based on the theoretical proposition that the reflective indicators are exchangeable in the sense that they are representative of a hypothetical population of all possible indicators (Zumbo, 2007). On the other hand, the causal indicators in formative models are not necessarily exchangeable and their covariances are therefore not expected to be fully accounted for by a single common latent variable. In other words, though the residual variance of the latent factor in a reflective model is based on the degree to which the covariances among the effect indicators remain unaccounted for, the residual variance in the formative model is based on the degree to which the causal indicators do not account for the variance in the latent factor as measured or brought about by other sources (this is analogous to residual variance in conventional regression models).

As pointed out by Fayers and others (Fayers, 2004; Fayers & Hand, 2002; Fayers & Machin, 2007), the distinction between causal and effect indicators is of central concern to

---

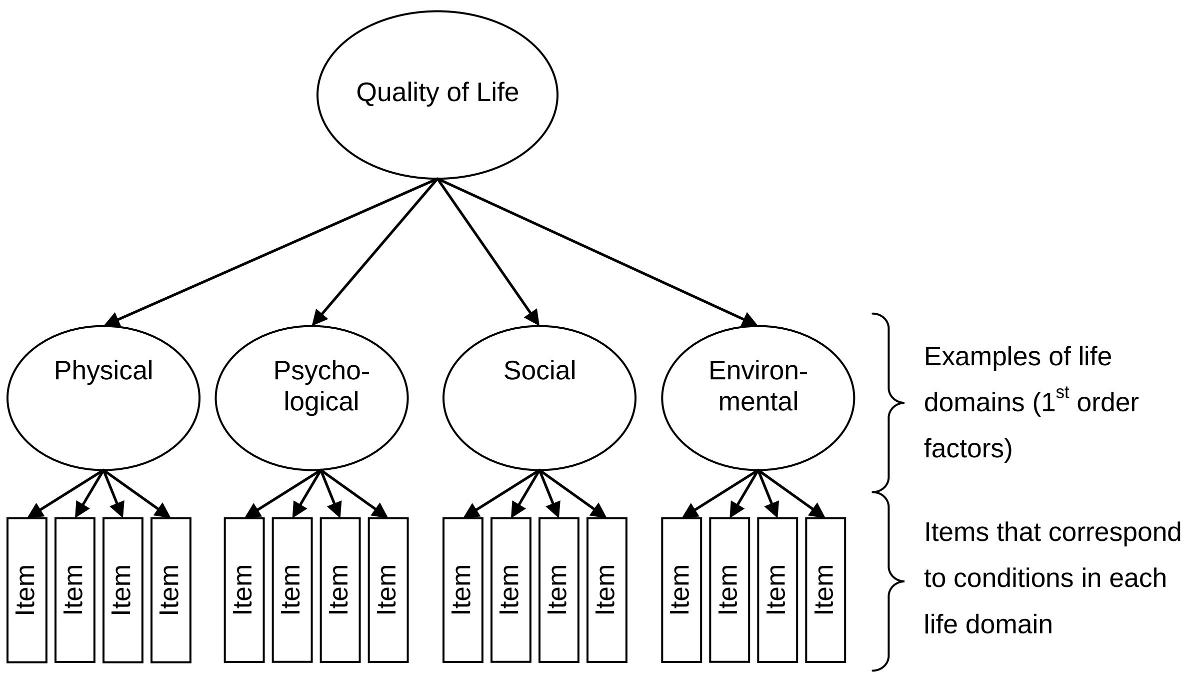
[2] Edwards and Bagozzi (2000) discussed the distinction between reflective measurement models, where a latent factor explains the covariances among effect indicators, and formative measurement models, where causal indicators are used to produce an endogenous (latent) variable. The effect indictors ($x_{i \ldots n}$) in the reflective model are defined as a function of the factor loadings for each of the items (λ_i) times the variance of the overarching latent variable (ξ_i) plus measurement error for the corresponding item ($\delta_{i \ldots In}$) as shown in the following equation: $x_i = \lambda_i \xi_i + \delta_i$. On the other hand, in the formative model the causal indicators (x_i) function as manifest explanatory variables as, for example, shown in the following equation: $\eta = \sum_i \gamma_{i \ldots n} \chi_{i \ldots n} + \zeta$, where ζ signifies a disturbance term for the residual variance of the endogenous variable, η. Thus, effect indicators essentially refer to observations (e.g., items in an instrument) that are specified as response variables that co-vary in relation to the construct being measured (as represented by a common latent factor), whereas causal indicators refer to observations that are specified as independent or explanatory variables that explain the variance in the construct of interest (e.g., Blalock, 1974; Bollen & Lennox, 1991; Fayers & Hand, 2002).

the measurement of quality of life in health research. This distinction relates closely to the methods by which an instrument is developed and evaluated and the purposes for which an instrument is used. Drawing from Feinstein's (1987) discussion of clinimetric indexes, Fayers (2004) argued that many quality of life instruments have been developed with specific clinical purposes in mind (see also Fayers & Hand, 2002; Fayers & Machin, 2007). These instruments do not necessarily measure an overarching theoretical construct, as in a reflective model, but rather provide an index of phenomena that are deemed relevant in a particular clinical context. Fayers and Machin (2007) raised the concern that clinically derived instruments might not conform to the assumptions underlying reflective measurement structures because they may include causal indicators that contribute to a common construct such as quality of life (i.e., in the form of a formative model). In other words, though effect indicators in psychometric models are carefully selected according to the degree to which they consistently reflect a common construct (with some error tolerated), causal indicators are selected based on the degree to which the indicators are expected to explain the construct of interest (e.g., the importance of the indicators in explaining quality of life). Similar arguments can be applied to the relationships between satisfaction with various life domains and global QOL. As pointed out earlier, satisfaction with life domains can be viewed as pertaining to those conditions in life that have the potential to contribute to global QOL. From this perspective, satisfaction with life domains would be viewed as formative indicators that explain global QOL, rather than reflective indicators that measure it.

In addition to distinguishing between formative and reflective measurement models, we need to distinguish direct and indirect relationships between a construct and the perceived phenomena by which its presence or magnitude is inferred (i.e., the items by which a measurement is obtained). Edwards and Bagozzi (2000) explained that the relationship between items and the corresponding construct may be mediated or moderated by another latent variable. To distinguish between direct and indirect measurement, they suggested that "as a general rule, if a measure describes the inherent attributes of a construct, the relationship between the construct and the measure should be considered direct, whereas if the measure refers to a cause or an effect of the construct, their relationship is indirect, spurious, or unanalyzed" (pp. 167-168). In our conceptual model of the relationships between health status, satisfaction with life domains and global QOL, the relationships between the observed indicators of the life domains and global QOL are spurious in the sense that these relationships are partially confounded by the latent factors that represent each of the life domains (see Figure 2 on page 14).

The above considerations have far-reaching implications for the measurement of health status, satisfaction with life domains, and global quality of life. Many multidimensional instruments designed to measure quality of life are based on the assumption that scores pertaining to various conditions in life (as, for example, represented by responses to particular items in a questionnaire) can be combined so as to obtain summary measures of various life domains, which can subsequently be combined to obtain an overall (general) quality of life score. The corresponding measurement model, which is illustrated in Figure 3, has been characterized as an "indirect reflective model" (Edwards & Bagozzi, 2000, p. 162). This model is based on the psychometric theory that the first-order latent factors (e.g., life domains) consistently reflect a second-order factor (e.g., quality of life), and that the correlations among these first-order latent factors are fully accounted for by the second-order factor (i.e., the residual variances of the first-order factors are uncorrelated with each other and with the second-order factor) (Edwards & Bagozzi, 2000; Fayers & Machin, 2007).[3] In other words, this model is based on the assumption that measures pertaining to each of the life domains co-vary in a consistent manner with respect to a common cause. Another way of saying this is that a change in any of the life domains, on

Figure 3 Indirect reflective model for the measurement of quality of life



Notes: The term *items* is meant to refer to the questions in multidimensional instruments designed to measure quality of life.

average, coincides with corresponding changes in all other life domains. It should be noted that this model is fundamentally different from the previously described spurious indicator model and the corresponding theoretical proposition that the life domains constitute

[3] The assumptions mentioned here focus on the relationships between the first-order and second-order factors as implied by the indirect reflective measurement model. There are additional assumptions underlying the indirect reflective measurement model that are explained in detail in the chapter on methodology and methods.

conditions that contribute to quality of life (i.e., the life domains in the spurious model are formative indicators of global quality of life) (see Figure 2).

Concerns with indirect reflective measurement models have been extensively discussed in the disciplinary contexts of psychology and sociology (e.g., Blalock, 1974; Bollen & Lennox, 1991; Edwards, 2001; Edwards & Bagozzi, 2000), and a few researchers have raised similar concerns about the measurement of quality of life and health outcomes (Beckie & Hayduk, 1997, 2004; Donaldson, 2005; Fayers & Hand, 2002; Fayers & Machin, 2007; Feinstein, 1987). Fayers and Hand (Fayers, 2004; Fayers & Hand, 2002) questioned the almost exclusive reliance on factor analysis for the measurement of diverse health outcomes and emphasized the need to carefully distinguish between effect and causal indicators. Similarly, Donaldson (2005) recommended that researchers use an indirect formative measurement model, rather than exclusively relying on an indirect reflective measurement model to examine the relationships among health outcomes (Ware et al., 1993). The general recommendation is that measurement structures should be more rigorously tested before they are used for the measurement of health and quality of life.

Our concerns arising from the predominant reliance on indirect reflective models for the measurement of quality of life include: (a) the assumption that variables pertaining to various life domains co-vary in a consistent manner in relation to a common overarching construct is incongruent with the notion that the various life domains refer to conditions that contribute to quality of life and (b) combining health outcomes measures with measures of global QOL, wellbeing, and life satisfaction, within the same measurement model, results in a conflation of concepts that are fundamentally different in nature. In other words, we would not expect the life domains in multidimensional quality of life instruments to co-vary in a consistent manner with respect to quality of life. Rather, we would expect these variables to contribute to the explanation of variance in quality of life.

Considering the above concerns, we suggest that an indirect reflective measurement model as described above is not a theoretically defensible model for the measurement of quality of life. Despite the wide-spread use of this model, the theoretical premises of combining measures of different life domains in a common factor structure are often not clearly articulated and examined. In particular, it is often unclear how the life domains, which purportedly comprise dimensions of quality of life, are expected to relate to one another (Ferrans, 2005). Nevertheless, the indirect reflective model is based on strong theoretical assumptions about the relationships among these variables. Our concern is therefore that some of the assumptions underlying the multidimensional measurement of quality of life in the form of an indirect reflective model may not be warranted, and that these approaches to the measurement of quality of life could therefore lead to erroneous conclusions about the relationships among quality of life, satisfaction with particular life domains, and health status.

1.6.2 Additional considerations pertaining to measurement validity

Thus far we have explained the indirect reflective measurement model in terms of the implied relationships among the reflective indicators. Support for these types of relationships is generally viewed as a necessary (but not sufficient) condition for construct validity, which can be partly evaluated in terms of the degree to which the model fits in a particular sample (e.g., using confirmatory factor analysis). A well-fitting model would provide support for the theoretical proposition that the reflective indicators are indeed exchangeable in that sample (while allowing for a certain degree of error). However, another important aspect of validation pertains to the degree to which the measurement model is equivalent in different samples drawn from the target population. Zumbo (2007) referred to this as “the exchangeability of sampled and unsampled units (i.e., respondents) in the target population” (p. 59). This aspect of validation refers to the degree to which individuals interpret and respond to the items, which correspond to each of the life domains, in a consistent and comparable manner. It is a necessary condition for the generalizability of inferences pertaining to the measurement structure of a particular instrument (Zumbo). If this assumption holds for the respondents in a particular sample, then that sample can be said to be homogeneous with respect to the measurement model. That is, if respondents interpret and respond to items in a consistent manner, then the relationships among the corresponding indicators will be distributed in the sample in a uniform manner.

This notion of homogeneity with respect to a particular measurement structure requires more attention in research pertaining to the measurement of quality of life. Considering the diversity of items that comprise most multidimensional quality of life instruments, it is plausible that respondents would diversely interpret some items because of, for example, differences in age or culture, or because of contextual differences such as different living environments, experiences resulting from disease or illness, or other life circumstances. Although some of these differences may be observed, there may also be unobserved differences that result in various interpretations and inconsistent responses to some items. Clearly, the validity of inferences drawn from quality of life instruments is contingent on the assumption that the respondents interpret the items in a consistent and comparable manner.

Therefore, in addition to the raised concerns about the indirect model for the measurement of quality of life, researchers ought to be concerned that people may not interpret and respond to some of the items of so-called multidimensional quality of life instruments in a consistent fashion. Quality of life instruments are often developed for the purpose of making generalizable inferences pertaining to the impact of disease or illness, medical treatments and other health-care interventions, and health policy in general populations consisting of diverse individuals. It is therefore imperative to examine the degree to which such generalizability is warranted and to develop instruments that consist of items that are interpreted in the most consistent manner by people in the target population.

1.7 Purpose and analytic objectives

To further explore these concerns, we designed a study to test some assumptions underlying the use of indirect reflective models for the measurement of quality of life (as exemplified in Figure 3), and to propose an alternative model to examine the relationships among satisfaction with various domains of life, global QOL, and health status (see Figure 2). A relatively large and diverse population was required so that we could compare model parameters for the measurement of quality of life across different sampling units (or subgroups) of respondents who might not respond to some indicators of quality of life in a consistent manner. The problem of inconsistent responses to survey questions could potentially apply to all populations, and has been identified as a particular concern for research involving children and adolescents due to differences in cognitive development and language abilities (Barnette, 2000; Borgers, Hox, & Sikkel, 2004; Fletcher & Hattie, 2005; Marsh, 1986; 1996; Wallander, 2001). We therefore decided to specifically examine the measurement of quality of life and its relationship to health status in a geographically, developmentally, and culturally diverse population of adolescents.

1.7.1 Quality of life in adolescents

Quality of life is increasingly viewed as an important consideration in research on adolescents' health (Dannerbeck, Casas, Sadurni, & Coenders, 2004; Huebner, Nagle, & Suldo, 2003; Huebner et al., 2004; Kaplan, 1998; Koot & Wallander, 2001; Raphael, 1996; Raphael, Brown, Rukholm, & Hill-Bailey, 1996; Topolski, Edwards, & Patrick, 2004; Topolski et al., 2001; Wallander, Schmitt, & Koot, 2001). Accordingly, several multidimensional instruments for the measurement of quality of life have been developed for the purpose of examining the impact of disease and chronic illness on various life domains that are considered to be of importance to adolescents (Drotar, 1998; Edwards, Patrick, & Topolski, 2003; Hinds & Haase, 2003; Spieth & Harris, 1996). Quality of life instruments have also been used in population health surveys of adolescents to examine the impact of health policies and health promotion initiatives (Bradford, Rutherford, & John, 2002; Huebner et al., 2004; Kaplan, 1998; Raphael, 1996; Raphael et al., 1996; Topolski et al., 2004; Topolski et al., 2001). These instruments typically consist of subscales that represent life domains that are of particular importance to children and adolescents, including: (a) perceptions of self (e.g., self-esteem), (b) relationships with friends and family, (c) school experiences, and (d) the living environment. In addition, some instruments include questions that specifically assess physical and mental health status, or physical, emotional, and social functioning.

Some researchers have developed conceptual models describing the relationships between these various life domains in children and adolescents. For example, based on their qualitative study of adolescents' quality of life, Edwards, Huebner, Connell, and Patrick (2002) developed a conceptual model with the domains of "social relationships," "sense of self," and "environment" (p. 283) as important contributing conditions of adolescents' global QOL. This model provided the basis for the development of the Youth Quality of Life Instrument – Research Version (YQOL-R) (Patrick, Edwards, & Topolski, 2002), which was used to obtain summary scores for each of the life domains. Principal components analysis was used to provide support for the use of an overall score (Patrick et al.). Huebner (1997) posited a model that consists of similar life domains of importance to children and adolescents (satisfaction with family, peers, school, self, and living environment). However, in contrast to viewing these life domains as contributing conditions for quality of life, his "Multidimensional Life Satisfaction Model" is based on the theoretical premise that the life domains constitute dimensions that arise from general life satisfaction. He accordingly developed the Multidimensional Students' Life Satisfaction Scale (MSLSS) (Huebner, 1994) based on an indirect reflective measurement structure wherein general life satisfaction is specified as a second-order factor that accounts for the correlations among the five life domains (Huebner, Laughlin, Ash, & Gilman, 1998).

In consideration of the theoretical and empirical developments pertaining to the quality of life of adolescents, we designed a test of the indirect reflective measurement structure of the MSLSS (Huebner, 1994) and of an alternative spurious model of the relationships among adolescents' satisfaction with various domains of life, global QOL, and health status (see Figure 2 on page 14). We selected this instrument because the developers used factor analysis techniques for item selection and to validate its factor structure (construct validity), and because an indirect reflective measurement structure for this instrument was purportedly substantiated by at least three confirmatory factor analyses (Gilman, 1999; Gilman, Huebner, & Laughlin, 2000; Huebner & Gilman, 2002; Huebner et al., 1998). In addition, the MSLSS is an example of an instrument that measures satisfaction with life domains in a manner that is not contaminated by items measuring physical and mental health status. Thus, this instrument provided a suitable example for our study purposes.

1.7.2 Analytic objectives

In accordance with the general purposes of our study and our particular focus on the measurement structure of the MSLSS in adolescents, we formulated the following analytic objectives to guide our analyses: (a) to test the assumptions of the putative indirect reflective measurement structure of the MSLSS with the goal of assessing its reliability and validity with respect to the measurement of adolescents' satisfaction with their family, friends, living environment, school and self, and their general life satisfaction, (b) to determine the degree to which the dimensions of life satisfaction explain global QOL, and (c) to examine whether perceived mental health status, perceived physical health status, or both contribute to global QOL, and whether the dimensions of life satisfaction mediate the relationship(s). However, despite our focus on adolescents, we emphasize that our study was not intended to address theoretical considerations in adolescent developmental psychology, but rather was designed to examine problems pertaining to the measurement of quality of life and its relationship with health status, and to propose theoretically defensible solutions.

1.8 History and use of the MSLSS

Considering our explicit focus on the measurement structure of the MSLSS, it is informative to provide some background on the history and use of this instrument. The

MSLSS was developed by Huebner (1994) as an instrument for the measurement of general life satisfaction and satisfaction with particular life domains considered to be of importance in the lives of children and adolescents (Huebner, 2001). The measured satisfaction with life domains were based on theoretical developments and empirical studies related to life satisfaction in adults and children, interviews with elementary school children, student essays, and exploratory and factor analyses. The instrument was subsequently examined, and purportedly validated, in adolescents (Huebner, 2001; Brantley, Huebner, & Nagle, 2002; Gilman, 1999; Gilman et al., 2000; Huebner & Gilman, 2002; Park et al., 2004). As a result, the instrument has been used for a variety of research purposes in samples of children and adolescents with ages ranging from 8 to 19 years (grades 3 to 12).

To locate studies that used the MSLSS,[4] we searched for the terms "MSLSS," "Multi(-)dimensional Students' Life(-)satisfaction," "SLSS," and "students' life satisfaction scale" in the titles, keywords, and abstracts of several databases (including CINAHL, PubMed, Embase, PsychINFO, Social Sciences Citation Index, Science Citation Index, Dissertation Abstracts International, and Health and Psychosocial Instruments). This search strategy produced results that included the scale of interest, the MSLSS, as well as its precursor, the SLSS (Huebner, 1991), and a more recent abbreviated 6-item version, the BMSLSS (Seligson, Huebner, & Valois, 2003). We also reviewed the reference lists of several review articles about the MSLSS and we completed forward citation searches in Social Science Citation Index and PsychINFO for three key articles about the development of the MSLSS (Gilman et al., 2000; Huebner, 1994; Huebner & Gilman, 2002). This strategy identified 184 citations, published between 1982 and 2007, that were subsequently screened

[4] Because the MSLSS is cited frequently, we do not refer repeatedly to the corresponding original reference by Huebner (1994).

to determine whether findings related to the MSLSS were indeed reported. We identified 39 independent MSLSS studies that were published between 1994 and 2007 of which 4 were not published in English (we relied on the available citation information and the abstracts, which were published in English, for our review of the studies that were not published in English).

Based on a review of the identified MSLSS studies we observed that this instrument has mostly been used to examine the relationships between life satisfaction and a variety of variables of relevance to developmental psychology and positive psychological wellbeing in children and adolescents. This is congruent with one of the goals of the development of the MSLSS, which was to focus on the experiences of children and adolescents as an alternative to the predominant emphasis on psychopathological symptoms (Huebner & Gilman, 2002). For example, Greenspoon and Saklofske (2001) designed a study to substantiate an integrated system for the assessment of mental health in children by examining a variety of variables (e.g., "domains of temperament, personality, self-concept, interpersonal relations, and locus of control" (p. 84)) that were associated with subjective wellbeing and psychopathology in children. The MSLSS was used to classify 407 children in grades 3 to 6 as having high or low subjective wellbeing. Their study provides an example of the distinction between subjective wellbeing and psychopathology and the importance of including both of these variables in a framework for mental health assessment.

Other researchers have used the MSLSS to examine the degree to which life satisfaction is associated with different social behaviors of children and adolescents in their relationships with parents and peers. For example, Nickerson and Nagle (2004) found that attachment with parents and peers was associated with differences in overall life satisfaction ($R^2 = 0.47$, $p < 0.001$) and most of the satisfaction with life domains, as measured by the

MSLSS, in children and early-adolescents (*N* = 303). In particular, life satisfaction was positively correlated with trusting relationships with peers and parents, and negatively correlated with delinquency as measured by the People in My Life instrument by Cook (as cited in Nickerson and Nagle, 2004). They also found that, in comparison with the younger children in the sample, the early-adolescents reported being less satisfied with school and family.

Gilman (2001) examined the degree to which life satisfaction was associated with social interest and participation in extra-curricular activities in adolescents (*N* = 321). The term *social interest* was used to refer to the degree to which the adolescents saw themselves as engaging in “prosocial behavior (e.g., helpful, compassionate)” (p. 175) as measured by the Social Interest Scale (Crandall, 1975). The results indicated that overall life satisfaction was, to some degree, explained by social interest and participation in extracurricular activities ($R^2 = 0.06$, $p < 0.001$). Specifically, increased social interest was associated with increased satisfaction with friends and family, and increased participation in extracurricular activities was associated with increased satisfaction with school. Gilman did not find statistically significant differences with respect to the adolescents’ satisfaction with self or their living environment.

The MSLSS also has been used to compare life satisfaction scores across different populations of children and adolescents. Gilman and Ashby (2003) examined the relationship between perfectionism and life satisfaction in adolescents (*N* = 132) and found a significant positive association between perfectionism and general life satisfaction and satisfaction with self. The adolescents who set high personal standards reported significantly greater satisfaction with self in comparison with those who set lower standards. However, it was not clear whether these differences were representative of actual differences in life satisfaction or inconsistencies in how adolescents in the different groups interpreted and responded to the life satisfaction items in the MSLSS. Interestingly, the researchers did not find differences in satisfaction with family, friends, school and living environment.

Gilman, Easterbrooks, and Frey (2004) used the MSLSS to compare life satisfaction in children and adolescents (aged 8 to 18 years) who were deaf or hard of hearing (N = 88) with that of those who had no hearing impairments (N = 71). They found that those who were deaf or hard of hearing reported lower satisfaction with most life domains and that the correlations among the dimensions of life satisfaction were not equivalent to those of the children who had no hearing impairments.

Several researchers have used the MSLSS to compare students with different learning abilities. Huebner, Brantley, Nagle, and Valois (2002) examined differences in life satisfaction among adolescents with mild mental disabilities (N = 80) and a matched sample of so-called "normally achieving" (p. 29) adolescents (N = 80). They also compared adolescents' ratings of their life satisfaction with proxy ratings that were obtained from their parents. The findings of their multi-trait multi-method analysis revealed that the parents' and adolescents' ratings of life satisfaction were congruent for the group of normally achieving adolescents. However, they found differences within the group of adolescents with mild mental disabilities and concluded that "the meaning of the reports of normally achieving students may be different for students with special needs (e.g., mental disability)" (p. 27). In another analysis of the same sample, Brantley, Huebner, and Nagle (2002) obtained lower correlations, and thus lower reliability estimates, for all five life satisfaction domains in the adolescents with mild mental disabilities in comparison with the normally achieving adolescents. These findings were replicated by McCullough and Huebner in their comparison of 80 adolescents with learning disabilities with 80 adolescents without learning disabilities. Taken together, the differences in reliability estimates and satisfaction with life domain correlations across various groups of adolescents suggest that there may be differences in how adolescents interpret and respond to questions about their satisfaction with the five life domains.

Before drawing further conclusions, it is important to consider that most of the MSLSS studies were based on relatively small samples in particular geographic locations. Of the 39 MSLSS studies that we identified, 4 were based on reported sample sizes of 100 or less, 17 on samples sizes between 100 and 300, 8 on sample sizes between 300 and 600, and 9 on sample sizes greater than 600 (the sample size for one study could not be determined because it was not published in English). Most studies were based on children and adolescents in schools in the USA: 8 studies were of students from only one school in a Southeastern state, 14 studies were of students from two or more schools in a Southeastern state, 1 study was of students in a particular residential treatment facility in the USA (Gilman & Handwerk, 2001), and 2 studies were of students from schools in other regions of the USA. We found 3 studies that involved a comparison of students in the USA with students in other countries including South Korea (Park, 2000, 2005; Park, Huebner, Laughlin, Valois, & Gilman, 2004) and Croatia (Gilman, Ashby, Sverko, Florell, & Varjas, 2005). Other studies of children or adolescents outside of the USA included 3 studies of school-aged children and adolescents in Hong Kong (Chan et al., 2003; Chang, McBride-Chang, Stewart, & Au, 2003; Leung, McBride-Chang, & Lai, 2004) (one of these studies used only the school subscale of the MSLSS), 2 studies of children in primary or secondary schools in Western Canada

(Greenspoon & Saklofske, 1997, 1998; Greenspoon & Saklofske, 2001), and studies of students in China (Tian & Liu, 2005), the Slovak Republic (Medved'ova, 1999), and Israel (the sampled population for two additional studies could not be determined because they were not published in English). The MSLSS has been translated into Korean (Park, 2000, 2005; Park et al., 2004), Croatian (Gilman et al., 2005), Chinese (Tian & Liu, 2005), Hebrew (Schiff, Nebe, & Gilman, 2006), and Catalan for use in Spain (Casas, as cited in Huebner, 2004). Nevertheless, although the MSLSS has been used in several different samples, it has primarily been used in studies of a relatively small number of students in schools in the Southeastern USA. A study that would allow for the examination of the measurement structure of the MSLSS in a relatively large and culturally, geographically, and developmentally diverse sample could greatly contribute to the validity of using this instrument in adolescent populations.

Our review of the studies that utilized the MSLSS revealed that this instrument has been predominantly used in psychological research. Is has, however, been recommended specifically for health research purposes (Huebner et al., 2003; Huebner et al., 2004). Huebner et al. (2003) argued that instruments for the measurement of satisfaction with life domains, such as the MSLSS, can be used as health outcomes assessments to evaluate the impact and effectiveness of health promotion programs and interventions, for screening purposes so as to identify children and adolescents who may be at risk for developing health problems, and for research purposes to examine differences in various life domains for children and adolescents with chronic illness in comparison with those who have no chronic illness. Accordingly, health-care professionals can use the appraisals of various aspects of the lives of adolescents with physical or mental health challenges to target appropriate supportive services. Nevertheless, we were unable to locate a study wherein the MSLSS was examined in relation to the health status of children or adolescents. Our study was therefore specifically designed to examine the relationships among adolescents' health status and their satisfaction with various life domains with the goal of validating the use of this instrument for health research purposes.

1.9 Concluding comments

In this chapter we discussed several potential concerns related to the measurement of multidimensional quality of life. We specifically questioned the commonly used approaches of combining measures of different life domains into a common factor structure for the purpose of obtaining an overall, or general, measurement of a person's quality of life. We suggested that the specified relationships among life domains, as implicated by the indirect reflective structure underlying these approaches to quality of life measurement, are theoretically implausible. As an alternative approach, we suggested that quality of life can be conceptualized as a global and unidimensional concept that is partially explained by various life domains. We further suggested that health status can be viewed as a condition that has the potential to contribute to global QOL and satisfaction with each of the life domains. Our analytical objectives are to examine these theoretical propositions using a multidimensional life satisfaction instrument, the MSLSS, in a diverse sample of adolescents.

The methods of data collection and statistical analysis are discussed in detail in the next chapter. We particularly emphasize the statistical theory behind confirmatory factor analysis and factor mixture analysis with the purpose of explaining how these methods were used to test the measurement assumptions. The third chapter includes a detailed presentation of the findings. We suggest that the findings have important theoretical implications for the measurement of quality of life and its relationship to health status in adolescents. These theoretical implications are discussed in chapter four. We also provide several recommendations pertaining to the use of the Multidimensional Students' Life Satisfaction Scale (MSLSS) (Huebner, 1994), and we discuss some methodological recommendations pertaining to the use of confirmatory factor analysis and factor mixture analysis for the purpose of testing assumptions pertaining to the measurement of quality of life. Considering the methodological focus of our analyses, we believe the findings to be of interest to health researchers who are concerned about issues pertaining to the measurement of quality of life, and to health researchers who are interested in examining the relationships between quality of life and health status in adolescents.

2 METHODOLOGY AND METHODS

In this chapter we provide a detailed account of the methods used in our study. The chapter begins with an overview of our analyses and a general discussion of the methodological approaches used to address the analytic objectives of our analyses. We then discuss the survey methods and approaches to measurement used to collect the data, and we continue with a discussion of the statistical methods. Several mathematical formulas underlying our statistical methods are discussed in some detail with the purpose of describing the assumptions and theoretical implications of the models that we estimated. Our discussion of the statistical methods concludes with an overview of the techniques that we used to account for missing data.

2.1 General overview of the analytical approach

Our analyses involved three projects pertaining to the measurement of multidimensional life satisfaction and global QOL and their relationships with perceived mental and physical health status. The first project focused on the measurement of multidimensional life satisfaction in adolescents. The measurement structure derived from this project was then adopted for the second project, which was designed to examine the degree to which each of the dimensions of life satisfaction explains global QOL. The model was expanded in the third project so as to examine whether perceived mental or physical health status contributes to global QOL, and whether the dimensions of life satisfaction mediate those relationships.

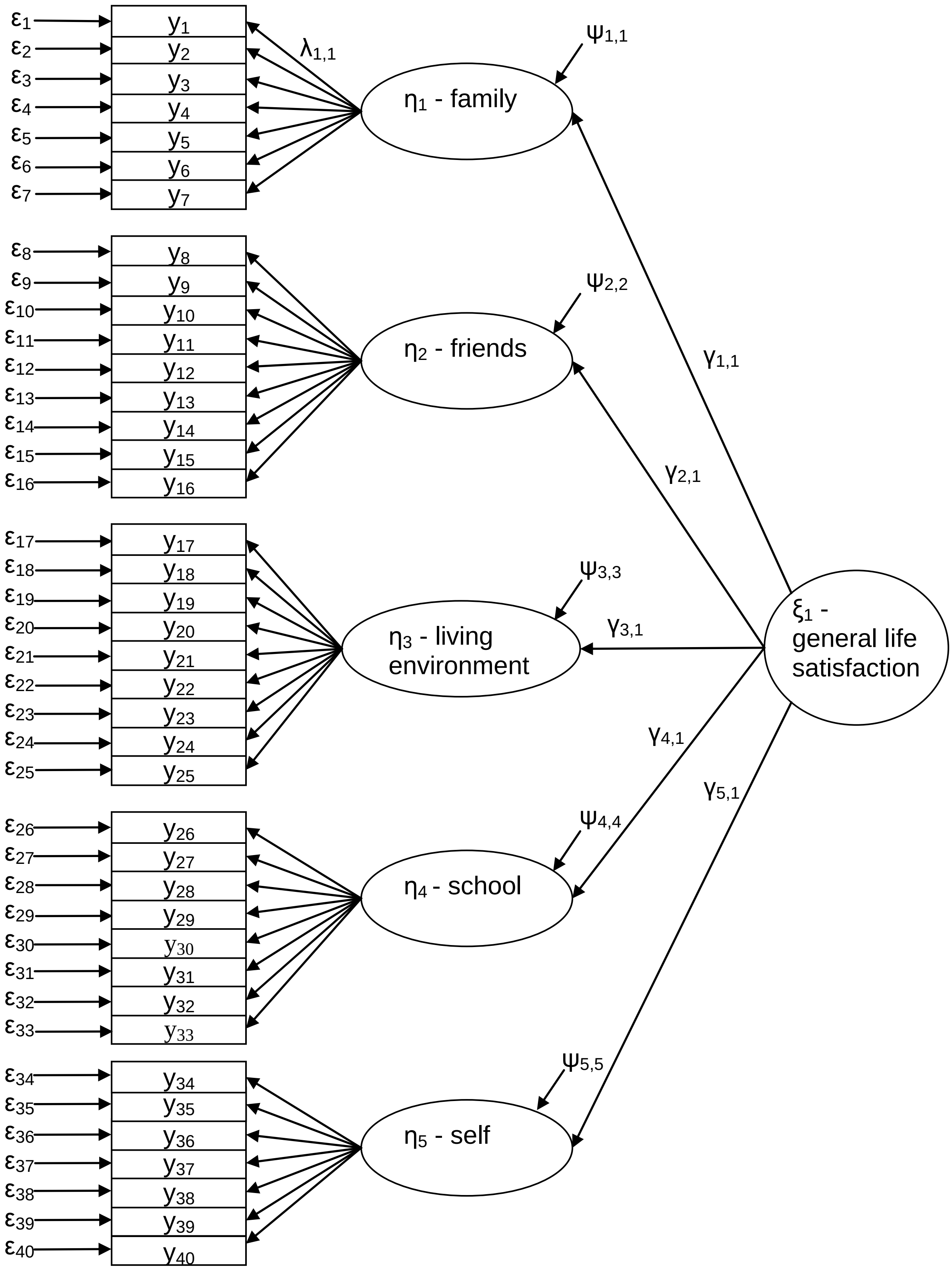
2.1.1 The measurement of multidimensional life satisfaction in adolescents

The purpose of the first project was to test the assumptions of the putative measurement structure of the Multidimensional Students' Life Satisfaction Scale (MSLSS) with the goal of determining whether the instrument is reliable and valid with respect to the measurement of adolescents' satisfaction with their family, friends, living environment, school and self, and their general life satisfaction. Based on the results of several confirmatory factor analyses, it has been suggested that the reliability and construct validity of the MSLSS, for the measurement of satisfaction with family, friends, living environment, school and self, is supported by having confirmed the fit of a measurement structure with five correlated first-order factors (Gilman, 1999; Gilman & Ashby, 2003; Greenspoon & Saklofske, 1998; Huebner et al., 1998; Park, 2000; Park et al., 2004). In addition, the construct validity of the MSLSS for the measurement of general life satisfaction was purportedly supported by fitting a second-order factor that accounted for the correlations among the five first-order factors (Gilman, 1999; Gilman et al., 2000; Huebner & Gilman, 2002; Huebner et al., 1998). The corresponding second-order measurement structure, which is depicted in Figure 4, can be characterized as an "indirect reflective model" (Edwards & Bagozzi, 2000, p. 162). Here, the term *reflective* refers to the presence of one or more latent factors that account for the covariances among the observed variables. And the term *indirect* is meant to refer to a hierarchical factor structure where a second-order factor accounts for the covariances among the first-order factors.

2.1.1.1 Assumptions of the indirect reflective measurement structure

The assumptions underlying the putative indirect reflective measurement structure can be summarized as follows: (a) respondents interpret and respond to the items in a consistent

Figure 4 The indirect reflective measurement structure of the MSLSS



manner (i.e., the thresholds representing the probability distributions for the ordinal response options are invariant across known and unknown classes of respondents within the sample) (Assumption one), (b) the observed variables of each subscale consistently reflect a single latent factor (i.e., they provide a measure of their respective first-order latent factor such that the factor loadings are invariant) (Assumption two), (c) the correlations among the observed variables are accounted for by the first-order factors (Assumption three), (d) the first-order latent factors consistently correlate because of a second-order latent factor (i.e., the second-order factor loadings are invariant) (Assumption four), and (e) the correlations among the first-order latent factors are fully accounted for by a second-order latent factor (i.e., the residual variances of the first-order factors are uncorrelated with each other and with the second-order factor) (Assumption five) (Edwards & Bagozzi, 2000; Fayers & Machin, 2007). Assumptions three and five imply that there are no significant correlations among the observed variables and among the first-order latent factors conditional upon the measurement model. Assumptions two and four imply that samples are homogeneous with respect to the relationships between the latent factors and the items as implied by the measurement structure. Assumption one implies homogeneity with respect to the item responses. Whereas assumptions three and five relate to the question of whether the measurement model fit the data provided by the overall sample, assumptions one, two, and four imply invariance of the model's structural parameters across observed or unobserved differences in the sample (known or unknown classes of respondents). In other words, assumptions one, two and four relate to the question of whether the adolescents that provided the data for these analyses responded to the items in a consistent manner such that their responses were comparable. Invariance of item responses (assumption one) implies that the probability distributions for the ordinal response options (i.e., the thresholds) of each item were representative of all the respondents in the sample. If this were not the case, then the corresponding factor score would not be comparable across the subsamples that were characterized by different item response patterns. Similarly, invariance of the model parameters (assumptions two and four) implies that a single set of factor loadings and factor variance(s) would be representative of the entire sample. If this were not the case, then the factor scores could not be compared across subsamples for which the factor loadings or factor variance(s) differed.

The assumption that the adolescents responded to the items in a consistent manner may not have been warranted. Adolescents may interpret some items differently because of cultural, developmental, or personality differences, or because of contextual differences such as different living environments, employment experiences or other life circumstances. Adolescents also may not share a common frame of reference for responding to various items. Besides the potential for differences in interpretation, adolescents may adopt different response styles with respect to the rating scales of each item. For example, they may respond to some items by upward or downward comparisons of their life circumstances to those of their peers, or to previous life circumstances, or to some ideal or model life circumstance. The resulting differences in “response style behaviors” can greatly distort the measurement of psychological constructs (Moors, 2003, p. 278; Viswanathan, 2005). Although some of these differences may be observed, it is more plausible that there are various unobserved groups of adolescents that differ in how they respond to the questions of the MSLSS because of various combinations of individual and contextual factors.

2.1.1.2 Methodological approaches to the examination of assumptions

Our methodological approaches for testing the assumptions underlying the indirect reflective measurement structure of the MSLSS involved five steps. The first three steps addressed assumptions three and five and the related question of whether the second-order measurement model, as well as the nested first-order measurement models, fit the data provided by the overall sample of adolescents. The remaining steps pertained to assumptions one, two and four and the related question of whether the sample was homogeneous with respect to the parameters implied by the measurement structure.

Did the implications of the structural parameters of the measurement model fit the data provided by the overall sample? The first step was to test the model fit of the indirect reflective model as shown in Figure 4. Good model fit could only have been obtained, in this instance, if assumptions two to five were justified.[5] We tested the second-order factor model as well as the implicit model with five correlated first-order factors. If we were to find that the latter model fit well, but the former model was found not to fit well, we would have had to conclude that assumptions four or five were not justified. We would then have proceeded by examining the residual correlations among the first-order latent factors to consider possible explanations for the second-order model's misfit. If both models were found to not fit well, then we would have concluded that one or more of the other assumptions were not justified.

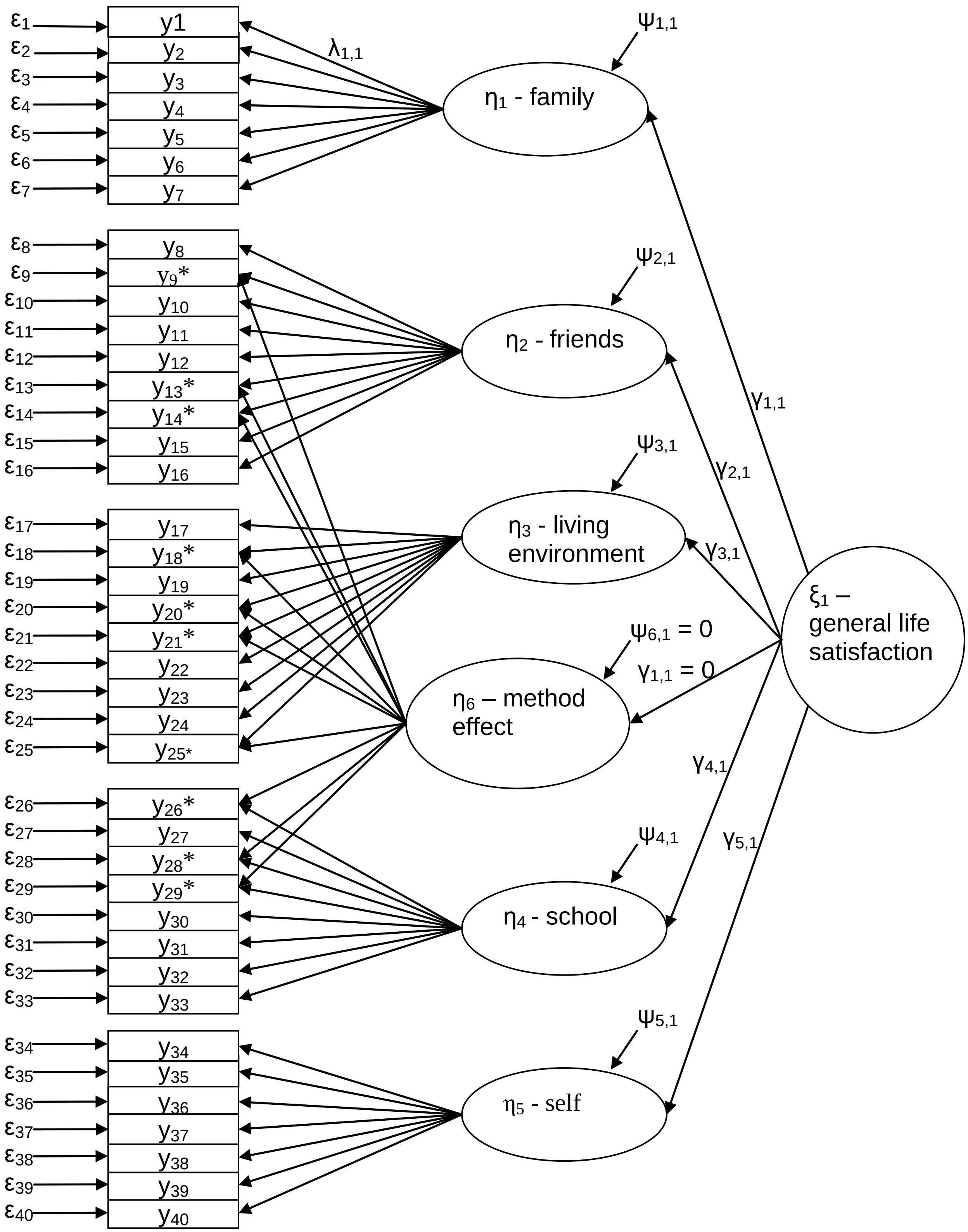
[5] Assumption one is needed to determine whether the adolescents responded to the questions in a reliable (consistent) manner. However, this assumption is not needed to achieve good model fit in a CFA. There may be systematic inconsistencies that inflate correlations among the observed variables, or random sources of error that attenuate these correlations (Viswanathan, 2005). Nevertheless, in both situations, good model fit can still be obtained. Assumption one therefore needs to be tested separately by examining the consistency of item response patterns (e.g., thresholds).

Having observed poor model fit for the original indirect reflective measurement structure (which is described in greater detail below), we continued by specifying a modified measurement model of five correlated first-order factors and one additional independent first-order factor to account for the correlations among the negatively worded items present in the MSLSS (see Figure 5). It is widely recognized that people may respond to negatively worded items differently than they do to positively worded items (e.g., Horan, DiStefano, & Motl, 2003; Tomas & Oliver, 1999). We modelled this additional source of covariance by specifying paths between the negatively worded items and an additional independent first-order factor which we labeled *method effect*. This approach is similar to some of the published analyses of the Rosenberg Self-Esteem Scale (Rosenberg, 1989), which also has some negatively worded items (Horan et al.; Tomas & Oliver, 1999; Wang, Siegal, Falck, & Carlson, 2001).

We proceeded by examining the residual correlations and factor loadings as the basis for finding other potential explanations for poor model fit. By doing so, we were able to identify the model parameters that were associated with those correlations among the observed variables (the MSLSS items) that were poorly accounted for by the first-order factor structure (assumption three).[6] We mapped the residual correlations to identify any observed variables that could be associated with more than one factor, and we used principal components analysis to determine whether the residual correlations could be largely

[6] The measurement model implies a unique factor loading for each observed variable in relation to one of the latent factors. This constraint may be too restrictive if some of the observed variables relate to more than one latent factor. For example, although the question, "I like being in school" is designed to measure the concept, satisfaction with school, it may also reflect some aspect of another concept such as, for example, satisfaction with friends (considering that at least some friends are likely to be at the same school). In this case, the question does not uniquely measure a single concept. The measurement model may then be modified by excluding the question or by allowing the variable to load on more than one factor.

Figure 5 Second-order structure of the MSLSS with a method effect



Notes: The manifest variables with an asterisk are reverse-scored (negatively worded).

accounted for by adding one or two additional dimensions or factors to the model (Zumbo, 2002).

We then continued our analyses by examining the model fit of each of the first-order factor models (i.e., each subscale) independently so as to more specifically test assumption three. This step was necessary to examine the unidimensionality of each of the subscales and to identify any observed variables that might have been poorly associated with their putative corresponding latent factor. For example, an unreliable item might have had an unacceptably small factor loading, or the correlations associated with one particular item might have been poorly accounted for by a single latent factor (i.e., large residual correlations would have been present).

Did the adolescents in this sample respond to the items in a consistent manner? The analytical approach described thus far reflects the assumption that the structural parameters of the measurement model were consistent in the overall sample. However, evidence of lack of model fit may have been associated with inconsistencies in item responses or other model parameters across different subgroups of respondents in the sample. That is, the parameters of the measurement model may not have been invariant with respect to any observed or unobserved differences in the sample. The next steps in our analyses were therefore designed with the purpose of examining whether the assumptions of consistent item responses and factor loadings were warranted. We examined each subscale independently to specifically test assumptions one and two. Factor mixture analysis (FMA) was used to examine sample heterogeneity with respect to the measurement model by allowing the thresholds of the observed variables, as well as the factor loadings, to vary across two or more latent classes (Lubke & Muthén, 2005). By comparing the thresholds and factor loadings across the latent classes, we could identify those observed variables for which the thresholds or factor loadings were least invariant across two or more latent classes.

Based on the results of the FMAs, we subsequently modified the first-order measurement models by selecting the four items with the most consistent response patterns (invariant thresholds) and invariant factor loadings across two or more latent classes (this is further explained below in the section on the statistical methods employed). We sought to isolate at least four items for each subscale so that we could meaningfully assess model fit with a minimum of two degrees of freedom to test a common factor structure (Mulaik, 2004).[7] Thus, we obtained abridged versions for each MSLSS subscale by including those four items that were most reliable in this sample of adolescents.

After determining adequate model fit for each of the abridged subscales, we combined the first-order factor models for each subscale into a comprehensive, albeit reduced, measurement model with correlated first-order factors, as implied by the original indirect reflective measurement model. This step was based on the rationale that a second-order factor model was only justified if the model with correlated latent factors fit well. If this latter model were found to not fit well, we would have concluded that assumption four or assumption five, or both, was not warranted (based on the premise that the measurement models for each first-order factor fit well). That is, if the latent factors were found to not correlate because of a single common factor, then the conclusion would have been that each of the factors did not represent a common overarching construct, and would best be described

[7] Although it has been suggested that three indicators are sufficient to test a single factor model, Mulaik (2004) convincingly argued that at least four indicators are needed to test a common factor structure because the use of only three indicators per factor could not lead to conclusive inferences about the relationships between the factor and other latent variables included in the model.

as distinct concepts. If the correlated factor model were found to fit, we could have further tested the assumption by specifying a second-order latent factor model (an indirect reflective model). A factor mixture analysis of the indirect reflective measurement structure would have provided a direct test of assumption four.[8]

Even if model fit were found to be acceptable for a second-order factor, we would have still needed to consider the magnitude of the indirect relationships between the second-order factor and the observed variables (obtained by multiplying the first-order factor loadings and the second-order factor loadings). In all cases, second-order factors account for a smaller proportion of the variance in the observed variables than any of the first-order factors (by virtue of being distal to the observed variable). The amount of variance in each observed indicator that is accounted for by a second-order factor should be taken into account when determining whether each observed indicator indeed provides a meaningful measure of the second-order factor. For example, if an observed indicator has a standardized first-order factor loading of 0.50, and the regression of the first-order factor on the second-order factor is also 0.50, then the indirect effect that specifies the relationship between the observed variable and the second-order factor is 0.25, and the resulting amount of explained variance that is attributable to the second-order factor is only 6.25%. This would not be desirable if the purpose is to obtain a valid measure of general life satisfaction based on a second-order factor.

In summary, we tested each of the five assumptions underlying the indirect-reflective measurement model of the MSLSS by using CFA to examine the second-order factor model as well as each of the first-order factor models independently, and by using FMA to examine

[8] Although this was attempted, our sample size was unfortunately not large enough to reliably estimate the parameters for this model.

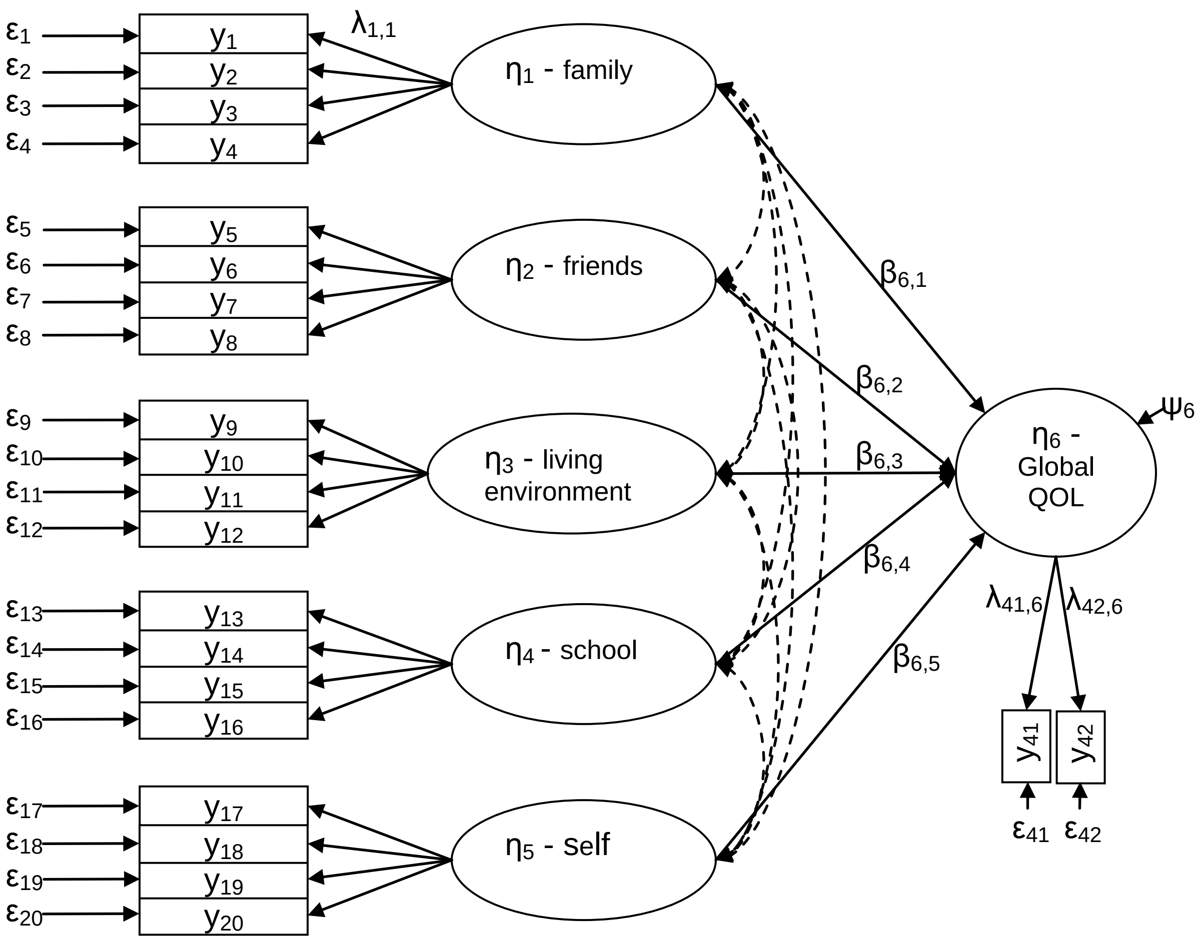
potential sample heterogeneity with respect to the measurement structures for each of the subscales. Through this process we developed abridged versions for each of the MSLSS subscales. The abridged subscales were then specified in a second-order measurement model and evaluated to determine the degree to which the first-order factors reflected a common second-order latent factor.

2.1.2 A spurious indicator model for the relationships between multidimensional life satisfaction and global QOL

The next step in our analyses was to examine the relationships between the life satisfaction dimensions and global QOL as measured by two additional indicators (see Figure 6). This model was specified by regressing the global QOL factor on the five latent life satisfaction dimensions. [9] The specified structural relationships between the MSLSS indicators and the global QOL variables are analogous to the relationships of a "spurious indicator model" as defined by Edwards and Bagozzi (2000, p. 166). This name derives from the fact that the specified relationships between the indicators of the MSLSS and the overarching construct of interest (i.e., global QOL) were spurious with respect to the dimensions of life satisfaction as represented by the first-order factors of the MSLSS. In this sense, global QOL was defined as a composite of various dimensions of life satisfaction that contributed to it. Modeling the variables in this fashion permitted us to evaluate the magnitude of the relationships between the life satisfaction dimensions and global QOL and to determine the relative importance of each of the dimensions in relation to global QOL.

[9] In Figure 6, the model is shown based on the MSLSS measurement structure with five correlated latent factors. Obviously, if we determined that the indirect reflective measurement structure for the MSLSS fit the data well, we then would have specified a second-order factor in this model so as to account for the correlations among the first-order factors.

Figure 6 A spurious indicator model of the relationships between multidimensional life satisfaction and global QOL

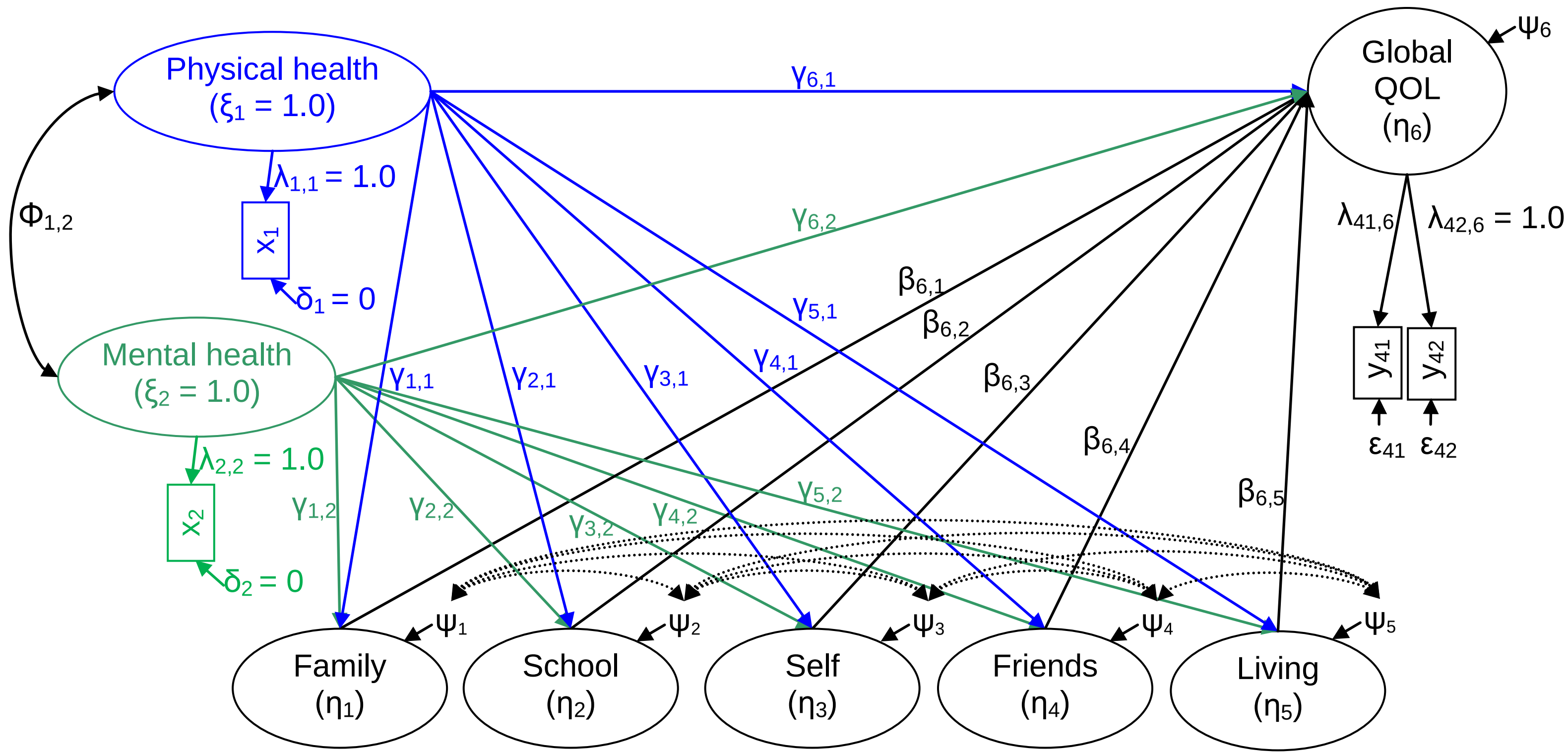


2.1.3 Examining the relationships between health status and QOL

To this point, the analytic approach focused on the measurement of multidimensional life satisfaction and its relationship with global QOL. Our final set of analyses focused on extending these models for the purposes of examining whether perceived mental or physical health status contributed to global QOL and whether the life satisfaction dimensions mediated those relationships (see Figure 7). Thus, in this 'mediation' model, global QOL was regressed on the five life satisfaction dimensions and on perceived mental and physical health status. Structural equation modeling (SEM) was used to fit the model and to obtain the parameter estimates. The Pratt index (Thomas, Hughes, & Zumbo, 1998) was used to determine the relative importance of perceived mental and physical health status and the dimensions of life satisfaction in relation to global QOL.

In summary, we designed multiple analyses to examine the relationships among global QOL, multidimensional life satisfaction, and perceived physical and mental health status in adolescents. We used CFA and FMA to test the assumptions underlying the indirect reflective structure for the measurement of multidimensional life satisfaction in adolescents. SEM techniques subsequently were used to examine the degree to which satisfaction with family, friends, school, living environment and self contribute to global QOL. The Pratt index (Thomas, Hughes, & Zumbo, 1998) was used to determine the relative importance of these dimensions of life satisfaction. Similar approaches were used to examine whether perceived mental or physical health status also contribute to global QOL. In addition, we examined whether the relationships between mental and physical health status and global QOL were mediated by the life satisfaction dimensions.

Figure 7 Structural parameters for the mediation model



Notes: The measurement structures for the dimensions of life satisfaction and global QOL are the same as those in Figure 6. Mental and physical health status were both specified as latent variables so as to be able to specify their ordinal nature, as discussed in the section on statistical methods.

We discuss the statistical methods for the above analyses in detail in section 2.5 because these methods represent the theoretical premises upon which our analyses were based. However, before doing so, we first provide a description of the survey methods and the items measured to obtain the data employed in our study.

2.2 Survey methods

The data for our analyses were taken from the British Columbia Youth Survey on Smoking and Health 2 (BCYSOSH II) which was completed for a research project entitled "Exploring Gender Differences in Tobacco Dependence among Adolescents".[10] Ethical approval was obtained from the Behavioral Research Ethics Board at the University of British Columbia (see Appendix A). This was a cross-sectional health survey of adolescents in grades 7 to 12 who were attending schools in regional school districts throughout the province of British Columbia (BC), Canada. At the time, there were 60 school districts in BC and 19 of these were contacted to obtain permission for schools within their jurisdiction to participate. This resulted in a sample of 89 eligible schools from 14 school districts that provided permission (five school districts did not provide permission). Out of the 89 schools, 49 (57%) agreed to participate. These included 42 secondary schools, 5 alternative schools, and 2 middle schools that included grade 8 students. A non-probability sampling approach was used within each of the schools. Sampling of adolescents in 22 of the 49 was completed by inviting the entire student body or students in particular grades to participate. In the remaining 27 schools, the selection of participants primarily occurred by recruiting students in courses that were taken by most students. We anticipated that these sampling strategies

[10] The description of survey methods was primarily taken from publications by Tu, Ratner, & Johnson (in press) and Richardson et al. (in press). Dr. Joy L. Johnson, who co-authored these publications, was the principal investigator for the survey.

would result in a sample of adolescents who were at various stages of development, who lived in diverse geographical locations and living conditions, and who reflected the diversity of cultures represented throughout the province.

The survey was administered by trained research assistants during class-time hours in pen and paper format (79.6%) or through an online format (20.4%). The format in which the survey was administered was primarily determined by the availability of a computer lab in the various schools. The research personnel recorded class attendance so that a response rate within each school could be determined. The students were informed that their responses to the survey questions were confidential and that they had the right not to participate or to refrain from answering any of the questions. Passive parental consent was obtained by providing parents with a letter that informed them about the survey. Less than 1% of the students refused to participate and the response rate within schools was 84%, on average (non-response was mostly due to student absenteeism) (Richardson et al., in press; Tu et al., in press). The resulting sample consisted of 8,225 adolescents who completed the survey. Most adolescents (51%) completed the survey in 37 minutes or less.

2.3 Measurement

The survey questionnaire included items pertaining to smoking, alcohol and other drug use, family life, health status, quality of life and various demographic characteristics. The variables of interest to our study were primarily derived from questions pertaining to the health status, life satisfaction, and global QOL of adolescents. In addition, we used several demographic variables and variables pertaining to the adolescents' experiences at school. The sections of the paper version of the questionnaire that included the variables of interest to our analyses are included in Appendix B.

2.3.1 Perceived mental and physical health status

Perceived health status was assessed by asking the adolescents to rate their physical health and their mental or emotional health as “excellent, very good, good, fair, or poor.” The resulting variable was therefore ordinal in nature. The validity of measuring perceived health status in this manner is supported by findings from studies of researchers who have used adolescents’ ratings of their health status to examine its relationship with various health outcomes, including physical activity, nutrition, health-risk behavior, disability (defined as the number of days of limited activity) and a composite health status index (Vingilis, Wade, & Adlaf, 1998; Vingilis, Wade, & Seeley, 2002; Wade, Pevalin, & Vingilis, 2000; Wade & Vingilis, 1999). Study findings have consistently revealed that increased perceived health status in adolescents is associated with less health-risk behaviors and fewer days of limited activity (Vingilis et al., 1998; Wade et al., 2000). Vingilis et al. (2002) found that a composite index of “eight different dimensions of health including vision, hearing, speech, mobility, dexterity, cognition, pain and discomfort, and emotion” (p. 196) had the largest effect on adolescents’ perceived health status when controlling for differences in age, sex, geographic location, income, family structure, disability, nutrition, psychological distress, physical activity and health-risk behavior.

To validate whether the adolescents in our sample distinguished between physical and mental health status in responding to these questions, we examined the relative importance of the two variables in explaining the variance in depressive symptoms, measured with 12 items of the Center of Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977) (see Appendix B, F10 items on p. 197). It was hypothesized that the explained variance in the total scores of the 12-item CES-D would be mostly attributed to mental health status. The results, as reported in Table 1, confirmed that these two variables explained 35.5% of the variance in depressive symptoms and that 94% of this explained variance could be attributed to mental health status relative to physical health status.[11]

Table 1 Relationships between perceived physical and mental health status and depressive symptoms (12-item CES-D)

Variable	*b*	*SE b*	*β*	*r*	*d*
Physical health	-0.46	0.09	-0.06	-0.33	6%
Mental health	-3.73	0.08	-0.56	-0.59	94%

Notes: *r* = estimated polyserial correlations for the relationships between ordinal perceived physical and mental health status variables and a continuous variable derived from the 12-item CES-D total score, *d* = Pratt Index. R^2 = 0.36. *N* = 7,985.

We similarly examined the relative importance of perceived physical and mental health status in explaining the variance in the adolescents' responses to the following question: "On how many of the last 7 days did you exercise or participate in sports activities for at least 20 minutes that made you sweat and breathe hard?" It was hypothesized that physical health status would be more strongly predictive of participating in physical activities than would mental health status. The results showed that, although physical and mental health status accounted for only 7.7% of the variance in frequency of physical activity, 82% of this explained variance could be attributed to physical activity, whereas only 18% was

[11] The 12-item CES-D total score was regressed on the perceived physical and mental health status variables. We used the Mplus 4.2 (B. Muthén & L. K. Muthén, 2006) software package to estimate the model parameters based on full information maximum likelihood estimation. The Pratt index (Thomas et al., 1998) was used to determine the relative importance of the perceived physical and mental health status variables in the model. In this case, the Pratt index values were based on the estimation of polyserial correlations. Estimation methods and methods for calculating the Pratt index are discussed in the section on statistical analyses.

attributed to mental health status (see Table 2). These findings provide further support for the claim that the adolescents in our sample distinguished between mental and physical health status.

Table 2 Relationships between perceived physical and mental health status and frequency of physical activity

Variable	*b*	*SE b*	*β*	*r*	*d*
Physical health	0.68	0.04	0.24	0.27	82%
Mental health	0.18	0.03	0.07	0.19	18%

Notes: *r* = estimated polyserial correlations for the relationships between the ordinal perceived physical and mental health status variables and the frequency of physical activity which was specified as a censored variable in the model, *d* = Pratt Index. R^2 = 0.08. *N* = 7,033.

2.3.2 Multidimensional life satisfaction

Huebner's 2001 version of the MSLSS was used to assess the adolescents' satisfaction with family (7 items), friends (9 items), school (8 items), living environment (9 items), and self (7 items). We followed Huebner's recommendation to use a 6-point response format with response options ranging from "strongly disagree" to "strongly agree." The items were included in the survey questionnaire in the order recommended in the MSLSS manual (Huebner, 2001) (see Table 3). Although Huebner (1994) originally recommended using a 4-point response format, subsequent studies provided support for the use of a 6-point response format with older children (grades 8 - 12) (Gilman et al., 2000; Huebner et al., 1998). The MSLSS includes 10 negatively worded items that were reverse scored prior to analysis. The

Table 3 Variables and item-wording for the MSLSS

Item # [1]	Variable name	Family subscale
7	fam1	I like spending time with my parents
8	fam2	My family is better than most
18	fam3	I enjoy being at home with my family
19	fam4	My family gets along well together
21	fam5	My parents treat me fairly
28	fam6	Members of my family talk nicely to one another
30	fam7	My parents and I do fun things together
		Friends subscale
1	frnd1	My friends are nice to me
4	frnd2*	I have a bad time with my friends
11	frnd3	My friends are great
12	frnd4	My friends will help me if I need it
16	frnd5	My friends treat me well
23	frnd6*	My friends are mean to me
24	frnd7*	I wish I had different friends
29	frnd8	I have a lot of fun with my friends
38	frnd9	I have enough friends
		Living environment subscale
15	lenv1	There are lots of fun things to do where I live
27	lenv2*	I wish I lived in a different house
31	lenv3	I like my neighborhood
32	lenv4*	I wish I lived somewhere else
34	lenv5*	This town/city is filled with mean people
36	lenv6	My family's house is nice
37	lenv7	I like my neighbors
40	lenv8	I like where I live
39	lenv9*	I wish there were different people in my neighborhood
		School subscale
3	schl1*	I feel bad at school
6	schl2	I learn a lot at school
9	schl3*	There are many things at school I don't like
13	schl4*	I wish I didn't have to go to school
20	schl5	I look forward to going to school

Item # [1]	Variable name	School subscale (cont'd)
22	schl6	I like being in school
25	schl7	School is interesting
26	schl8	I enjoy school activities
		Self subscale
2	self1	I am fun to be around
5	self2	There are lots of things I can do well
10	self3	I think I am good looking
14	self4	I like myself
17	self5	Most people like me
33	self6	I am a nice person
35	self7	I like to try new things

[1] Item number based on the order of the item on the survey questionnaire.

* Negatively worded items.

so-called α-coefficients of internal consistency (Cronbach, 1951),[12] in our sample were 0.92 for the family subscale, 0.91 for the friends subscale, 0.87 for the school subscale, and 0.85 for both the living environment and self subscales.[13] These coefficients were similar in

[12] This coefficient has generally been referred to as Cronbach's α. However, in his publication entitled "My current thoughts on coefficient alpha and successor procedures," Cronbach (2004) indicated that he never intended an eponymous coefficient. He wrote the following: "It is an embarrassment to me that the formula became conventionally known as Cronbach's α" (p. 397). Unfortunately, he did not provide alternative terminology. We therefore have respected his concern by referring to this coefficient as the α-coefficient of internal consistency.

[13] As discussed subsequently, the MSLSS items were treated as ordinal variables in our analyses. Conventionally, the α-coefficient of internal consistency is based on Pearson correlations. However, these would be negatively biased in our study. The α-coefficient of internal consistency was therefore based on polychoric correlations (rather than Pearson correlations) using the following formula: $\text{Alpha} = k\bar{r}/(1+(k-1)\bar{r}$, where k = the number of items for the subscale, and $\bar{r}$ = the average of the inter-item polychoric correlations. However, we also recognize that the α-coefficient of internal consistency may not be the best index of reliability considering the findings of our subsequent analyses.

magnitude to those reported in other reliability analyses of the MSLSS (Huebner & Gilman, 2002).

2.3.3 Global quality of life

Global QOL was measured using Cantril's (1966) self-anchoring ladder (referred to as the QOL-ladder) and a question that asked the respondents to rate their satisfaction with their quality of life. Cantril's self-anchoring ladder approach has been used to measure various global concepts including quality of life (Andrews & Robinson, 1991). The adolescents in our survey were asked to rate their life using a picture of a ladder with 8 rungs with the top rung representing "the best possible life for you" and the bottom rung representing "the worst possible life for you" (see Appendix B).[14] Global QOL was also measured by asking the adolescents to rate their agreement with the statement "I am satisfied with my quality of life" with four response options ranging from "strongly disagree" to "strongly agree."

The use of two observed variables as measures of global QOL allowed for the estimation of error variances for each of the observed variables (Schumacker & Lomax, 2004). Both global QOL variables were included as ordinal variables in our analyses. Their polychoric correlation was 0.53 (N = 7,528).

[14] An error in the web version of the questionnaire resulted in 10 rungs, instead of 8 rungs, being presented for the QOL-ladder. To remedy this, we re-scored the QOL-ladder for the web and paper versions to their common denominator by multiplying the web version QOL-ladder by 0.8 and rounding the resulting scores to zero decimals. In addition, 109 adolescents indicated a value between rungs on the paper version. Their responses were randomly scored upward or downward to the nearest integer.

2.3.4 Additional variables used to describe the sample

The adolescents were asked to indicate their age and gender, and to answer several questions related to their ethnic identity, their living arrangements, and their experiences at school. Ethnic identity was determined by asking the question "How would you describe yourself?" with 12 response options (including other) provided; the options were adapted from the classification of race by Statistics Canada (2003) (e.g., "white/Caucasian", Aboriginal/First Nation, Chinese, South East Asian). The adolescents could select multiple responses. Their responses were combined and collapsed into the following 4 categories: "white/Caucasian", Asian (including Chinese, Japanese, Korean, South East Asian, and Filipino), Aboriginal/First Nation, and other. The adolescents were also asked to report their country of birth and whether they spoke "another language" at home on a regular basis.

We provide a description of the various living arrangements of the adolescents because they might have responded differently to some of the MSLSS questions depending on their particular living arrangement (e.g., the word *family* in the family subscale might have been variously interpreted). They were asked, "Which parent or parents do you currently live with most of the time?" with 8 response options (i.e., mother, father, step-mother, step-father, guardian(s), foster parent(s), grandparent(s), and other please specify). Multiple responses were allowed.

Similarly, the adolescents might have responded differently to some of the items in the MSLSS (e.g., items in the school subscale) depending on how they compared themselves to other students at school. We therefore describe the adolescents' responses to questions about their experiences at school. We asked them to rate their school performance on a 7-point scale (ranging from "far below average" to "well above average") in response to the question, “Compared with other students in your school, how do you rate yourself in the school work you do?” We also asked them to evaluate the degree to which they felt like an outsider at school on a 5-point scale (ranging from “all the time” to “never”) in response to the question, “How often do you feel like an outsider (or left out of things at your school)?”

2.4 Data screening

The data were screened for implausible response patterns and missing responses. Data for 72 respondents that had the same value for all positively- and negatively worded MSLSS items were excluded from the analyses because these responses were deemed to be untrustworthy. Of the remaining 8,153 respondents that completed the survey, 5,269 (64.6%) completed all the MSLSS items, 1,056 (13.0%) had a missing value for only one of the items, 980 (12.0%) had missing values for more than one of the items, and 848 (10.4%) did not complete any of the items. A comparison of various characteristics of the missing data subsamples is presented in Table 4. In comparison with the other subsamples, the subsample with all MSLSS responses missing had a higher percentage of adolescents that were male, Aboriginal, in grades 7 or 8, reported below average school performance, and had lower satisfaction with their quality of life. Various statistical methods to account for missing data are discussed at the conclusion of the following section on the statistical methods of the analysis.

Table 4 Descriptive statistics of full sample and missing data subsamples

	Number of MSLSS Responses Missing				
	Complete Data	1 Missing	More than 1 Missing	All Missing	Total Sample
Variable	(*n* = 5,269)	(*n* = 1,056)	(*n* = 980)	(*n* = 920)	(*N* = 8,225)
Ethnicity: $\chi^2_{(9)} = 164.87$, $n = 7{,}882$, $p < .05$*					
White	75.8%	71.2%	68.7%	56.6%	72.6%
Asian	5.7%	6.1%	5.6%	6.7%	5.9%
Aboriginal	13.9%	16.2%	20.2%	30.7%	16.5%
Other or mixed	4.6%	6.4%	5.5%	6.1%	5.1%
Sex: $\chi^2_{(3)} = 81.65$, $n = 8{,}074$, $p < .05$*					
Male	47.4%	49.2%	50.5%	64.0%	49.8%
Female	52.6%	50.8%	49.5%	36.0%	50.2%
Grade: $\chi^2_{(12)} = 152.59$, $n = 8{,}163$, $p < .05$					
Grades 7 or 8	19.5%	26%	30%	35.1%	23.2%
Grade 9	19.7%	19.9%	19.8%	16.6%	19.4%
Grade 10	24.9%	24.2%	23%	16.9%	23.7%
Grade 11	22.6%	19%	16.3%	20.1%	21.1%
Grade 12 or "other"	13.4%	11%	10.9%	11.3%	12.6%
Perceived School Performance: $\chi^2_{(3)} = 54.7$, $n = 7{,}060$, $p < .05$*					
Average or below average	45.5%	52.0%	57.4%	62.0%	48.0%
Above average	54.5%	48.1%	42.6%	38.0%	52.0%
Satisfied with quality of life: $\chi^2_{(9)} = 52.3$, $n = 7{,}606$, $p < .05$*					
Strongly disagree	4.1%	4.9%	4.5%	10.3%	4.6%
Disagree	12.5%	12.5%	14.9%	16.2%	13.0%
Agree	53.3%	53.6%	50.9%	48.2%	52.7%
Strongly agree	30.1%	29.1%	29.8%	25.3%	29.6%

* Chi-square comparing the four subsamples of respondents classified by the number of missing MSLSS responses.

2.5 Statistical methodology and methods

Our statistical methods of analysis were based on statistical theory underlying confirmatory factor analysis (CFA), structural equation modeling (SEM), and factor mixture analysis (FMA). We begin with a discussion of the CFA techniques used to examine the measurement structure of the MSLSS. We discuss several key mathematical formulas in some detail with the purpose of explaining the statistical relations and the basis of the previously discussed assumptions underlying the indirect reflective measurement structure. The statistical relations are illustrated using path diagrams. We also discuss the estimation methods and guidelines for the assessment of model fit, which were used to fit the CFA models and the SEMs. FMA techniques are subsequently introduced as a method to examine any sample heterogeneity with respect to the MSLSS subscales. Again, a few key formulas are discussed with the purpose of illustrating the statistical relations in these models. We then introduce the Pratt index (Thomas et al., 1998) as a measure of the relative importance of the variables in our structural equation models. Our discussion of statistical methods concludes with an overview of the techniques that were used to account for the missing data in our analyses.

2.5.1 Confirmatory factor analysis methods

Confirmatory factor analysis can be defined as a statistical technique for determining whether a set of observed variables relate to one or more latent factors that are specified *a priori*. A latent factor can be defined as an unobserved variable that accounts for the correlations among two or more observed variables. The observed variables are often referred to as indicators of the latent factor(s). Our discussion of CFA focuses on model specification of the indirect reflective measurement structure (also referred to as a second-order factor structure) for ordinal variables, estimation methods, and guidelines for the evaluation of model fit.

2.5.1.1 Specification of the indirect reflective measurement structure

The indirect reflective measurement structure (see Figure 4 on page 38) for the 40 MSLSS indicators was specified by combining a first-order factor CFA model for the relationships among the observed indicators and the five latent factors (Equation 1), and a structural model specifying the relationships among the five first-order factors and the second-order factor (Equation 2). The first-order factor structure is shown in the following equation (as, for example, discussed by Byrne (1998)):

$$\mathrm{Y} = \Lambda_y \eta + \varepsilon \,, \tag{1}$$

where Y is a 40 x 1 vector of the 40 variances of the observed variables, Λ_y is a 40 x 5 matrix representing the regression coefficients between the 40 observed variables and their corresponding latent factor, η is a 5 x 1 vector of the latent factor variances, and ε is a 40 x 1 vector of the residual variances (i.e., measurement error) for the observed variables (which are assumed to be independent). Each observed variable (y_i) relates to only one of the five latent factors (η) and the relationships between y_i and the other latent factors are therefore specified as zero. In this sense, the first-order factor structure represents assumptions two and three underlying the indirect reflective measurement structure as discussed in the general overview of our analytical approach.

The relationships between the first-order and second-order factors (which reflect assumptions four and five) can be defined as follows (Byrne, 1998):

$$\eta = \Gamma \xi + \zeta \,, \tag{2}$$

where Γ is a 5 x 1 vector of the regression coefficients between the second-order exogenous factor and each of the first-order endogenous factors, ξ is a 1 x 1 vector of the second-order exogenous factor variance, and ζ is a 5 x 1 vector representing the residual variance for each of the first-order latent factors (also referred to as disturbance terms, which are independent in the model).

By substituting η in Equation 1 (first-order factor structure) with the right side of Equation 2 (second-order factor structure) we obtain the following equation for the indirect reflective model:

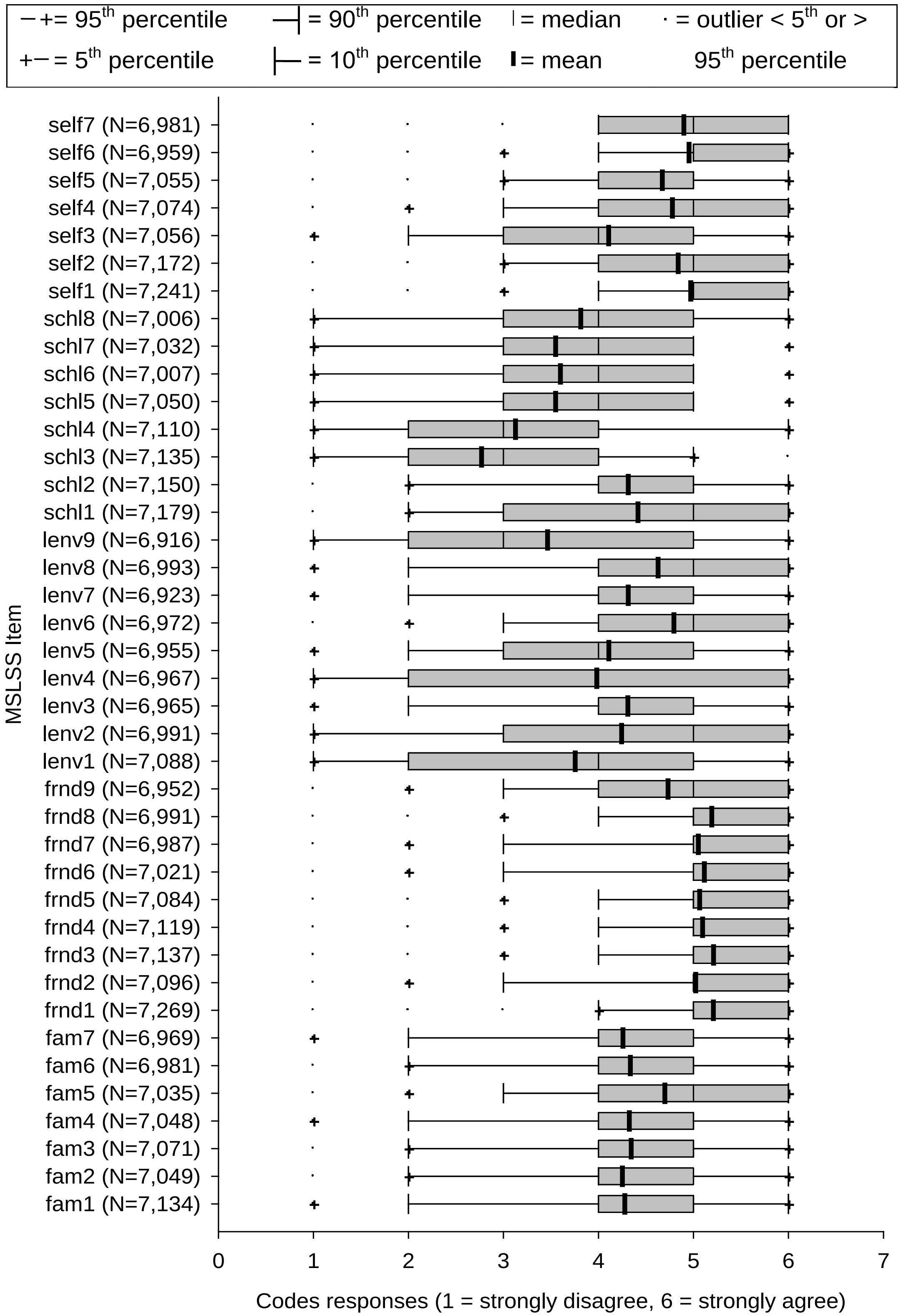
$$\Upsilon = \Lambda_y(\Gamma\xi + \zeta) + \varepsilon \qquad (3)$$

Thus, we specified 40 observed indicators that were uniquely associated with only one of five first-order latent factors, which in turn were associated with one common second-order factor. The variances for all of the latent factors were specified to equal one to avoid indeterminancy (Schumacker & Lomax, 2004).

2.5.1.2 Representation of ordinal variables

We examined the distributions of the ordinal variables to determine whether they could be treated as continuous variables in our analyses (see Figure 8). The distributions of the MSLSS variables were negatively skewed in the direction of increased satisfaction for most items (skewness < -1.9) except for three negatively worded items that were positively skewed (skewness ranging from 0.1 to 0.6). Kurtosis ranged from -1.3 to 5.3. Although these values may be considered to be within an acceptable range for confirmatory factor analysis with maximum likelihood estimation (e.g., skewness < |2.0| and kurtosis < |7.0|) (Fabrigar, MacCallum, Wegener, & Strahan, 1999), the boxplots, histograms and QQ plots clearly revealed the non-normal and discrete nature of the data (see boxplot graphs in Figure 8).The

Figure 8 Box plots of the distributions of the MSLSS variables



resulting deviations from normality, which most certainly violated the assumption of multivariate normality, could have led to inflated chi-square statistics, underestimated parameter estimates, and downwardly biased standard errors thereby increasing the risk of type I error (i.e., rejecting a well-fitting model) (Finney & DiStefano, 2006; Flora & Curran, 2004). Although the skewness and kurtosis for some of the other variables in our analyses were somewhat more acceptable (e.g., the QOL-ladder as well as the perceived physical and mental health status variables), their distributions still clearly reflected their underlying ordinal nature. We therefore modeled all variables as ordinal variables in our analyses.

The ordinal nature of the observed MSLSS indicators (as well as the other variables in our subsequent analyses) was modeled by specifying discrete representations of an underlying normally distributed latent response variable (denoted as y^*) for each observed variable (see, for example, V. E. Johnson & Albert, 2004; Millsap & Yun-Tein, 2004; Muthén, 2004). As was aptly explained by Muthén (2004), “[This] latent response variable formulation focuses on the linear relation between y^* and x instead of the non-linear relationship between y and x” (p. 3). The relationship between the observed response y and y^* can then be expressed as follows:

$$y = \mathrm{C}, \text{ if } \tau_c < y^* \leq \tau_c + 1,$$

for categories $c = 0, 1, 2, \ldots, \mathrm{C} - 1$ and $\tau_0 = -\infty$ and $\tau_c = \infty$, where C represents the ordinal response categories and t_c represents the C - 1 thresholds. Thus, in Equation 3, the vector Y can be replaced by Y^* to indicate that the distributions of y^* are not observed but estimated based on the observed ordinal-item responses.

The estimated distribution of y^* can be obtained by using the probit link function to estimate the cumulative probabilities for $\mathrm{Y} \geq c$ ($c = 1, \ldots, \mathrm{C}$). The expected cumulative probability (θ) for a particular item response *c* for a given individual *i* is shown in the following equation, adapted from Johnson and Albert (2004):

$$\theta_{ic} = \pi_{i1} + \pi_{i2} + ... + \pi_{ic}\,, \tag{4}$$

where π_{ic} is the probability of individual *i* responding with category C - 1 for a particular item (θ_{ic} is defined by C - 1 response categories because the cumulative probability of the last response category is, obviously, equal to 1.0). Using the probit link function, the cumulative probability can be expressed as follows (Johnson & Albert; Skrondal & Rabe-Hesketh, 2004):

$$\theta_{ic} = \Phi(-t_c + \eta'\lambda)\,,^{15} \tag{5}$$

where Φ is the standard normal cumulative distribution function, t_c represents the C - 1 thresholds for the ordinal response categories of the observed variables, η is a vector of the explanatory variables (in our case, η is a single number representing the first-order latent factor that is associated with the observed variable), and λ represents the regression parameters for each of the observed variables (in our case, λ is an element in Λ_y from Equation 1). The probit regression coefficients therefore represent a one unit increase in the *z* scores of the distribution of the cumulative probabilities for each of the observed variables. These regression coefficients are the same across all ordinal categories for each of the observed ordinal variables (also referred to as the "parallel slopes" assumption (Borooah, 2002)).

[15] Probit regressions are more typically expressed as follows: $\theta_{ic} = \Phi(-t_c + x'\beta)$ (e.g., V. E. Johnson & Albert, 2004; Skrondal & Rabe-Hesketh, 2004). In Equation 5, *x* is replaced by η and *β* is replaced by λ with the intent of clearly linking the probit regression model to the CFA model in Equations 1 to 3.

2.5.1.3 Estimation methods and guidelines for the evaluation of model fit

We used a robust mean and variance adjusted weighted-least squares estimation method (WLSMV) (Muthén, 2004), which is included in the Mplus 4.2 software package (B. Muthén & L. K. Muthén, 2006). This estimation method has been recommended for use with ordinal data resulting in more robust χ^2 statistics, less bias in the parameter estimates, and lower Type I error rates (Beauducel & Herzberg, 2006; Finney & DiStefano, 2006). Models were also estimated using a robust full information maximum likelihood estimation method (MLR) to account for the missing data (see discussion of missing data techniques).

Model fit was assessed by examining the chi-square statistic and the global fit indices, and by comparing the differences between the implied and the observed polychoric correlation matrices. The chi-square statistic reflects the difference between these two matrices and should therefore ideally be statistically non-significant.[16] However, it is widely recognized that the chi-square statistic is biased with increased degrees of freedom, sample size and deviations from multivariate normality (e.g., Schumacker & Lomax, 2004). We therefore also used the following non-statistical fit indices to examine model fit: (a) the root mean square error of approximation (RMSEA), (b) the comparative fit index (CFI), and (c) the standardized root mean residual (SRMR). Some general guidelines for evaluating model fit include that the RMSEA should not exceed 0.08, and that a value less than 0.06 is indicative of a well-fitting model, that the CFI should not be smaller than 0.95, and that the

[16] The number of degrees of freedom for the evaluation of the statistical significance of the chi-square statistic is normally based on the difference between the number of distinct elements in the covariance matrix and the number of free parameters (e.g., Schumacker & Lomax, 2004). However, the degrees of freedom for some robust chi-square statistics (e.g., the Satorra-Bentler chi-square and the WLSMV chi-square) are different because they are adjusted to account for a scaling correction factor (Muthén, du Toit, & Spisic, 1997). In the tables presented, we provide the number of estimated as well as the model-implied degrees of freedom.

SRMR should not exceed 0.08 (e.g., Hu & Bentler, 1999; Schumacker & Lomax; Vandenberg & Lance, 2000).

Although the above guidelines for model fit were initially based on the estimation of models with continuous observed variables, several published simulation studies provide support for the application of similar guidelines to models based on ordinal categorical variables. For example, Beauducel and Herzber (2006) found that the RMSEA based on WLSMV estimation was very similar, but slightly larger, compared with the ML based RMSEA when variables with five or six categories were used. They also found that the CFI was the same when using WLSMV and ML estimation in their CFA simulation study involving ordinal variables with five or six categories. Similarly, results of a simulation study by Yu (2002) indicated that the Comparative Fit Index (CFI) performed better than the RMSEA for binary and non-normal continuous data and that rejection rates were acceptable with a cutoff value for the CFI close to 0.96. Yu found that the SRMR did not perform well for binary data, and recommended that a weighted root mean residual (WRMR) should be used instead. However, Beauducel and Herzberg (2006) found that the WLSMV based SRMR was very similar to the ML based SRMR for factor models that included variables with five or six ordinal categories. Use of the SRMR therefore seems to be justifiable in our study, which included variables with four to eight ordinal categories. The results of these simulation studies suggest that the previously mentioned cutoff guidelines for fit indices based on continuous variables and ML estimation could be cautiously applied to the use of WLSMV estimation based on ordinal variables with six categories.

Although we acknowledge these rules of thumb for evaluating model fit, we also concur with Marsh, Hue, and Wen (2004) who cautioned against the overgeneralization and uncritical application of cutoff values for goodness of fit indices. Ultimately, the assessment of model fit should be based on a direct comparison of the observed and implied correlation matrices. To facilitate this comparison, we mapped the residual correlations (i.e., observed minus implied correlations) so as to reveal potential areas of misfit. We also examined the range of residual correlations as well as the percentage of residual correlations with absolute values greater than 0.1 to evaluate the degree of misfit.[17] We found that these values provided a useful summary of the distribution of correlations in our study. We caution, however, that good fit is reflected in the overall pattern of residual correlations, and that a model may therefore still not fit well even if all the residual correlations have absolute values smaller than 0.1.

In addition to these guidelines for assessing overall model fit, we also examined the difference in fit between competing models. To do so, we used the widely recommended technique of comparing the chi-square statistics of two nested models to test the significance of the differences between the models (Schumacker & Lomax, 2004).[18] We also examined the fit of competing models by comparing the range of residual correlations and the

---

[17] The value of 0.1 was merely chosen as a means of providing more information about the distribution of the residual correlations. Clearly, the magnitude of a residual correlation needs to be interpreted in light of the magnitude of the corresponding observed correlation. Nevertheless, although no specific cut-off criteria for evaluating residual correlations exist; residual correlations greater than 0.1 would certainly indicate that the model does not fit well.

[18] Although the difference in chi-square can readily be evaluated based on the corresponding difference in degrees of freedom based on the number of free parameters and distinct values in the covariance matrix, it is important to point out that the chi-square differences based on WLSMV estimation are not distributed as a chi-square and therefore cannot be directly evaluated (L. K. Muthén & B. Muthén, 2006). The techniques for estimating the significance of difference in WLSMV chi-square statistics for comparing nested models were discussed by Asparov and Muthén (2006). We used the DIFFTEST in Mplus 4.2 (B. Muthén & L. K. Muthén, 2006) software to implement this technique.

percentage of residual correlations with absolute values greater than 0.1. In addition, the Bayesian Information Criterion (BIC) was used to compare the models that were derived with full information maximum likelihood estimation.[19] The BIC is based on a comparison of the maximum likelihood value while taking the number of free model parameters into account and is particularly useful for the comparison of models that are not nested (Muthén, 2004). A smaller BIC value corresponds to a larger likelihood and is therefore indicative of better model fit.

2.5.2 Factor mixture analysis methods

Factor mixture analysis was used to examine potential sample heterogeneity with respect to the measurement structure for each subscale of the MSLSS. A factor mixture model can be viewed as a combination of a latent class model and a factor model (Lubke & Muthén, 2005). In a latent class model, individuals are clustered so as to maximize local independence among the observed variables conditional upon a latent class variable (Hagenaars & McCutcheon, 2002; Magidson & Vermunt, 2004). The general form of a latent class model can be defined as follows for three observed categorical variables (Hagenaars & McCutcheon):

$$\pi_{c_1c_2c_3k}^{Y_1Y_2Y_3X} = \pi_k^X \pi_{c_1k}^{Y_1|X} \pi_{c_2k}^{Y_2|X} \pi_{c_3k}^{Y_3|X}, \quad \text{for } c_1, c_2 \text{ and } c_3 = 1, \ldots, 6;\ k = 1, \ldots, K \tag{6}$$

where Y_1, Y_2 and Y_3 are the ordinal variables with response categories denoted by c_1, c_2 and c_3 having values ranging from 1 to 6, and X is a latent class variable with k classes. In this model, $\pi_{c_1c_2c_3k}^{Y_1Y_2Y_3X}$ is the probability of independently obtaining a particular set of responses for variables Y_1, Y_2 and Y_3 conditional upon the latent class variable X. The baseline model of no

[19] As explained by Muthén (2004), the BIC is calculated based on the following formula: $BIC = -2 \log L + r \ln n$, where r refers to the number of free parameters and n refers to the sample size.

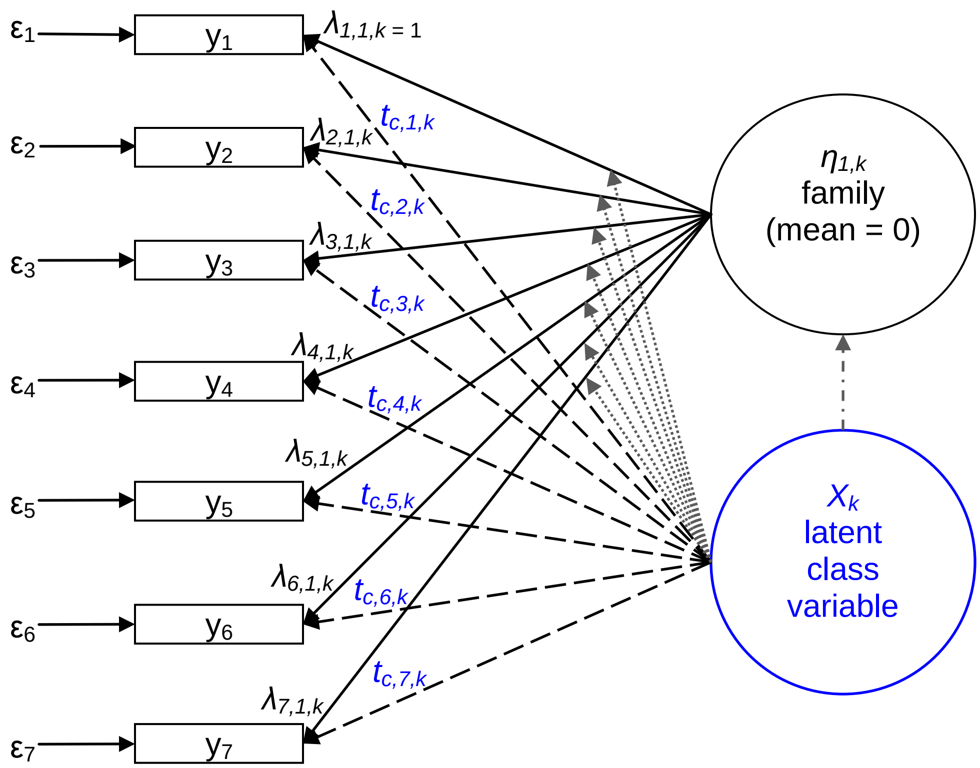
latent classes is defined by mutual independence among the ordinal variables (i.e., $\pi_{c_1 c_2 c_3}^{Y_1 Y_2 Y_3} = \pi_{c_1}^{Y_1} \pi_{c_2}^{Y_2} \pi_{c_3}^{Y_3}$) (Magidson & Vermunt). Therefore, as in a latent factor model where the observed variables are specified as being independent conditional on a common factor (i.e., a continuous latent variable), the observed variables in a latent class model are independent for all individuals within the same latent class (i.e., conditional on a categorical latent variable).

By combining the factor analysis model with a latent class model, one obtains a clustering of individuals so as to maximize independence among the observed variables conditional on the class-specific measurement model parameters (Lubke & Muthén, 2005). The factor mixture model specified in our analyses was used to specifically address the following two related issues: (a) whether the sample was heterogeneous with respect to the measurement structure and (b) whether the measurement model parameters were invariant across two or more latent classes. The first issue was addressed by determining the number of latent classes needed to achieve the best model fit, and the second issue was addressed by comparing the model parameters across the latent classes to determine which parameters were least invariant.

We specified the factor mixture models for each subscale by allowing the thresholds, factor loadings and factor variances to vary across two or more latent classes (see Figure 9). The path coefficients from the factor to the observed variables are given by Equation 2, with the addition that these are now conditional on latent class membership as represented by the dotted lines connecting the categorical latent class variable, C, to each of the factor coefficients. In other words, the factor loadings and factor variances are conditional on the *unordered* categorical latent class variable, as is heuristically shown in the following equation for a first-order factor structure:

$$\mathrm{Y}^{*} = \Lambda_{yk}\eta_k + \varepsilon_{yk}, \text{ for } k = 1, \ldots, K \text{ latent classes.} \qquad (7)$$

Figure 9 A diagram of the factor mixture model for the family subscale



λ parameters for relationships between the latent factor and the observed variables.

t parameters for c – 1 thresholds that are conditional on latent class variable X_k, k = 1, …, K.

Indicates that the λ parameters are conditional on latent class variable X_k, k = 1, …, K.

Indicates that the variance of the latent factor, η, is conditional on latent class variable X_k, k = 1, …, K – 1.

K is the number of latent classes and C is the number of ordinal categories for the observed variables.

It is important to recall that the observed variables in our study were treated as *ordinal* variables. The factor loadings can therefore only be interpreted when also taking into account the thresholds of each observed item, and we cannot assume that these thresholds are indeed invariant (Millsap & Yun-Tein, 2004).[20] These thresholds were therefore allowed to vary across the latent classes (thereby testing the first assumption). This conditional dependence of the thresholds on latent class memberships is represented by the paths from the latent class variable to each of the observed variables in Figure 9.

When we initially used the 4.0 version of the MPlus software (B. Muthén & L. K. Muthén, 2006), it was not possible to specify probit regression coefficients for latent class mixture models. Polynomial (ordinal) logistic regression was therefore used to model the relationships between the observed variables and their corresponding first-order latent factors. The cumulative probabilities of Equation 4 were defined in terms of a polynomial logistic regression as shown in the following equation (as, for example, discussed by Skrondal and Rabe-Hesketh (2004)):

$$\theta_{ic} = \exp(-t_c + \eta' \lambda) \tag{8}$$

Although not explicitly shown in the above equation, all the parameters for the ordinal logistic regression are conditional on the latent class variable as shown in Figure 9 (page 73). The factor mean was constrained to equal zero so that the thresholds could be meaningfully compared across latent classes. That is, the thresholds represented the average cumulative log odds for a particular item response given that the average factor score was zero. In addition, one factor loading was constrained to equal 1.0 in all latent classes to obtain a just- or over-identified model. Further details about model specification and the number of estimated parameters are provided in Table 5.

---

[20] As pointed out by Millsap and Yun-Tein (2004), the conditions for measurement invariance of a single factor structure with ordinal variables are that the parameters of the factor model (i.e., the factor loadings and factor variance), as well as the thresholds for the ordinal variables are invariant across two or more sample populations (in our case, latent classes).

Table 5 Specification of a single-factor mixture model with ordinal variables

Parameters	Constraints	# of parameters
Thresholds	All free to vary across classes	$pk(c-1)$
Unique factor loadings	One factor loading is set to 1 for each latent class	$k(p-1)$
Factor means	Constrained to zero for all latent classes	0
Factor variances	Free to vary across all latent classes	k
Latent class means	One is constrained	$k-1$

[1] Notes: p = number of observed variables, c = number of ordinal response categories, k = number of latent classes.

The model parameters were estimated using a robust maximum likelihood estimation method (MLR) available in the Mplus software package (B. Muthén & L. K. Muthén, 2006). A common problem associated with the use of maximum likelihood estimation methods for mixture models is that a local maximum, rather than a global maximum, of the likelihood function can be obtained (McCutcheon, 2002). To address this problem, we used randomly generated seeds to specify 500 sets of parameter starting values. Of these, the most plausible sets of starting values were chosen based on 50 random seeds that resulted in the greatest likelihood values after 20 iterations using the expectation maximization (EM) algorithm. The random seed that resulted in the greatest log likelihood for the converged model was chosen to obtain a solution for the parameters in the model.

Ideally, the greatest log likelihood would be replicated to provide further support for the conclusion that a true global maximum had been obtained. However, it is possible that the greatest log likelihood may not be replicated even if a large number of randomly generated starting values are used. In these situations, we used starting values derived from additional random seeds to determine whether the log likelihood could be replicated. In situations where this was not the case, we examined whether the parameter estimates resulting from the greatest log likelihood were similar to those obtained from the next greatest log likelihood. Ultimately, the solution with the greatest log likelihood and with parameters that were similar to those obtained from a solution with the next greatest log likelihood was considered to be the best solution (McCutcheon, 2002; L. K. Muthén & B. Muthén, 2006).

Model fit was further assessed by comparing a solution for a model with *K* classes to a solution for a model with *K* - 1 classes to determine the best fitting and most parsimonious model (i.e., the model with the least number of classes). We relied primarily on the BIC for these comparisons (Nylund, Asparouhov, & Muthén, 2006). Nylund et al. compared various statistical information criteria and likelihood ratio indices that have been used to determine the correct number of latent classes. The results of their simulation study showed that the BIC consistently identified the correct number of latent classes for categorical observed variables for sample sizes of 1,000. The sample-size adjusted BIC was more reliable for smaller sample sizes. Considering the size of our sample, the BIC was considered to be adequate for our purposes.[21]

In addition to using the BIC to evaluate comparative fit, we compared the predicted proportions of class membership based on the posterior probabilities. Adding another latent

[21] Although the difference in -2 log likelihood values for *k* and *k* - 1 classes can also be examined using the Lo-Mendell-Rubin likelihood ratio test (LRT) or a bootstrapped LRT, these methods were not relied upon for our analyses because the simulation study by Nylund et al. (2006) showed that the Lo-Mendell-Rubin test had higher type I error rates for categorical observed variables and because the bootstrapped LRT resulted in a significant increase in computational time.

class may not be substantively meaningful if one of the class proportions is very small (Nylund et al., 2006). We also examined the internal consistency and the fit of the first-order factor structures of each subscale within each of the classes independently so as to determine whether within-class model fit of the measurement structure was substantially improved when more classes were specified.

2.5.2.1 Assessment of parameter invariance

We used non-parametric graphical methods to determine which observed indicators were least invariant across the latent classes. A comparison of the thresholds across latent classes was obtained by graphing the threshold values and 95% confidence intervals of each observed variable within the latent classes along the Y-axis and the threshold number along the X-axis. A relative lack of invariance could be identified by assessing those items with the most discrepant threshold values across thresholds. A difference in the mean of the latent variable underlying an observed indicator would be indicated if the discrepancies in the values of a particular threshold were of similar magnitude for all five thresholds of that observed indicator (i.e., the lines connecting the thresholds were parallel). A difference in the distribution of the latent variable underlying an observed indicator would be indicated if the discrepancies for the five thresholds of that observed indicator were not consistent in magnitude (i.e., the lines connecting the thresholds were not parallel).

In summary, the second-order factor structure, as defined by Equation 3, was first examined with CFA to specifically test assumption five in the overall sample (based on the premise that the other assumptions were justified). The first-order factor structures, for each subscale, as defined by Equation 1, were subsequently examined independently to specifically test assumption two (based on the premise that assumption one was justified).

FMA was used to test assumptions one and three by examining the degree to which the thresholds and factor loadings of the ordinal regressions for each observed indicator were consistent across two or more latent classes. For each subscale, the four observed indictors with the most consistent thresholds were chosen to construct new (abridged) measures that were relatively more reliable in our sample. These subscales were then recombined into an indirect reflective measurement model to determine whether they reflected a common second-order factor, which we labeled general life satisfaction. Assumption four was not directly tested in our analyses, but inferences about the plausibility of this assumption could be drawn based on our examination of the other four assumptions.

2.5.3 Additional statistical methods

The methods discussed thus far focused on fitting the latent variable models. Once good model fit was obtained, we could then use the estimated parameter estimates to calculate indirect effects, the degree of mediation, and the relative importance of the variables explaining global QOL. The "tracing rule" was applied to calculate the magnitude of the indirect effects (Cohen, Cohen, West, & Aiken, 2003, p. 461). According to this rule, the total effect for an explanatory variable, A, in relation to a response variable, B, can be calculated by adding all the direct and indirect effects for the relationships between A and B, and values for the indirect effects can be obtained by multiplying the coefficients for the relationships between the variables that comprise the indirect relationships between A and B. We used the Delta method to calculate the standard error associated with this indirect effect (Mackinnon & Dwyer, 1993).

We also were interested in evaluating the degree to which the relationships between physical and mental health status and general life satisfaction were mediated by satisfaction with family, friends, school, living environment and self. One approach is to define the degree of mediation as a percentage of the degree to which the total effect is mediated (Mackinnon & Dwyer, 1993; Mackinnon, Fairchild, & Fritz, 2007). This percentage mediation attributed to a mediating variable can be easily determined by dividing the indirect effect by the total effect (and multiplying by 100%). The total percentage mediation for all mediating variables combined can be determined by dividing the sum of all the indirect effects for that relationship by the corresponding total effect.

Finally, the Pratt index (Thomas et al., 1998) was calculated to determine the relative importance of the variables that were found to explain global QOL. The Pratt index provides an additive measure of the importance of explanatory variables in a general linear model. Although the Pratt index has been primarily used for observed variable regression models, Zumbo (2007) pointed out that the generic formulation of the Pratt index based on a general linear model can also be used for latent variables. A Pratt index value for each variable was obtained by multiplying the standardized regression coefficients by the corresponding correlations and dividing that value by the total explained variance as shown in the following equation:

$$d_k = \frac{\beta_k r_k}{R^2}, \quad \text{for } k = 1, \dots, n \tag{9}$$

where k represents one of the explanatory variables, β represents the standardized regression coefficient for that variable, r represents the corresponding correlation, and R^2 represents the total variance explained. The correlations among the latent factors were obtained by estimating a corresponding model with correlated latent factors. The resulting Pratt index values sum to 100% and therefore provide an estimate of the relative percentage of explained variance that can be attributed to each of the explanatory variables.

2.5.4 Missing data techniques

Missing data techniques constitute another important component of the statistical methods that we used. It is widely recognized that uncritical approaches to dealing with missing data can lead to biased parameter estimates and unwarranted conclusions. Conventional approaches, such as listwise and pairwise deletion, result in inflated standard errors and biased parameter estimates if the missing data are not missing completely at random (MCAR) (Allison, 2002; Enders, 2006). Other problems with these approaches are that listwise deletion can result in the exclusion of a large amount of information and that the sample size for pairwise deletion is unspecified (because every covariate pair is potentially based on a different sample size). The latter can lead to non-positive definite matrices in CFA and SEM.

We compared the results of our analyses based on the following three missing data techniques so as to assess the degree to which the assumptions underlying each technique might have influenced our conclusions: (a) single imputation using the EM algorithm for a subsample of those who had no more than one missing value for the MSLSS items (N = 6,325), (b) full information maximum likelihood (FIML) for the subsample of respondents with values for at least one of the observed variables (for the CFAs, this resulted in a sample size of 7,305), and (c) multiple imputation based on the expectation maximization (EM) to impute data based on all relevant available information. We did not rely on mean imputation because this missing data technique relies on the implausible assumption that the missing data were missing completely at random (MCAR). Our comparison of missing data subsamples already revealed that this was clearly not the case (see Table 4 on page 61). In addition, even if the assumption of MCAR were justified, simulation studies have shown that mean imputation invariably results in biased parameter estimates (Enders, 2006).

2.5.4.1 Single imputation using the EM algorithm

The first approach involved using the EM algorithm to impute data for those who had a missing value for no more than one MSLSS item. We used this approach to obtain a single dataset which could be used in our FMAs.[22] We desired to impute only a very small amount of data because single imputation based on the EM algorithm may result in biased standard errors and model fit indices. Enders (2006) explained that this bias arises because the sample size of the EM covariance matrix is undetermined. Thus, to avoid significant variation in sample size, we imputed only a single value for those 1,056 respondents who had a missing value for one and only one of the MSLSS items. This resulted in the most gain in sample size with a minimal percentage of imputed data (0.42% imputed data).

The use of the EM algorithm ensured that the most likely response for an item was imputed based on the values of the other 39 items for all individuals in the dataset (this is different from mean imputation where the imputed value would be based on the values of the other 39 items for the particular individual). However, this approach involved excluding 23.1% of the respondents by listwise deletion which is based on the assumption that the missing values for these respondents were missing completely at random (MCAR). In addition, the MSLSS items that these respondents did answer were ignored which resulted in a significant loss of available information. The degree to which these limitations affected our findings can be assessed by comparing the parameter estimates to those obtained when using other missing data techniques.

[22] Although multiple imputation may seem like a better alternative, we required a single dataset for the mixture analyses.

2.5.4.2 Full information maximum likelihood

We compared the results based on single imputation to those obtained when using FIML for the subsample of youth who completed at least one MSLSS item (N = 7,305 for the CFAs for the MSLSS)[23] so as to determine whether the findings might be biased by missing data mechanisms that were unaccounted for. When using FIML, all available values for the variables in the dataset are used to obtain maximum likelihood estimates of the parameter estimates and their standard errors (i.e., FIML does not actually involve imputation of data) (Enders, 2006). FIML is based on the assumption that data are missing at random (MAR), which is less restrictive than the previously mentioned MCAR assumption.[24] Another advantage of this approach is that the ordinal nature of the observed variables can be taken into account by specifying these as ordinal variables in the model to be analyzed. However, a limitation is that it is difficult to include so-called "auxiliary variables" in the estimation (Enders, 2006, p. 320). Auxiliary variables are those that are not part of the intended model but that may provide information that relates to the patterns of missingness among the variables in the model. As pointed out by Enders, biased parameter estimates may result if the patterns of missingness are not conditioned on auxiliary variables associated with the observed variables that have missing responses.

---

[23] Obviously, the sample size varies depending on the variables that are included in the model.

[24] The MAR assumption is less restrictive than the previously mentioned MCAR assumption. As explained by Allison (2002), "The data on Y are said to be missing completely at random (MCAR) if the probability of missing data on Y is unrelated to the value of Y itself or to the values of any other variables in the data set" (p. 3). In contrast, "Data on Y are said to be missing at random [MAR] if the probability of missing data on Y is unrelated to the value of Y, after controlling for other variables in the analysis" (p. 4). Enders (2006) explained that the MAR assumption can be relaxed by including auxiliary variables that relate to the probability of having missing data on Y. The MCAR assumption is, therefore, clearly more restrictive.

2.5.4.3 Multiple imputation

In addition to FIML and single imputation, we used multiple imputation techniques that do allow for the inclusion of auxiliary variables. Inclusion of variables that are correlated with the analysis variables of interest can help to reduce bias and therefore improve the accuracy of the imputations (Enders, 2006). We specifically included the following variables in our multiple imputation model: all the MSLSS variables, the demographic variables (sex, ethnicity, grade), the two global QOL variables, the perceived mental and physical health status variables, depressive symptoms, and two variables pertaining to the adolescents' experiences at school in comparison with other students. The demographic variables were included because missingness differed to some degree across various demographic groups (see Table 4). The global QOL variables were included because they are theoretically related to the items of the MSLSS. Perceived mental and physical health status were included because of the associations between these variables and life satisfaction in adolescents (e.g., Zullig, Valois, Huebner, & Drane, 2005). We similarly conjectured that depressive symptoms may have been associated with the adolescents' responses to the health status, global QOL variables, and some MSLSS variables. Two additional variables, the adolescents' ratings of their school performance and the degree to which they felt respected in comparison with their peers, were included because they may have been be related to missingness on some of the other variables in the analyses (in particular those pertaining to satisfaction with school and self).

Although multiple imputation using the EM algorithm allows for the inclusion of auxiliary variables, a limitation of this approach is that the ordinal nature of the observed variables could not be accounted for in the imputation model. Nevertheless, in his review of guidelines for multiple imputation, Allison (2002) suggested that there is substantial evidence that multiple imputation techniques are robust to deviations from normality. He also suggested that the precision in estimation can be improved by limiting the range of imputed values to be equal to the observed range for each variable. Based on this restriction, the imputation process can be repeated until the random draws result in values that are within the observed range. The imputed values can then be rounded to zero decimals to match the discrete nature of the ordinal variables (Allison). A similar approach can be used to include nominal (in our case, sex and ethnicity) or ordinal (in our case, grade) as dummy-coded auxiliary variables in the imputation algorithm.[25]

We used the SAS 9.1 software package (SAS Institute, 2005) to create 10 imputed datasets. A single Markov chain Monte Carlo (MCMC) method was used with starting values based on the EM maximum likelihood estimates. We followed guidelines that were summarized by Allison (2002) and Enders (2006) to assess convergence by examining the degree to which the parameter estimates (means and covariances) changed between iterations. To do so, we evaluated the change in parameter estimates across subsequent iterations as displayed in a time series plot, and we used an autocorrelation plot to evaluate the degree to which the parameter estimates were correlated with those obtained from subsequent iterations. Convergence in our analyses was supported by the finding that the time series plot of the worst linear function (WLF) did not reveal any particular trend (see Figure 10). In addition, the autocorrelation plot of the worst linear function revealed that

[25] Allison (2002) provided further guidelines for converting the values of dummy-coded variables back to the original nominal variable based on the most likely classification. This, however, was not necessary because the imputed values for the nominal variables were not used in the subsequent analyses; they were only used for the purpose of increasing the accuracy of the imputed values for the variables that were included in our analyses (none of which was nominal).

correlations between iterations with a lag of 1 to 20 were all within sampling error (see Figure 11).[26]

Figure 10 Worst linear function time-series plot for multiple imputation

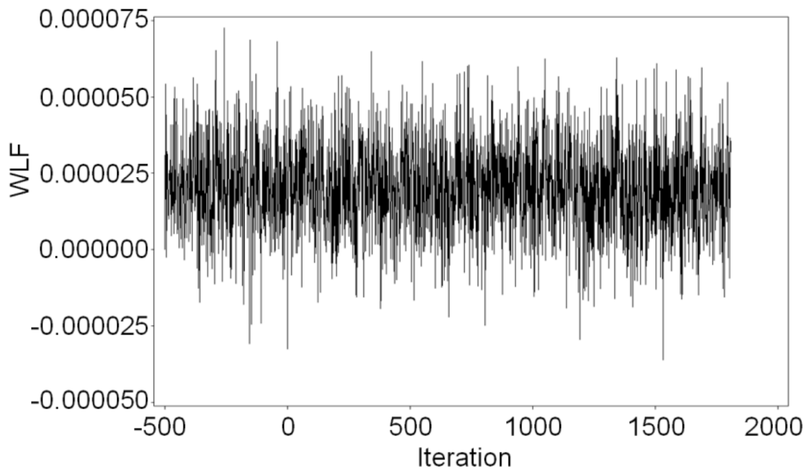
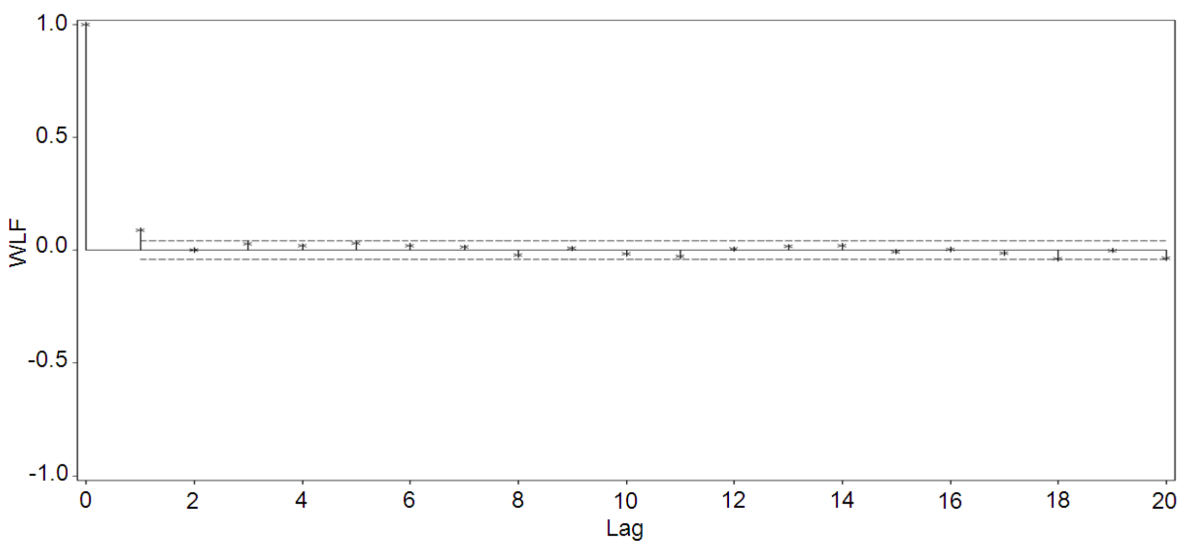


Figure 11 Auto-correlation plot with 95% confidence intervals for multiple imputation



[26] These results are based on 500 iterations prior to estimating values for the first imputed dataset (also referred as 500 burn-in iterations) and 200 iterations between each subsequently imputed dataset.

The parameter estimates for our analyses subsequently were averaged across the imputed datasets and their standard errors were obtained by adjusting for the corresponding within and between imputation variances as explained by Enders (2006). In this way, the amount of missing information was taken into account when estimating the standard errors (Allison, 2002). Model fit indices were similarly averaged across the imputed datasets. These techniques were implemented by using the multiple imputation feature in the Mplus 4.2 software package (B. Muthén & L. K. Muthén, 2006).

In summary, we used several approaches to account for the missing data in our analyses. The first approach was based on the premise of imputing a minimal amount of data for those who had a missing value for only one of the MSLSS items. This approach was necessary to obtain a single complete dataset for the factor mixture analyses. The second approach involved FIML, which allowed us to account for the ordinal nature of the variables. The third approach was based on multiple imputation techniques based on all available information. This technique allowed us to include auxiliary variables in the imputation process. The rationale was that consistency in the results across these three techniques would provide support for a conclusion that the results were not substantially influenced by the different assumptions underlying each of these approaches. A comparison of the findings derived from these missing data techniques is discussed in the results section.

2.6 Concluding comments

We have provided a detailed overview of the methods underlying our analyses so as to clearly describe how these methods were used to address the substantive purposes of the study. In addition to these substantive purposes, our methodological approaches were greatly influenced by several methodological objectives. Our first methodological objective was to model the ordinal categorical nature of the variables as closely as possible. Our preliminary analyses demonstrated that there was a great deal of variability with respect to the ordinal thresholds for each of the variables. This diversity in response patterns would have been ignored if the ordinal variables were simply assumed to be normally distributed in our analyses.

The second methodological objective was that we desired to specify a model that was least affected by the observed and potentially unobserved heterogeneity in the sample. For example, our selection of MSLSS items (in the abridged measurement structure) was partially guided by identifying those with parameters that were most invariant across latent classes. An alternative approach may have been to exclude those respondents that would have been considered statistical outliers. This approach would have resulted in the exclusion of a significant number of respondents given that the sample was found to be relatively heterogeneous in the confirmatory factor analyses.[27] Rather than excluding respondents in our analyses, we opted to adjust the model so as to best represent the diversity or heterogeneity within the sample. An obvious limitation is, of course, that these methods were relatively exploratory in nature.

A third methodological objective was to avoid unnecessarily excluding cases with data that may have been informative. This had direct implications for the missing data techniques that we used. These techniques were chosen to prevent needless bias by including,

[27] For example, if we were to calculate the Mahalanobis distance for the identification of multivariate outliers (as, for example, provide by Tabachnick and Fidell (2007)), we would have found that, of the subsample of 6,325 respondents that had a missing value for only one of the MSLSS items, 571 (6.9%) had a value greater than the critical value of 73.4 based on a chi-square distribution with 40 degrees of freedom ($p < 0.001$). These respondents would have possibly been excluded as outliers in our analyses. This finding is not surprising considering that the other published CFA results of the MSLSS were primarily completed with smaller and more homogeneous samples (e.g., samples that were limited to a particular school or region).

as much as possible, the information that was at our disposal. For example, rather than relying on methods that would emphasize the average response to a particular question (e.g., as in group mean imputation), without regard for the diversity that was represented in the sample, we used other methods to impute the most likely responses based on all potentially relevant information (i.e., as in multiple imputation).

Our missing data management techniques were driven not only by statistical considerations. We were concerned that the exclusion of a large amount of data would be questionable from an ethical perspective (e.g., as in listwise deletion). Considering the time and resources used to collect these data, it may have been deemed improper to have ignored a large percentage of respondents who missed only a few questions.

Thus, the statistical methods that we used were congruent with the methodological objectives underlying the study. The results of these methodological approaches, as discussed in the following sections, provide a detailed representation of the adolescents' responses to the survey questions and the hypothesized relationships among the variables based on the theoretical propositions of interest.

3 FINDINGS

The findings in this chapter are organized so as to specifically address the following three purposes: (a) to test the assumptions of the putative measurement structure of the Multidimensional Students' Life Satisfaction Scale (MSLSS) (Huebner, 1994, 2001) with the goal of assessing its reliability and validity with respect to the measurement of adolescents' satisfaction with their family, friends, living environment, school and self, and their general life satisfaction, (b) to determine the degree to which the dimensions of life satisfaction in the MSLSS explain global QOL, and (c) to examine whether perceived mental health status, perceived physical health status, or both contribute to global QOL, and whether the dimensions of life satisfaction mediate the relationship(s). We begin by providing a description of the sample and the distributions of the ordinal variables in our analyses. We then focus on the first purpose by presenting the findings resulting from the CFAs and the FMAs of the MSLSS. The SEM results of the spurious indicator models are presented next, and we conclude with an overview of the results pertaining to the role of life satisfaction dimensions in mediating the relationships between mental and physical health status and global QOL. We provide a comparison of findings resulting from different missing data techniques in relation to each of the analyses.

3.1 Sample description

The sample (N = 8,225) consisted of 4,064 boys (49.8%) and 4,099 girls (50.2%) (62 did not identify their gender) in grades 7 to 12 (see last column in Table 4 on page 61). The average age was 15.2 years (SD = 1.5, N = 8,054) with 7,964 adolescents being between the ages of 12 and 18 years. Of those adolescents who identified their ethnicity (N = 7,882), most described themselves as "white/Caucasian" (N = 5,721, 72.6%); the sample also included

1,301 (16.5%) Aboriginal adolescents, 461 (5.9%) Asian adolescents (Chinese, Japanese, Korean, Filipino or South-East Asian), and 399 (5.1%) adolescents that described themselves as belonging to one or more other ethnic groups. This ethnic diversity is further reflected in the finding that, of the 7,994 adolescents that provided information about the language(s) they spoke at home, a sizeable percentage indicated regularly speaking a language other than English (N = 1,384, 17.3%). In addition, of the 8,058 adolescents that indicated their country of birth, 557 (6.9%) reported being born in a country other than Canada.

The adolescents also were asked to respond to several questions about their living arrangements and their experiences at school. In terms of their living arrangements, among those that provided information about whom they lived with most of the time (N = 7,582), 59.9% (N = 4,542) reported living with their mother *and* father, 25.7% (N = 1,945) lived with their mother *and not* their father (9.2% lived with their mother and another person who was not their father, and 16.5% lived only with their mother), 7.8% (N = 590) lived with their father but not with their mother (2.6% lived with their father and another person who was not their mother, and 5.2% lived only with their father), and 6.7% (N = 505) did not live with their mother or father. Among the adolescents that answered these questions (N = 7,582), 3.4% (N = 261) reported living with a foster parent or guardian. These numbers reveal the diverse living arrangements of the adolescents in this sample.

The adolescents also differed with respect to their experiences at school. They were asked to rate their school performance on a 7-point scale (ranging from “far below average” to “well above average”) in response to the question, “Compared with other students in your school, how do you rate yourself in the school work you do?” Of the 7,060 adolescents that answered this question, 18.4% (N = 1,298) reported that their school performance was below average, 52.1% (N = 3,675) viewed their school performances as being above average, and 29.6% (N = 2,087) reported average school performance. The adolescents also were asked to evaluate the degree to which they felt like an outsider in their school environment. In response to this question, most of the adolescents (63.5%, N = 4,378) rarely or never felt like an outsider, 36.5% (1,715) felt like an outsider some of the time, and 11.6% (N = 1,101) saw themselves as an outsider most or all of the time (1,335 adolescents did not answer this question).

The above description reveals substantial diversity with respect to the age, ethnicity, living arrangements, and school experiences of the adolescents in this sample. These and other potential differences may have influenced how they interpreted and responded to some of the survey questions. We discuss the responses to the health status and global QOL questions next, and then continue with a detailed discussion of the measurement structure of the MSLSS (Huebner, 1994) while considering the possibility of heterogeneity in the adolescents' responses to the items.

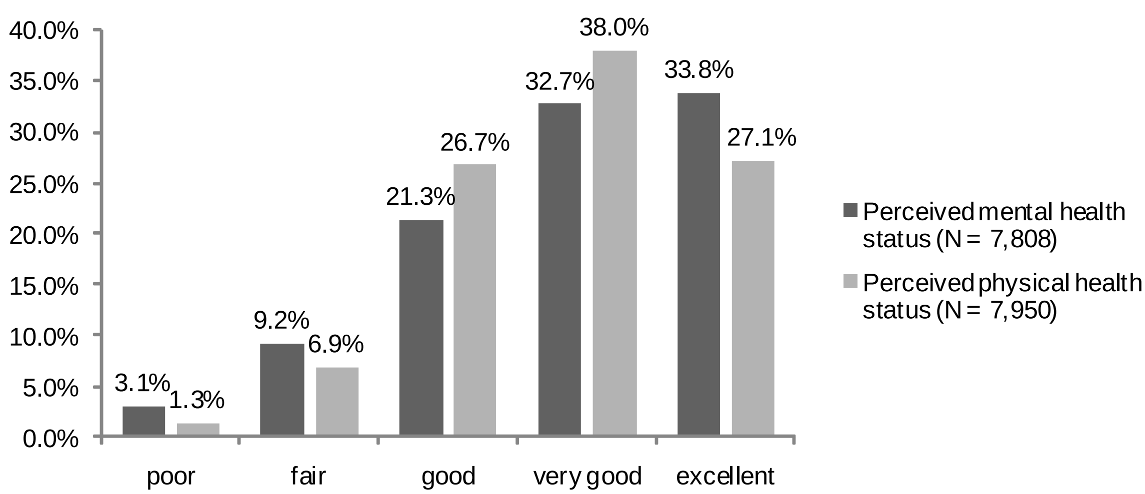
3.2 The distributions of the ordinal variables

A comparison of the adolescents' ratings of their perceived physical and mental health status is provided in Figure 12. The distributions of these variables revealed their ordinal and discrete nature.[28] Most of the adolescents reported very good (N = 3,018, 38.0%)

[28] Had we assumed that these variables were normally distributed, we would have found that the mean and standard deviations of both variables were very similar (for physical health $\mu = 3.83$, $\sigma = 0.95$; for mental health $\mu = 3.85$, $\sigma = 1.08$). However, the differences in the tails of the distributions (e.g., the difference in the percentages of poor or fair mental health compared with those of poor or fair physical health) would have been ignored if normal distributions had been assumed. Consequentially, it is not surprising that the polychoric correlation ($r = 0.55$) for these variables is somewhat larger than their Pearson correlation ($r = 0.48$) (which is based on the assumption of normally distributed variables).

or excellent (N = 2,156, 27.1%) physical health status; 2,124 (26.7%) of the adolescents rated their physical health as good, and relatively few rated their physical health as fair (N = 546, 6.9%) or poor (N = 106, 1.3%) (275 adolescents did not respond to this question). In comparison, a slightly larger percentage of the adolescents reported poor (N = 239, 3.1%) or fair mental health (N = 718, 9.2%), and most rated their mental health as good (N = 1,660, 21.3%), very good (N = 2,554, 32.7%), or excellent (N = 2,637, 33.8%) (417 adolescents did not respond to this question).

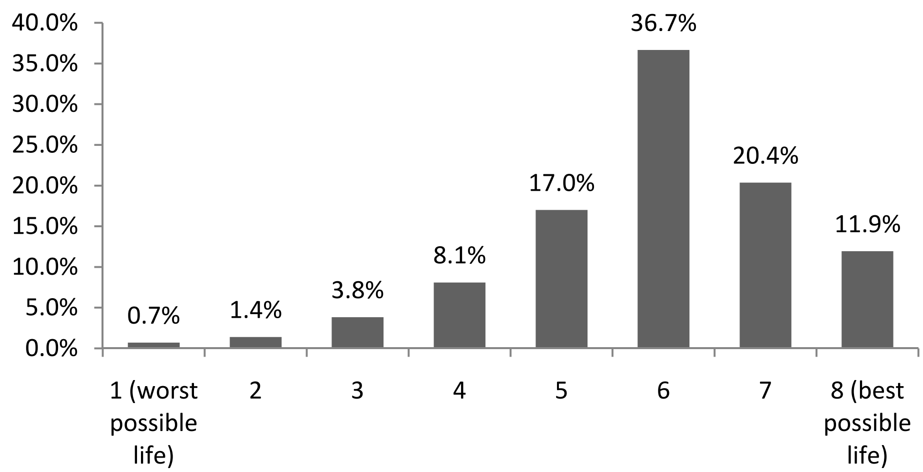
Figure 12 Responses to the perceived mental and physical health status questions



With respect to their global QOL, most of the adolescents agreed (N = 4,011, 52.7%) or strongly agreed (N = 2,252, 29.6%) that they were satisfied with their quality of life. Nevertheless, a sizeable percentage of the adolescents disagreed (N = 990, 13.0%) and some strongly disagreed (N = 353, 4.6%) with the statement (619 adolescents did not respond to this question). A similarly negatively skewed distribution was found for the QOL-ladder variable, which was rated on a scale of 1 (worst possible life) to 8 (best possible life) (see

Figure 13). Evidently most adolescents were more satisfied than dissatisfied with their quality of life.

Figure 13 Responses to the QOL-ladder



The percentages for each of the response options of the MSLSS (Huebner, 2001) items are presented in Table 6 (see Figure 8 on page 65 for their corresponding distributions). After reverse scoring the negatively worded items, the modes of all the items were at or above the response option corresponding to mild agreement. This suggests that most of the adolescents were satisfied rather than dissatisfied with respect to the MSLSS questions (the modal response was mildly agree for 7 items, agree for 21 items, and strongly agree for 12 items). However, the distributions were not uniform. In particular, there appeared to be several discrepancies in how the adolescents responded to the negatively worded items in comparison with the positively worded items. For example, 32.1% of the adolescents strongly agreed that they wished there were different people in their neighborhoods. This response seemed to contradict their responses to the other living environment items, which suggested that they were satisfied with their living environment. Similarly, 22.9% of the

Table 6 Item response percentages for MSLSS indicators

Variable	Percentage of item responses					
Family subscale	strongly disagree	disagree	mildly disagree	mildly agree	agree	strongly agree
fam1 (N=7,134)	5.5%	6.8%	10.7%	26.6%	32.4%	18.1%
fam2 (N=7,049)	4.4%	7.3%	12.7%	27.9%	29.4%	18.2%
fam3 (N=7,071)	4.8%	6.1%	10.8%	26.0%	32.7%	19.7%
fam4 (N=7,048)	5.2%	7.1%	11.9%	21.6%	34.3%	20.0%
fam5 (N=7,035)	3.3%	4.0%	7.5%	17.9%	39.7%	27.7%
fam6 (N=6,981)	4.9%	6.9%	11.7%	22.3%	34.5%	19.7%
fam7 (N=6,969)	5.5%	7.7%	11.4%	26.5%	29.0%	20.0%
Friends subscale						
frnd1 (N=7,269)	1.4%	1.0%	2.1%	9.1%	43.1%	43.2%
frnd2 (N=7,096)*	47.1%	30.8%	9.0%	6.2%	4.2%	2.7%
frnd3 (N=7,137)	1.3%	1.0%	2.9%	11.8%	35.6%	47.2%
frnd4 (N=7,119)	1.8%	1.6%	3.7%	13.9%	35.9%	43.1%
frnd5 (N=7,084)	1.4%	1.2%	3.6%	13.9%	42.9%	37.0%
frnd6 (N=7,021)*	50.5%	29.2%	9.0%	6.1%	2.9%	2.4%
frnd7 (N=6,987)*	54.3%	21.8%	9.2%	7.5%	4.0%	3.3%
frnd8 (N=6,991)	1.1%	1.5%	3.4%	11.2%	36.6%	46.2%
frnd9 (N=6,952)	3.3%	4.2%	7.8%	16.0%	38.4%	30.4%
Living environment subscale						
lenv1 (N=7,088)	12.2%	13.5%	14.9%	22.0%	21.1%	16.4%
lenv2 (N=6,991)*	34.8%	19.5%	12.1%	12.7%	10.7%	10.1%
lenv3 (N=6,965)	6.7%	6.7%	10.9%	21.6%	32.9%	21.2%
lenv4 (N=6,967)*	28.9%	18.6%	12.1%	15.3%	12.6%	12.5%
lenv5 (N=6,955)*	18.0%	27.8%	22.0%	17.9%	8.2%	6.2%

Living environment subscale (continued)	strongly disagree	disagree	mildly disagree	mildly agree	agree	strongly agree
lenv6 (N=6,972)	2.5%	3.4%	6.1%	18.5%	39.2%	30.4%
lenv7 (N=6,923)	6.6%	5.8%	10.7%	23.4%	33.5%	20.0%
lenv8 (N=6,993)	5.5%	4.6%	7.7%	18.3%	31.9%	32.1%
lenv9 (N=6,916) *	14.0%	16.6%	15.7%	22.4%	18.1%	13.2%
School subscale						
schl1 (N=7,179)*	25.4%	32.2%	15.5%	16.4%	6.5%	4.0%
schl2 (N=7,150)	4.7%	6.1%	9.7%	28.2%	35.6%	15.8%
schl3 (N=7,135)*	4.8%	8.8%	13.5%	25.2%	27.1%	20.7%
schl4 (N=7,110)*	11.1%	13.7%	14.2%	21.8%	16.3%	22.9%
schl5 (N=7,050)	12.4%	11.7%	19.4%	30.6%	17.4%	8.7%
schl6 (N=7,007)	13.2%	10.5%	16.5%	30.9%	20.7%	8.2%
schl7 (N=7,032)	13.2%	10.6%	18.2%	31.2%	19.8%	7.1%
schl8 (N=7,006)	10.3%	9.7%	16.2%	27.8%	23.9%	12.1%
Self subscale						
self1 (N=7,241)	1.3%	1.6%	3.2%	16.2%	48.0%	29.7%
self2 (N=7,172)	2.0%	2.8%	5.3%	18.8%	41.7%	29.5%
self3 (N=7,056)	6.1%	8.1%	12.7%	30.4%	27.8%	15.0%
self4 (N=7,074)	3.0%	3.3%	6.0%	17.9%	40.2%	29.6%
self5 (N=7,055)	2.0%	2.8%	6.7%	22.6%	46.2%	19.7%
self6 (N=6,959)	1.9%	1.8%	4.0%	15.4%	45.4%	31.6%
self7 (N=6,981)	1.6%	2.1%	5.0%	19.5%	40.2%	31.8%

Notes: Percentages based on original responses. The shaded fields indicate the mode for the corresponding item.

* Negatively worded items.

adolescents strongly agreed with the statement that they wished they did not have to go to school, whereas only 13.2% strongly disagreed with the statement that they liked being at school. These discrepancies, albeit understandable (adolescents would be expected to strongly agree with a statement about not having to go to school), suggested that the adolescents may not have been responding to some of the negatively worded items in a manner congruent with their responses to other items in the subscale.

On the whole, the discrete nature of the data is evident in the item response distributions. Most of the adolescents in this sample viewed themselves as being healthy and satisfied with their family, friends, living environment, school and self as well as with their global QOL. However, some adolescents also provided very different ratings for some or all of these variables.

3.3 Examining the measurement structure of the MSLSS

Our examination of the measurement structure of the MSLSS was based on the following CFAs and FMAs: (a) CFAs to test the original indirect reflective factor structure and various plausible modifications thereof, (b) CFAs to test the implied first-order factor structures for each of the subscales, (c) FMAs to determine the variance of the item responses with respect to the measurement structure of each subscale, and (d) CFAs to test the first-order and second-order factor structures of an abridged version of the MSLSS.[29]

[29] For purposes of consistency, the results of the analyses using the dataset with only one imputed MSLSS value (*N* = 6,325) are presented in the text. The results of the analyses based on the other missing data techniques are presented in corresponding tables.

3.3.1 Confirmatory factor analyses

The CFA results of the multidimensional measurement structures for the MSLSS are presented in Table 7. Neither the indirect reflective measurement structure (see Figure 4 on page 38) nor the measurement structure with five correlated latent factors was supported in our sample. The fit indices for both models did not fall within the recommended ranges and the residual correlations were very large (ranging from -0.23 to 0.35 for the less restrictive correlated five-factor model) with 107 (13.7%) of the 780 residual correlations for the correlated five-factor model having absolute values greater than 0.10.

We used principal components analysis of the residual correlations (resulting from the correlated five factor structure) to identify whether they could be attributed to one or more unspecified dimensions (Zumbo, 2002). The results indicated that only 6.6% of the total residual variance could be attributed to the first component, which had an eigenvalue of 2.60. We concluded that the residual correlations would not be substantially reduced (or model fit would not be substantially improved) by adding another factor or factors to the model. It is therefore not surprising that, although the model with a method factor led to a substantial improvement in the model's fit ($\Delta \chi^2 = 2{,}306.39$, $\Delta df = 5$, $p < 0.001$), this model still did not fit the data very well (WLSMV $\chi^2 = 1{,}4752.07$, $df = 239$, $p < 0.001$, RMSEA = 0.098, CFI = 0.759). The residual correlations for this model ranged from -0.19 to 0.25 with 8.6% of the residual correlations having absolute values larger than 0.10 (see third model in Table 7).

Further examination of the pattern of residual correlations for the correlated five-factor model revealed that 93 (87%) of the 107 residual correlations with absolute values larger than 0.10 were associated with the negatively worded items, many of the covariate

Table 7 CFA of multidimensional measurement structures for the MSLSS

Missing data technique	*N*	WSLMV χ^2	*df*	CFI	RMSEA	SRMR	Residual correlations	
							range	% > |0.1|
2nd order factor model with five 1st order factors								
Single EM imputation[1]	6,325	17,500.98	195* 735**	0.713	0.118	0.077	-0.24 to 0.32	18.5%
Multiple imputation[2]	7,305	20,614.76 20,808.48	n/a	0.700	0.118	0.078	-0.26 to 0.32	19.0%
Correlated five-factor model								
Single EM imputation[1]	6,325	17,336.59	215* 730**	0.716	0.112	0.070	-0.23 to 0.35	13.7%
Multiple imputation[2]	7,305	20,192.62 20,429.31	n/a	0.706	0.112	0.071	-0.22 to 0.35	14.0%
Correlated five-factor model with method factor								
Single EM imputation[1]	6,325	14,752.07	239* 720**	0.759	0.098	0.060	-0.19 to 0.25	8.6%
Multiple imputation[2]	7,305	16,816.32 16,995.02	n/a	0.757	0.096	0.06	-0.19 to 0.26	9.0%
Correlated five-factor model with 30 items (excluding negatively worded items)								
Single EM imputation[1]	6,325	8,977.99	167* 395**	0.813	0.091	0.049	-0.15 to 0.20	4.1%
Multiple imputation[2]	7,305	10,272.69 10,437.62	n/a	0.803	0.090	0.048	-0.15 to 0.21	3.7%

Notes: χ^2 and fit indices are based on mean and variance adjusted weighted least squares (WLSMV). RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index;

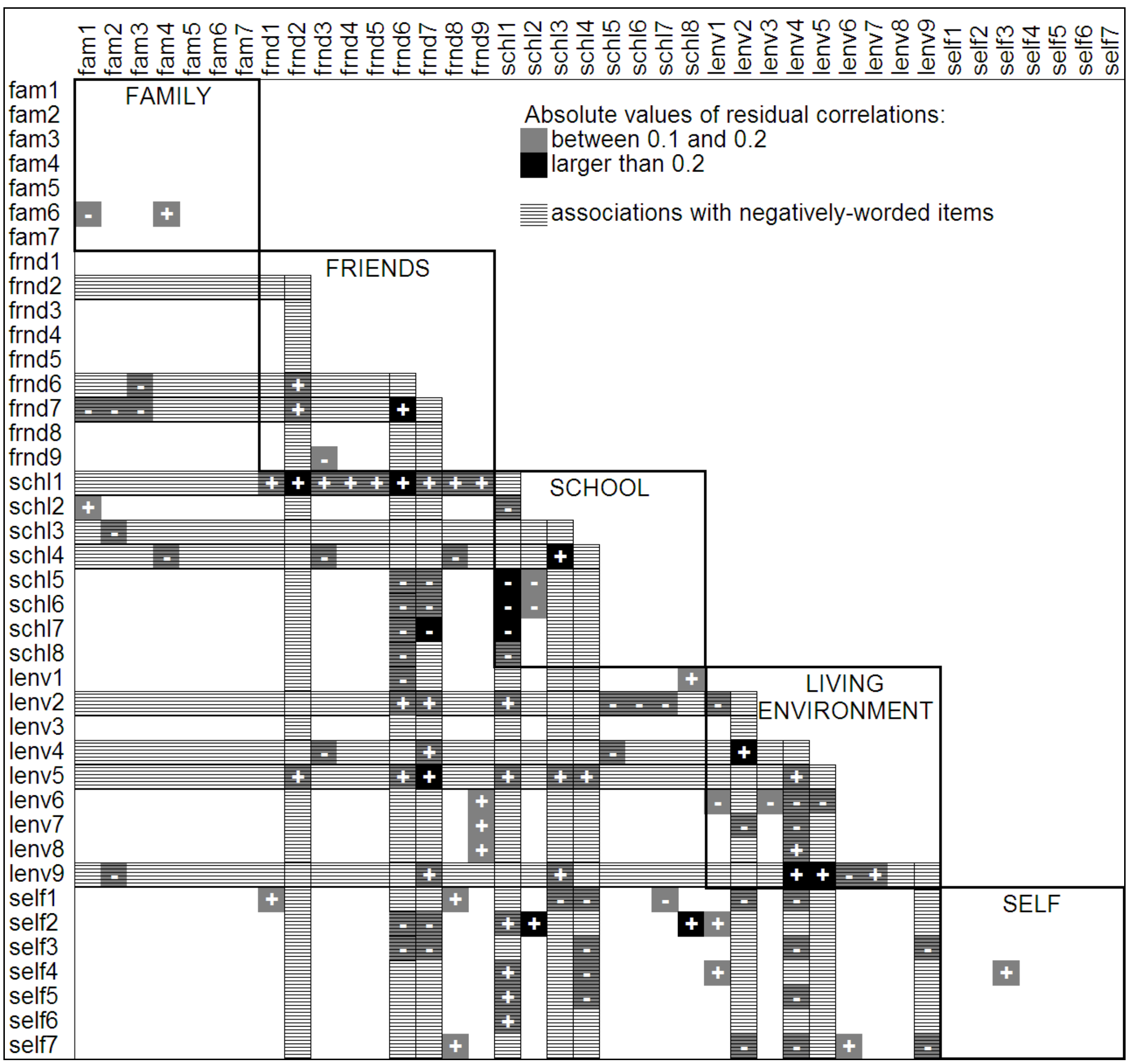
* *df* based on WLSMV estimation. ** *df* based on number of free parameters.

[1] Single imputation for those with one missing value.

[2] Analyses of multiple imputed datasets for those who completed at least one MSLSS item.

pairs being between negatively worded and positively worded items (see Figure 14). We therefore concluded that the observed model misfit was largely attributable to the negatively worded items. The adolescents likely did not respond to this type of items in a consistent manner. This was congruent with our earlier observations about the inconsistency in item response patterns for some of the negatively worded items in comparison with their corresponding positively worded items.

Figure 14 Residual correlations for the correlated five-factor model

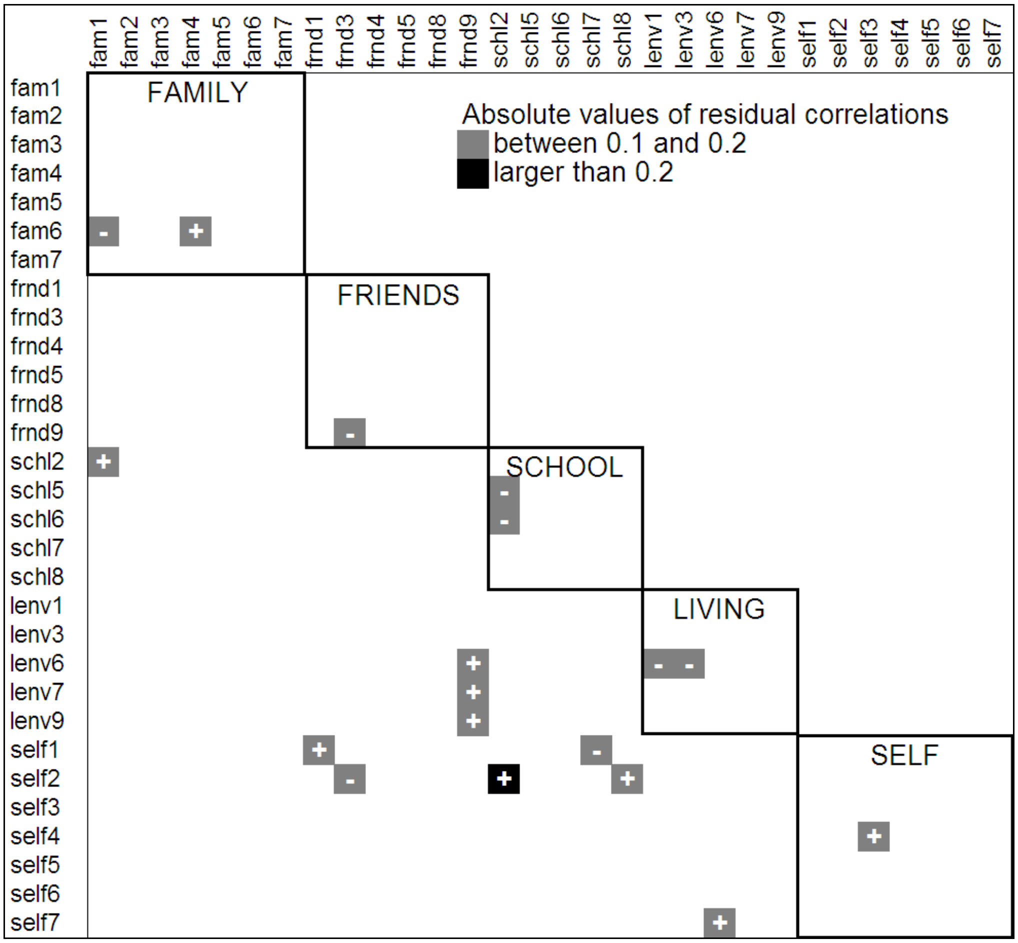


Notes: Residual correlations = observed polychoric correlations minus the estimated correlations of the correlated 5-factor model (*N* = 6,325).

Our inspections of the pattern of residual correlations and of the modification indices, however, did not reveal a simple structural modification of the measurement model by which the residual correlations could be accounted for. We hence continued our analyses by removing the negatively worded items altogether. This resulted in a better fitting model (WLSMV χ^2 = 8,977.99, *df* = 167, *p* < 0.001, RMSEA = 0.091, CFI = 0.813) (see last model in Table 7). The magnitude of the residual correlations was substantially reduced as indicated by: (a) a smaller SRMR value of 0.049, (b) a smaller range of residual correlations (from -0.15 to 0.20), and (c) a smaller proportion of residuals with absolute values greater than 0.10 (only 18 (4.14%) of 435). Although the magnitude of the remaining residual correlations and the large chi-square statistic suggested that problems with model fit persisted, the map of the residual correlations did not reveal any particular item or subscale as the predominant source of model misfit (see Figure 15).

Similar results were obtained for all of the above analyses when they were repeated with multiple imputation and full information maximum likelihood. The FIML BIC values provided additional information pertaining to the comparison of the CFA results of the reduced MSLSS (excluding the negatively worded items) with those of all the MSLSS items because the BIC can be used to compare the log likelihood across non-nested model. The log likelihood was substantially smaller for the CFA excluding the negatively worded items than for the CFA involving all MSLSS items (see Table 8).

Figure 15   Residual correlations excluding negatively worded items



Notes: Residual correlations = observed polychoric correlations minus the estimated correlations of the correlated 5-factor model excluding the negatively worded items (*N* = 6,325).

Table 8 Full information maximum likelihood CFAs of the MSLSS

Model	Log likelihood	Number of parameters	BIC
2nd order factor model with five 1st order factors	-37,1696.15	245	74,5571.91
Correlated five-factor model	-37,3137.94	250	74,8499.95
Correlated five-factor model with method factor	-37,1324.43	260	74,4961.91
Correlated five-factor model with 30 items	-26,9451.28	190	54,0592.85

Notes: BIC = Bayesian information criterion. *N* = 7,305.

3.3.1.1 Confirmatory factor analyses of subscales

We continued our analyses by examining each of the subscales of the MSLSS independently so as to specifically test assumption three pertaining to the unidimensionality among the observed ordinal item responses for each subscale. As shown in Table 9, none of the measurement structures for the subscales fit very well (i.e., assumptions two or three were not justified). However, model fit was better for the friends, living environment and school subscales when the negatively worded items were excluded, as indicated by improved values for each of the fit indices and smaller residual correlations. Nevertheless, the large chi-square statistics and fit indices (in particular the RMSEA) suggested that the unidimensional structures for each of the subscales still did not fit the data very well.

All the subscales also were evaluated using FIML so that we could more formally compare the non-nested models based on the inclusion of the negatively worded items (see Table 10). The results of these analyses confirmed that the fit improved substantially (as indicated by smaller BIC values) when the negatively worded items were removed.

Table 9 CFA models of the MSLSS subscales

Model	WLSMV χ^2	*df*	CFI	RMSEA	SRMR	Residual correlations	
						range	% > |0.01|
Family	3,020.71	11*	0.908	0.208	0.050	-0.16 to 0.10	19.1%
Self	976.17	13*	0.940	0.108	0.040	-0.09 to 0.08	0.0%
Friends (all items)†	3,334.86	18*	0.902	0.171	0.055	-0.10 to 0.19	13.0%
Friends (no neg.) ‡	643.48	8*	0.979	0.112	0.026	-0.05 to 0.07	0.0%
Living env. (all items) †	3,123.06	17*	0.865	0.170	0.077	-0.18 to 0.16	53.3%
Living env. (no neg.) ‡	412.50	5*	0.971	0.114	0.031	-0.06 to 0.04	0.0%
School (all items)†	2,504.89	15*	0.925	0.162	0.052	-0.07 to 0.18	3.5%
School (no neg.) ‡	1,270. 96	4*	0.958	0.224	0.042	-0.06 to 0.07	0.0%

Notes: Analyses based on EM imputations for those with one missing MSLSS values (*N* = 6,325).

* *df* based on WLSMV estimation.

† CFA based on all subscale items.

‡ CFA excluding negatively worded items.

Table 10 Full information maximum likelihood CFAs of MSLSS subscales

Subscale	# items	*N*	# par	Log likelihood	BIC
Family (all items)	7	7,240	42	-66,252.38	132,878.02
Self (all items)	7	7,291	42	-61,995.78	124,365.12
Friends (all items) †	9	7,300	54	-68,271.67	137,023.70
Friends (no neg.) ‡	6	7,300	36	-43,850.52	88,021.27
Living env (all items) †	9	7,175	54	-95,425.74	191,330.91
Living env (no neg.) ‡	6	7,174	30	-50,787.12	101,840.59
School (all items) †	8	7,281	48	-84,632.88	169,692.62
School (no neg.) ‡	5	7,246	30	-51,074.43	102,415.51

† CFA based on all subscale items.

‡ CFA excluding negatively worded items.

3.3.2 Factor mixture analyses of subscales

The next step in our analyses was to examine potential sample heterogeneity with respect to the unidimensional measurement structure for each of the subscales (excluding the negatively worded items) so as to specifically test assumptions 1 and 2. We specified factor mixture models for each subscale while allowing the item thresholds and factor loadings to vary across two or more latent classes (Figure 9 on page 73). By comparing these parameters across latent classes we were able to identify those items with the least consistent (and therefore potentially unreliable) response patterns. We discuss the FMA results for each of the subscales separately starting with the Family subscale.

3.3.2.1 Family subscale

The smallest BIC value for the family subscale was obtained for a four-class model (see Table 11), which suggested that the sample was heterogeneous with respect to the

Table 11 FMA of the family subscale

K	*P*	LL	Δ in LL[1]	BIC	Entropy	Class proportions[2]				
						C1	C2	C3	C4	C5
1	42	-58,211.65	0.000	116,790.89	1.000	1.00				
2	85	-56,863.62	0.003	114,471.18	0.579	0.40	0.60			
3	128	-56,258.82	0.000	113,637.94	0.613	0.28	0.49	0.29		
4	171	-55,899.78	0.004[3]	113,296.20	0.607	0.13	0.36	0.23	0.29	
5	214	-55,779.94	0.019	113,432.86	0.603	0.11	0.19	0.26	0.31	0.14

Notes: N = 6,325. LL = log likelihood. BIC = Bayesian information criterion.

[1] The change in LL compared with the next best LL obtained from the random starts. A value of zero indicates that the maximum likelihood was exactly replicated.

[2] Predicted class proportions based on the largest posterior latent class probability.

[3] Comparison of this solution to the solution with the next best LL revealed that, except for one threshold parameter corresponding to a very small cell size, the parameter estimates were very similar for both models and there was 99.1% congruence in predicted class membership based on the posterior probabilities (Cohen's kappa = 0.99, 95% CI: 0.984 to 0.990).

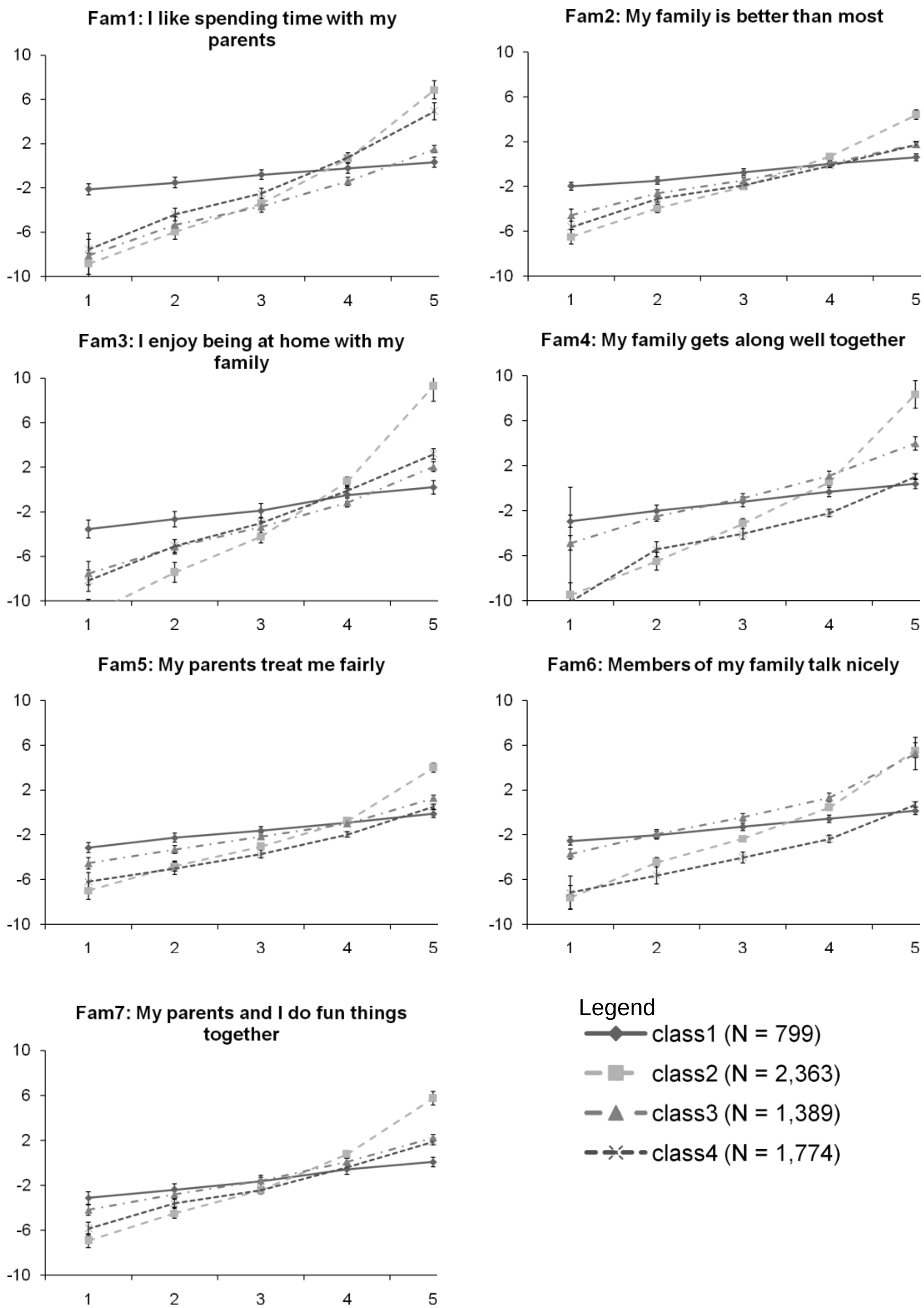
measurement structure. Although the log likelihood (LL) for the four-class FMA model was not exactly replicated (Δ in LL = 0.004), a comparison of this solution with the solution with the next best LL revealed that, except for one threshold parameter corresponding to a very small cell size, the parameter estimates were very similar for both models and there was 99.1% congruence in predicted class membership based on the posterior probabilities

(Cohen's kappa = 0.99, 95% CI: 0.98 - 0.99). Thus we were confident that a sufficient approximation of the global maximum of the log likelihood was obtained.

Inspection of the thresholds for each of the items in the four-class factor-mixture model, as shown in Figure 16, suggested that the thresholds for Fam1, Fam3, and Fam4 were least invariant across the four latent classes, whereas the thresholds for Fam2, Fam5, Fam6, and Fam7 were comparatively more invariant. Based on these findings, we specified a new single-factor measurement model for the family subscale based on the four items with the most invariant thresholds (i.e., Fam2, Fam5, Fam6, and Fam7). This resulted in remarkable improvement in model fit (WLSMV $\chi^2_{(2)} = 5.126$, RMSEA = 0.016, SRMR = 0.004, CFI = 1.00). The α-coefficient of internal consistency (based on polychoric correlations) for these four items was 0.83, and the standardized factor loadings ranged from 0.66 to 0.81. Subsequent CFAs of all of the 35 possible combinations of four variables confirmed that this particular combination of variables resulted in the best fitting model (see Table 12).[30]

[30] Although one might argue that it is not surprising that a model with only two degrees of freedom would result in a better fitting model, it is noteworthy that the chi-square values of the CFA for all 35 combinations of four variables ranged between 1,355.993 and the minimum value of 5.126. Thus it cannot be concluded that the improvement in model fit was a mere artifact of the reduction in degrees of freedom. A review of the best fitting models in Table 12 also clearly illustrated that the fit indices did not provide sufficient information to determine which combination of four variables was most congruent with assumptions one to three. For example, three of the four best fitting models included variables with thresholds that were clearly not consistent across the various latent classes.

Figure 16 Thresholds of family subscale variables in four latent classes



Notes: *N* = 6,325. Threshold values are scaled on the Y-axis as estimated cumulative log odds, threshold numbers are on the X-axis. The bars in each graph represent 95% confidence intervals.

Table 12 CFA results of the family subscale

Included variables				WLSMV χ^2	RMSEA	SRMR	CFI
Fam1	Fam2	Fam3	Fam4	488.083	0.196	0.031	0.985
Fam1	Fam2	Fam3	Fam5	101.021	0.088	0.015	0.997
Fam1	Fam2	Fam3	Fam6	281.239	0.149	0.028	0.989
Fam1	Fam2	Fam3	Fam7	33.517	0.050	0.007	0.999
Fam1	Fam2	Fam4	Fam5	269.487	0.145	0.025	0.983
Fam1	Fam2	Fam4	Fam6	1080.261	0.292	0.053	0.945
Fam1	Fam2	Fam4	Fam7	314.379	0.157	0.027	0.984
Fam1	Fam2	Fam5	Fam6	355.428	0.167	0.032	0.970
Fam1	Fam2	Fam5	Fam7	120.282	0.097	0.017	0.993
Fam1	Fam2	Fam6	Fam7	379.006	0.173	0.033	0.977
Fam1	Fam3	Fam4	Fam5	779.862	0.248	0.037	0.965
Fam1	Fam3	Fam4	Fam6	1260.782	0.446	0.079	0.965
Fam1	Fam3	Fam4	Fam7	388.938	0.175	0.024	0.986
Fam1	Fam3	Fam5	Fam6	850.681	0.259	0.045	0.968
Fam1	Fam3	Fam5	Fam7	231.205	0.135	0.018	0.991
Fam1	Fam3	Fam6	Fam7	679.012	0.231	0.036	0.981
Fam1	Fam4	Fam5	Fam6	531.514	0.205	0.036	0.975
Fam1	Fam4	Fam5	Fam7	250.716	0.140	0.022	0.989
Fam1	Fam4	Fam6	Fam7	1355.993	0.327	0.055	0.947
Fam1	Fam5	Fam6	Fam7	326.104	0.160	0.029	0.982
Fam2	Fam3	Fam4	Fam5	20.370	0.038	0.006	0.999
Fam2	Fam3	Fam4	Fam6	494.309	0.197	0.032	0.98
Fam2	Fam3	Fam4	Fam7	87.821	0.082	0.013	0.997
Fam2	Fam3	Fam5	Fam6	113.449	0.094	0.016	0.992
Fam2	Fam3	Fam5	Fam7	49.247	0.061	0.010	0.998
Fam2	Fam3	Fam6	Fam7	117.512	0.096	0.016	0.994
Fam2	Fam4	Fam5	Fam6	134.853	0.102	0.019	0.994
Fam2	Fam4	Fam5	Fam7	29.296	0.046	0.008	0.998
Fam2	Fam4	Fam6	Fam7	213.612	0.129	0.022	0.990
Fam2	Fam5	Fam6	Fam7	5.126	0.016	0.004	1.000
Fam3	Fam4	Fam5	Fam6	364.322	0.169	0.025	0.987
Fam3	Fam4	Fam5	Fam7	119.279	0.096	0.014	0.996
Fam3	Fam4	Fam6	Fam7	1035.349	0.286	0.041	0.968
Fam3	Fam5	Fam6	Fam7	160.863	0.112	0.017	0.993
Fam4	Fam5	Fam6	Fam7	286.420	0.150	0.022	0.989

Notes: Analyses based on all possible combinations of four variables for the family subscale. Rows for the four best fitting models are shaded. *N* = 6,325.

3.3.2.2 Self subscale

The smallest BIC value for the self subscale was obtained for a three-class model (see Table 13). However, the BIC values for the three and four class models were fairly close and the LL for both models was not exactly replicated. The posterior probabilities of the four-class model indicated that one latent class represented only a small proportion of the sample (9.4%). In addition, the parameter estimates for the four-class model did not replicate as well as for the three-class model. We therefore chose the three-class model as the most substantively meaningful and parsimonious model with the best model fit for our analyses.

Table 13 FMA of the self subscale

K	*P*	LL	Δ in LL[1]	BIC	Entropy	Class proportions[2]				
						C1	C2	C3	C4	C5
1	42	-54,173.18	0.000	108,713.96	1.000	1.00				
2	85	-53,019.88	0.064	106,783.69	0.579	0.31	0.69			
3	128	-52,758.93	0.012	106,638.14	0.517	0.41	0.46	0.14		
4	171	-52,600.71	0.011	106,698.07	0.520	0.18	0.52	0.10	0.20	

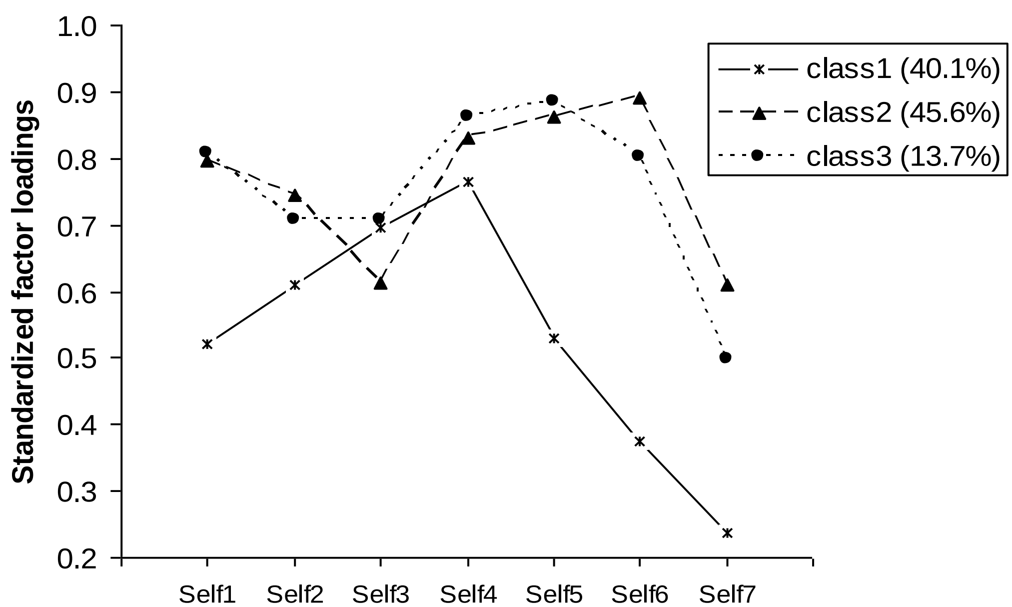
Notes: *N* = 6,325. LL = log likelihood. BIC = Bayesian information criterion.

[1] The change in LL compared with the next best LL obtained from the random starts.

[2] Predicted class proportions based on the largest posterior latent class probability.

Based on our comparison of the thresholds for each item of the self subscale, we concluded that the thresholds seemed least consistent for Self5 and Self6, and the most consistent thresholds were observed for Self3 and Self7. However, several other variables also appeared inconsistent and it was difficult to determine which combination of variables would be the most preferable. We subsequently examined the standardized factor loadings and found that these were least consistent across latent classes for variable Self6 (λ_{class1} = 0.24, λ_{class2} = 0.61, λ_{class3} = 0.50) (see Figure 17). The factor loadings were most consistent for Self2, Self3 and Self4. Although it seemed defensible to include these three variables and to exclude Self6 from a common factor model, it was unclear whether Self5 (with inconsistent thresholds) or Self7 (with inconsistent factor loadings) should have been included. Subsequent CFAs revealed that including variable Self5 led to substantial improvement in model fit (WLSMV $\chi^2_{(2)}$ = 16.967, RMSEA = 0.034, SRMR = 0.008, CFI = 0.999), whereas model fit was not adequate when Self7 was included (WLSMV $\chi^2_{(2)}$ = 122.276, RMSEA = 0.098, SRMR = 0.027, CFI = 0.987). We therefore chose the model that included variables Self2, Self3, Self4 and Self5 as the best fitting combination of items with the most consistent model parameters. The α-coefficient of internal consistency for the remaining four items was 0.81 and the standardized factor loadings ranged from 0.68 to 0.84.

Figure 17   Class-specific standardized factor loadings for the self subscale



Notes: *N* = 6,325.

School subscale

The FMA results also indicated that the sample was not homogeneous with respect to the measurement structure of the school subscale. However, we were unable to confidently identify the best model based on the BIC because the BIC continued to become smaller when more latent classes were specified (see Table 14). Nevertheless, we observed that the parameter estimates for the five class model were very unstable and our comparison of predicted latent class membership for the five class model with the best LL with that obtained with the next best LL revealed that the solutions were very different (Cohen's Kappa was 0.39 (95% CI: 0.37 to 0.40)). Clearly, the five class model did not fit the data well. On the other hand, the parameter estimates for the four class model were very similar in the two solutions with the best LLs (Cohen's Kappa for predicted latent class membership

Table 14 FMA of the school subscale

K	*P*	LL	Δ in LL[1]	BIC	Entropy	Class proportions[2]				
						C1	C2	C3	C4	C5
1	30	-44,727.06	0.001	89,716.70	1.000	1.00				
2	61	-43,333.15	0.003	87,200.19	0.696	0.29	0.71			
3	92	-42,996.10	0.008	86,799.20	0.505	0.32	0.46	0.23		
4	123	-42,772.53	0.590	86,621.59	0.568	0.36	0.30	0.19	0.15	
5	154	-42,670.15	4.206	86,688.14	0.583	0.13	0.21	0.28	0.12	0.26

Notes: *N* = 6,325. LL = log likelihood. BIC = Baysian information criterion.

[1] The change in LL compared with the next best LL obtained from the random starts.

[2] Predicted class proportions based on the largest posterior latent class probability.

was 1.00 (95% CI: 0.99 to 1.00)). We therefore chose the four class model as the best fitting model.

We compared the thresholds of the five items and found that the thresholds for item Schl6, “I like being at school,” were very different for each of the latent classes. Evidently the adolescents in this sample did not respond to this item in a consistent manner. Not surprisingly, model fit improved substantially when this item was excluded (WLSMV $\chi^2_{(2)}$ = 8.729, RMSEA = 0.023, SRMR = 0.004, CFI = 1.00). The α-coefficient of internal consistency for the remaining four items was 0.84 and the standardized factor loadings ranged from 0.69 to 0.89.

3.3.2.3 Friends subscale

The lowest BIC for the friends subscale was obtained with a two-class model (see Table 15). However, the LL of the three class model was very close in magnitude. The solution for the three class model with the best log likelihood was similar to the solution with the next best LL with 98% congruency in latent class membership and a Cohen’s Kappa of 0.97 (95% CI: 0.96 - 0.98). Consequently, a three-class model seemed to be justified. Both models were therefore used for our purpose of identifying items that were inconsistent with respect to the measurement model.

When examining the parameters of the measurement model within each of the latent classes, we found that the thresholds for Frnd9 seemed most consistent. The thresholds for Frnd4 and Frnd8 seemed more consistent than the thresholds of Frnd1, Frnd3 and Frnd5. However, we could not clearly determine which of the items had the least consistent thresholds and factor loadings. We examined the fit of various combinations of items and found that the model with items Frnd1, Frnd4, Frnd8 and Frnd9 led to the best model fit

Table 15 FMA of the friends subscale

K	*P*	LL	Δ in LL[1]	BIC	Entropy	Class proportions[2]				
						C1	C2	C3	C4	C5
1	36	-37,575.70	0.000	75,466.48	1.000	1.00				
2	73	-36,632.35	0.000	73,903.61	0.486	0.61	0.39			
3	110	-36,473.64	0.212	73,910.03	0.494	0.18	0.26	0.56		
4	147	-36,362.36	15.040	74,011.30	0.518	0.21	0.06	0.21	0.52	
5	184	-36,344.63	4.828	74,299.68	0.452	0.36	0.09	0.18	0.19	0.19

Notes: *N* = 6,325. LL = log likelihood. BIC = Bayesian information criterion.

[1] The change in LL compared with the next best LL obtained from the random starts.

[2] Predicted class proportions based on the largest posterior latent class probability.

(WLSMV $\chi^2_{(2)}$ =45.066, RMSEA = 0.058, SRMR = 0.013, CFI = 0.996). The α-coefficient of internal consistency for this combination of items was 0.82 and the standardized factor loadings ranged from 0.56 to 0.82.

3.3.2.4 Living environment subscale

For the living environment subscale we found that the lowest BIC value was obtained when three latent classes were specified (see Table 16). This subscale consisted of only five items, so our objective was to remove the item with the least consistent model parameters across the three latent classes. With respect to the item-thresholds, we observed that these were most homogeneously distributed for Lenv1, whereas the item thresholds of the remaining four items were inconsistently distributed across the three latent classes. The standardized factor loadings of Lenv7, “I like my neighbors,” were least consistent ranging from 0.37 in class 1 to 0.93 in class 3. Not surprisingly, a CFA of the unidimensional factor structure in the overall sample resulted in better model fit when this item was excluded (WLSMV $\chi^2_{(2)}$ = 28.191, RMSEA = 0.046, SRMR = 0.011, CFI = 0.998), and the resulting α-coefficient of internal consistency was 0.79.

Table 16 FMA of the living environment subscale

K	*P*	LL	Δ in LL[1]	BIC	Entropy	Class proportions[2]				
						C1	C2	C3	C4	C5
1	30	-45,283.93	0.000	90,830.43	1.000	1.00				
2	61	-44,073.01	0.000	88,679.91	0.600	0.42	0.58			
3	92	-43,856.87	0.008	88,518.95	0.521	0.46	0.27	0.27		
4	123	-43,757.53	0.081	88,591.59	0.535	0.22	0.29	0.21	0.29	
5	154	-43,699.27	1.856	88,746.39	0.550	0.09	0.29	0.16	0.29	0.17

Notes: *N* = 6,325. LL = log likelihood. BIC = Bayesian information criterion.

[1] The change in LL compared with the next best LL obtained from the random starts.

[2] Predicted class proportions based on the largest posterior latent class probability.

Despite these supportive findings, we remained concerned that the thresholds and standardized factor loadings of several items varied across latent classes in the original unidimensional factor structure with five items. We therefore re-analyzed the factor mixture model with the remaining four items. The results revealed that the sample remained heterogeneous with respect to a unidimensional measurement structure. The lowest BIC was obtained for a two-class model, but the BIC of the three-class model was also very close in magnitude. Although the distribution of thresholds across three latent classes appeared fairly similar for each of the items, the standardized factor loadings differed substantially for Levn1, ranging from 0.41 to 0.71, and Lenv6, ranging from 0.46 to 0.93. The standardized factor loadings for Lenv3 (ranging from 0.66 to 0.90) and Lenv8 (ranging from 0.89 to 1.00) were more similar across the latent classes.

We then estimated alpha coefficients of internal consistency within each of the latent classes based on the most likely latent class membership so as to determine the extent to which the differences in factor loadings might have affected the reliability of this subscale. Not surprisingly, we found that the internal consistency differed substantially in the three latent class subsamples with values of 0.87 in class 1 (N = 1,753), 0.98 in class 2 (N = 929) and 0.62 in class 3 (N = 3,643). We also estimated the internal consistency in the overall sample and in each of the latent class subsamples when excluding Lenv1 or Lenv6 and we found that the internal consistency remained almost the same. Taken together, these findings suggested that the four items for living environment may not have co-varied in a homogeneous manner with respect to a unidimensional factor structure (assumption two), and that the observed heterogeneity was predominantly associated with the lack of invariance for the parameters associated with Lenv1 and Lenv6.

On the whole, the results of the factor mixture analyses revealed that the sample was heterogeneous with respect to the unidimensional measurement structures for each of the subscales. We found substantial improvements in model fit when only the four most invariant items for each subscale were retained. However, several concerns with respect to two items in the living environment subscale remained. Similar results were obtained when we replicated the confirmatory factor analyses using FIML (see Table 17) and multiple imputation for the subsample of respondents who completed at least one of the MSLSS items (see Appendix C).

Table 17 CFAs of the abridged MSLSS subscales using full information maximum likelihood

Subscale	*N*	# parameters	Log likelihood	BIC
Family	7,228	24	-40,288.28	80,789.82
Self	7,254	24	-37,643.54	75,500.43
Friends	7,300	24	-32,393.39	65,000.28
Living environment	7,174	24	-41,060.66	82,334.40
School	7,245	24	-42,208.51	84,630.34

Notes: BIC = Bayesian information criterion.

3.3.3 Confirmatory factor analyses of the abridged MSLSS

After determining adequate model fit for each of the abridged subscales with only four items, we continued our analyses by combining the abridged MSLSS subscales into a multidimensional measurement model so as to test whether the first-order factors correlated in a consistent manner (assumptions 4 and 5). We first tested a model with five correlated first-order factors because this model underlies the more restrictive second-order factor model. Once we obtained acceptable model fit for the correlated first-order factor model, we could compare the findings for this model with those resulting from a second-order factor model so as to specifically test the unidimensional structure among the first-order factors.

3.3.3.1 CFA results for five correlated first-order factors

The CFA results of the model with five correlated first-order factors for the abridged MSLSS indicated that this model did not fit the data well (see Table 18). Our examination of the residual correlations (see Figure 18) revealed that the patterns of model misfit primarily involved several items from the living environment subscale and several items from the self

Table 18 CFA models of the abridged MSLSS

Missing data technique	*N*	WSLMV χ^2	Df	CFI	RMSEA	SRMR	Residual correlations	
							range	% > |0.1|
Correlated five-factor model								
Single imputation[1]	6,325	4,182.79	102* 160**	0.869	0.080	0.041	-0.13 to 0.18	2.1%
Multiple imputation[2]	7,305	4,834.84 4,971.76	n/a	0.864	0.080	0.042	-0.13 to 0.18	2.1%
Correlated five-factor model excluding two items of the living environment subscale								
Single imputation[1]	6,325	3,115.91	81* 125**	0.907	0.077	0.039	-0.09 to 0.18	1.3%
Multiple imputation[2]	7,305	3,561.06 3,691.96	n/a	0.894	0.077	0.039	-0.96 to 0.18	1.3%
Modified correlated five-factor model with two living environment items								
Single imputation[1]	6,325	1,368.26	80* 120**	0.958	0.050	0.026	-0.07 to 0.07	0.0%
Multiple imputation[2]	7,305	1,546.34 1,606.51	n/a	0.955	0.050	0.026	-0.07 to 0.07	0.0%
Modified 2nd order factor model including the above modifications								
Single imputation[1]	6,325	1,801.15	72* 125**	0.944	0.062	0.035	-0.09 to 0.09	0.0%
Multiple imputation[2]	7,305	2,111.82 2,158.08	n/a	0.938	0.062	0.035	-0.09 to 0.09	0.0%

Notes: χ^2 and fit indices are based on mean and variance adjusted weighted least squares (WLSMV). RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index;

* Df based on WLSMV estimation. ** Df based on number of free parameters.

[1] Single imputation for those with one missing value.

[2] Analyses of multiple imputed datasets for those who completed at least one MSLSS item.

subscale. With respect to the living environment subscale, we found that the correlation for Lenv6 with Lenv1 was overestimated ($r_{residual}$ = -0.13), and that the correlation for Lenv6 and Frnd9 was underestimated ($r_{residual}$ = 0.12). These findings were not surprising considering our previous finding that the adolescents did not respond to Lenv1 and Lenv6 in a consistent manner. It is possible that Lenv6, “My family’s house is nice,” measures something different from the other items that focus on one’s living environment rather than the specific environment of one’s family house. Similarly, Lenv1, “There are lots of fun things to do where I live,” is different from Lenv3, “I like my neighborhood,” and Lenv8, “I like where I live.”

Figure 18 Residual correlations for the abridged MSLSS with five latent factors

	fam2	fam5	fam6	fam7	frnd1	frnd4	frnd8	frnd9	lenv1	lenv3	lenv6	lenv9	schl1	schl5	schl7	schl8	self2	self3	self4	self5
fam2	**FAMILY**																			
fam5	0.01																			
fam6	0.01	0.03																		
fam7	-0.02	-0.02	0.00																	
frnd1	-0.02	0.01	0.01	-0.03	**FRIENDS**															
frnd4	-0.01	0.02	-0.02	-0.04	0.05															
frnd8	-0.03	0.01	0.03	0.03	0.00	0.04														
frnd9	-0.01	0.01	0.03	0.00	-0.07	-0.06	-0.04													
lenv1	-0.03	-0.04	-0.03	0.03	-0.03	-0.05	-0.03	-0.01	**LIVING**											
lenv3	-0.03	-0.01	-0.02	0.05	-0.04	-0.05	0.01	0.03	-0.04											
lenv6	0.05	0.01	0.04	0.04	0.01	0.02	0.04	0.12	-0.13	-0.08										
lenv9	-0.02	-0.03	-0.04	-0.02	-0.06	-0.07	-0.04	0.07	0.01	0.07	0.02									
schl1	0.05	0.04	0.00	0.03	0.07	0.06	0.02	0.01	0.05	0.03	-0.02	0.01	**SCHOOL**							
schl5	-0.02	0.00	0.00	0.01	0.01	0.01	-0.02	-0.01	0.06	-0.01	-0.06	-0.03	-0.07							
schl7	-0.04	-0.01	-0.04	-0.03	-0.05	-0.04	-0.07	-0.06	0.03	-0.02	-0.09	-0.03	0.00	0.06						
schl8	-0.03	0.00	-0.01	0.03	0.00	0.02	0.03	0.00	0.07	0.03	-0.02	-0.01	-0.08	-0.02	0.03					
self2	0.06	0.01	-0.02	0.01	-0.03	-0.07	-0.07	-0.02	0.06	-0.04	0.01	-0.02	0.18	0.00	-0.01	0.08	**SELF**			
self3	0.05	-0.06	-0.03	-0.06	-0.03	-0.05	-0.09	-0.01	0.02	-0.07	-0.01	-0.07	-0.02	-0.02	-0.09	-0.05	0.04			
self4	0.04	0.03	0.03	-0.01	-0.02	-0.05	-0.05	-0.01	0.07	-0.02	0.02	-0.01	0.03	-0.04	-0.07	-0.01	-0.01	0.11		
self5	-0.01	-0.01	-0.01	-0.03	0.09	0.07	0.08	0.08	0.05	-0.01	0.02	-0.03	0.00	0.01	-0.08	0.00	-0.08	0.02	-0.07	

Notes: Based on the single imputation subsample (*N* = 6,325).

Considering these challenges, we decided to remove Lenv1 and Lenv6 from our model and to only include the two remaining items that were shown to be more consistent in the analyses of the living environment subscale, as discussed earlier. The CFA of the model with five correlated first-order factors that included only two living environment items resulted in better overall model fit (WLSMV $\chi^2_{(81)}$ = 3,115.91, RMSEA = 0.077, CFI = 0.907) (see second model in Table 18 on page 117). However, the patterns of misfit involving items in the self subscale remained. The largest residual correlation was observed for the relationship between Self2, "There are lots of things I can do well," and Schl2, "I learn a lot at school" ($r_{residual}$ = 0.18). The correlation between Self2 and Schl8, "I enjoy school activities" was also underestimated ($r_{residual}$ = 0.08). To explain these observations, we speculated that the adolescents' responses with respect to their learning experiences at school might have been influenced by their perception of whether they could do things well in general (item Self2). We examined this by regressing the two items in the school subscale items on Self2. This resulted in a comparatively better fitting model (WLSMV $\chi^2_{(82)}$ = 2,233.82, RMSEA = 0.064, CFI = 0.930) (this model is not included in Table 18). The improvement in chi-square was statistically significant (ΔWLSMV $\chi^2_{(2)}$ = 626.75) and the residual correlations for Self2 with Schl2 ($r_{residual}$ = 0.01) and Schl8 ($r_{residual}$ = 0.01) were very small.

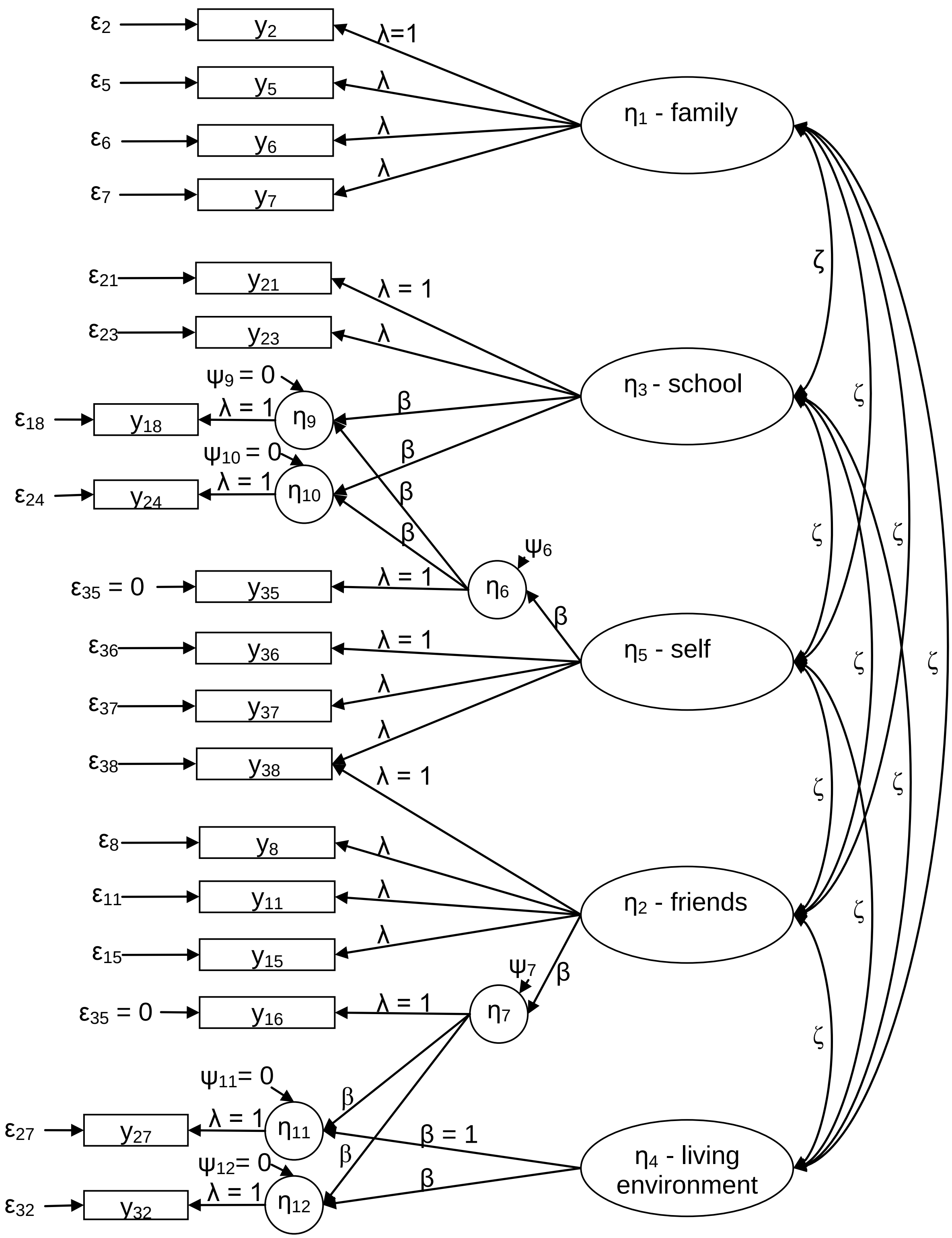
In addition to the patterns of misfit involving Self2, we observed that the correlations between Self5, "Most people like me" and the items of the friends subscale were underestimated ($r_{residual}$ ranging from 0.07 to 0.09). It is perhaps not surprising that the adolescents' experiences of being liked by people (Self5) would have been associated with the relationships they had with their friends and peers. We speculated that, although Self5 was an indicator of satisfaction with self, the item may have also reflected how the adolescents perceived their relationships with their peers and friends. To test this hypothesis, we specified a model that included a cross loading of Self5 on the latent factor for friends.

The resulting model fit the data well (WLSMV $\chi^2_{(81)}$ = 1,532.91, RMSEA = 0.053, CFI = 0.953), and overall model fit was substantially improved in comparison with the previously reported modified model (ΔWLSMV $\chi^2_{(1)}$ = 504.84) (this model is not reported in Table 18). In addition, the correlations among the items of the self subscale were better estimated ($r_{residual}$ ranging from -0.01 to 0.02).

Although the fit indices for this model were acceptable by most standards, we found that we could further improve the model by accounting for the remaining residual correlations between Frnd9, “I have enough friends” and the two living environment items ($r_{residual}$ for Lenv3 = 0.04; $r_{residual}$ for Lenv8 = 0.10 in the previously reported model).We speculated that the degree to which adolescents like where they live might be associated with the degree to which their living environment is conducive to having enough friends. We tested this by regressing the two living environment items on Frnd9. This resulted in further improvement in fit (WLSMV $\chi^2_{(80)}$ = 1,368.26, RMSEA = 0.050, CFI = 0.958) (the modified correlated five-factor model in Table 18 on page 117). The difference in fit between this model and the previously reported model was statistically significant (ΔWLSMV $\chi^2_{(2)}$ = 187.22), and the residual correlations ranged from -0.07 to 0.07. The standardized regression parameters for the relationships between Frnd9 and Lenv3 (β = 0.13) and Lenv8 (β = 0.18) were statistically significant albeit relatively small in comparison with those of the other parameters in the model.

The specification of the correlated five-factor model including all of the above modifications is displayed in Figure 19. All parameter estimates were statistically significant ($p < 0.05$). The standardized factor loadings were smallest for the relationships between Self9 and satisfaction with self and friends ($\lambda_{38,5}$ = 0.46, $\lambda_{38,4}$ = 0.46), and the remaining factor

Figure 19 Modified correlated five-factor structure of the abridged MSLSS



loadings ranged from 0.61 ($\lambda_{36,5}$) to 0.87 ($\lambda_{37,5}$). The correlations among the five latent factors were also all statistically significant with values ranging from 0.30 to 0.69 (see Table 19).

Table 19 Polychoric correlations among latent factors in the modified version of the abridged MSLSS

	Family	Friends	School	Living	Self
Family					
Friends	0.51 (0.49-0.54)				
School	0.47 (0.45-0.49)	0.30 (0.28-0.33)			
Living	0.69 (0.67-0.70)	0.41 (0.38-0.45)	0.45 (0.43-0.47)		
Self	0.60 (0.58-0.62)	0.60 (0.58-0.62)	0.37 (0.34-0.39)	0.52 (0.50-0.55)	

Notes: 95% confidence intervals are shown in brackets. These estimates were based on the sample of adolescents who had a missing value for no more than 1 MSLSS item (*N* = 6,325).

3.3.3.2 CFA results for the indirect reflective measurement structure

The purpose of the above modifications was to obtain a well-fitting model with correlated first-order factors so that we could subsequently examine whether these first-orders factors responded to a common second-order factor. We tested the second-order factor model including the above modifications and found that, although this model resulted in fairly reasonable fit (WLSMV $\chi^2_{(91)}$ = 1,801.15, RMSEA = 0.062, CFI 0.944), it did not fit as well as the correlated first-order factor, as indicated by a statistically significant difference in chi-square (ΔWLSMV $\chi^2_{(5)}$ = 372.19) (see last model in Table 18 on page 117).

We compared the structural parameter estimates of the two models to locate the differences in model fit. This comparison revealed that the parameters of the first-order factor structures were nearly identical in both models. However, some of the correlations between the first-order factors, as implied by the second-order factor structure, differed in comparison with the correlations that were observed in the model with five correlated first-order factors. In particular, the correlation between friends and self (r = 0.60) based on the correlated five-factor model, was larger than the correlation implied by the second-order factor structure (r = 0.49). In contrast, the correlations between friends and living environment (r = 0.414) and friends and school (r = 0.30) were smaller than those implied by the second-order factor structure (r = 0.52 and r = 0.35, respectively). Thus, the correlational structure of the first-order factors had changed somewhat as a result of the constraints implied by the second-order factor. These findings suggested that the correlated five-factor model provided a better representation of the dimensional structure than did the second-order factor model.

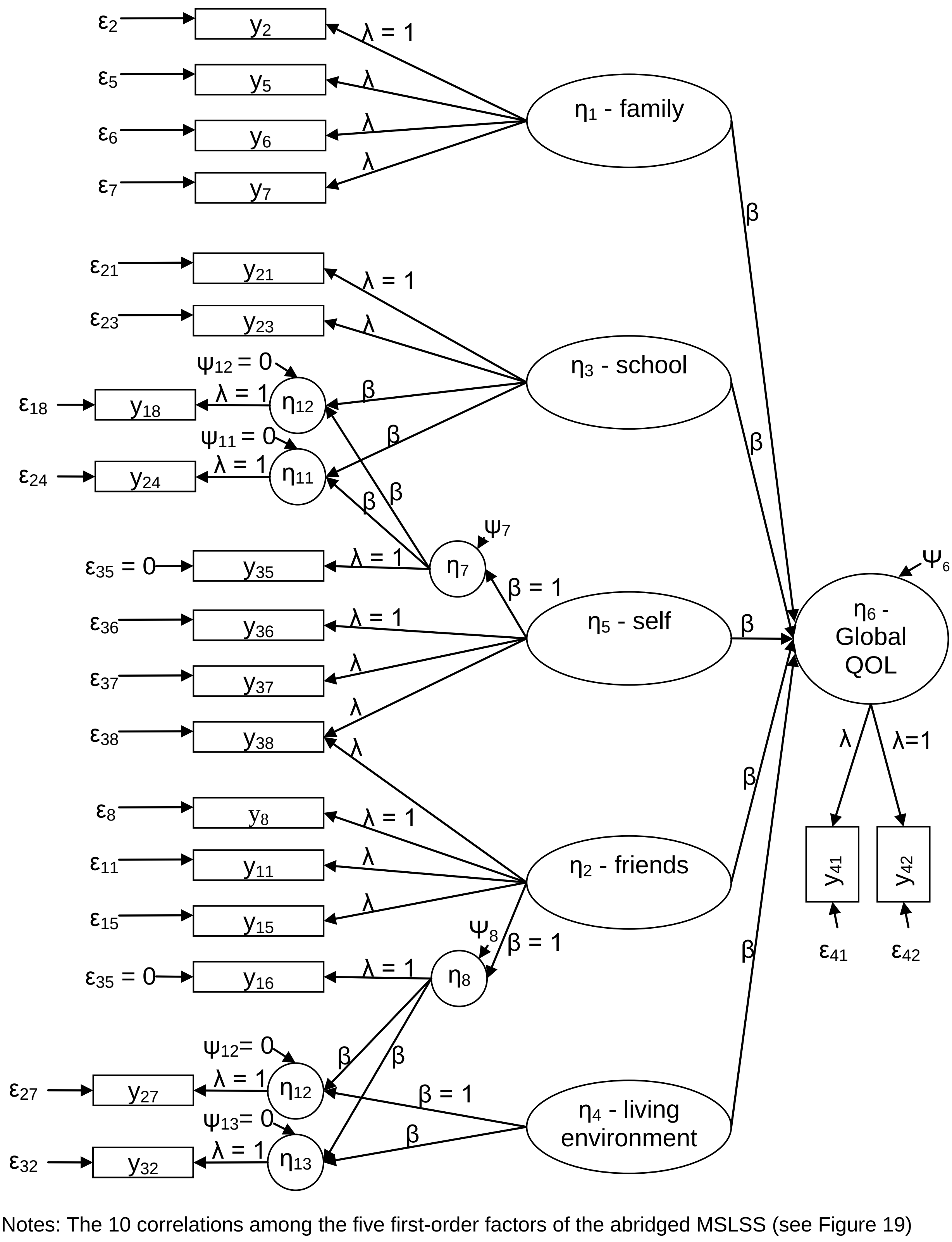
We also examined the relationships between the second-order factor and the first-order factors. Based on the magnitude of the standardized coefficients, we determined that the second-order factor accounted for a substantial percentage of the variance in family (R^2 = 0.68), self (R^2 = 0.75) and living environment (R^2 = 0.63). However, the explained variance in friends (R^2 = 0.43) and school (R^2 = 0.29) was much smaller. This indicated that the second-order factor predominantly represented satisfaction with family, self, and living environment and, to a lesser extent, satisfaction with friends and school.

The findings pertaining to the second-order factor structure suggested that the correlated five-factor model provided a better alternative than the second-order factor model for the purposes of examining relationships between the first-order factors and other variables of interest to our subsequent analyses. Nevertheless, the differences in fit and parameter estimates between the correlated five-factor model and the second-order model were not large.

3.4 Relationships between multidimensional life satisfaction and global QOL

We used the correlated five-factor structure of the abridged MSLSS to examine the degree to which satisfaction with family, friends, school, living environment and self explained the adolescents' perceptions of their quality of life as a whole (global QOL). We continued our analyses by testing a so-called "spurious indicator model" (Edwards & Bagozzi, 2000, p. 166) where the first-order factors explained the variance in global QOL as measured by the satisfaction with quality of life variable and the QOL-ladder (see Figure 20). The reported results for these analyses are based on the single imputation subsample of those who completed both the global QOL question and who had no more than one missing response to the MSLSS questions (N = 6,163). These results are compared with those based on the multiple imputation subsample of adolescents who completed at least one of the MSLSS questions (N = 7,305) (see Table 20). The results of the CFA factor analysis showed that this model fit the data well. Taken together, satisfaction with family, friends, school, environment and self accounted for 67.3% of the variance in global QOL.

Figure 20 A spurious indicator model of the abridged MSLSS and global QOL



Notes: The 10 correlations among the five first-order factors of the abridged MSLSS (see Figure 19) are not shown here.

Table 20 SEM results for the spurious indicator model

Missing data technique	*N*	WSLMV χ^2	Df	CFI	RMSEA	SRMR	Residual correlations	
							range	% > |0.1|
Single EM imputation	6,163	1,590.95	100* 150**	0.954	0.049	0.025	-0.06 to 0.07	0.0%
Multiple imputation	7,049	1,846.26 1,764.52	n/a	0.953	0.049	0.025	-0.06 to 0.07	0.0%

We subsequently used the Pratt index (*d*) (Thomas et al., 1998) to determine the relative importance of each of the life satisfaction dimensions in explaining global QOL (see Table 21). The results indicated that, although global QOL correlated substantially with all of the life-satisfaction dimensions, it was predominantly explained by satisfaction with self (*d* = 66%) and family (*d* = 26%), and, to a much lesser extent, by living environment (*d* = 7%)

Table 21 Relative importance of variables explaining global QOL

Variable	*B*	*SE B*	β	*r*	*D*
Family	0.29	0.02	0.26	0.67	26%
Friends	-0.01	0.02	-0.01	0.51	-1%
Living	0.09	0.02	0.08	0.57	7%
School	0.03	0.02	0.03	0.40	2%
Self	0.65	0.03	0.58	0.78	66%

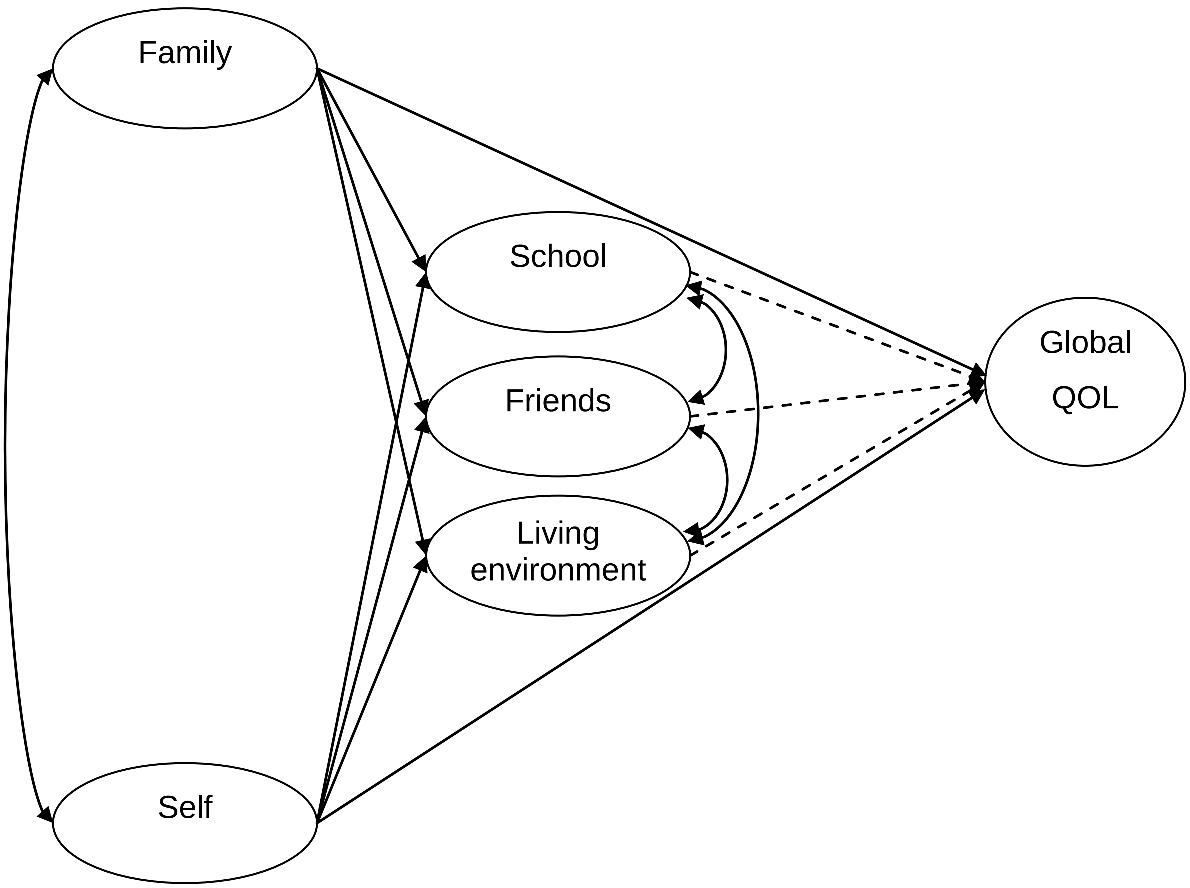
Notes: *r* = estimated correlation with the latent global quality of life variables, *d* = Pratt index. R^2 = 0.67. Analyses were based on the single imputation for subsample of respondents who completed both global quality of life variables and had no more than one missing MSLSS value (*N* = 6,163).

and school (*d* = 2%). It is particularly noteworthy that the relationship between friends and global QOL was not statistically significant in this model.

The above findings suggested that the relationships between global QOL and satisfaction with friends, school, and living environment were confounded by satisfaction with self and family. We tested this by specifying the corresponding regression parameters between these variables in another nested model (see Figure 21). In this model, global QOL and friends, school and living environment were regressed on satisfaction with self and family. The regression parameters of global QOL on satisfaction with friends, school and living environment were all specified to be zero so as to test whether they were fully explained by their associations with self and family. This model resulted in good fit (WLSMV $\chi^2_{(100)}$ = 1,551.38, RMSEA = 0.049, CFI 0.956). The fit indices were nearly identical to those that were obtained for the model with five correlated exogenous factor, and the chi-square difference between this model and the correlated five-factor model was small, albeit statistically significant (ΔWLSMV $\chi^2_{(3)}$ = 21.59, p = 0.0001).

These findings suggested that the explanatory relationships between global QOL and satisfaction with friends, school and living environment were indeed almost entirely explained by their common association with satisfaction with self and family (see Table 22). The Pratt indices suggested that satisfaction with friends was primarily explained by satisfaction with self, and that satisfaction with school and satisfaction with living environment were primarily explained by satisfaction with family. Satisfaction with self and family explained 39.9% of the variance in satisfaction with friends (β = 0.46 for self and β = 0.24 for family), 49.8% of the variance in satisfaction with living environment (β = 0.19 for self and β = 0.58 for family), and 23.9% of the variance satisfaction in school (β = 0.15 for self and β = 0.39 for family).

Figure 21 A heuristic diagram of the relationships among latent factors in a spurious indicator model



Notes: The relationships with dashed lines were constrained to be zero.[31]

---

[31] The dashed lines are shown to demonstrate that the model with these parameters would have been equivalent to the model with five correlated latent factors (our previous model). According to Hershberger (2006), "Equivalent models may be defined as a set of models that, regardless of the data, yield identical (a) implied covariance, correlation, and other moment matrices when fit to the same data, which in turn imply identical (b) residuals and fitted moment matrices, (c) fit functions and chi-square values, and (d) goodness-of-fit indices based on fit functions and chi-square" (p. 15). By fixing the parameters that correspond to the dashed lines at zero, we therefore obtained a model that was nested within the model with five correlated latent factors.

Table 22 Relative importance of satisfaction with self and family explaining satisfaction with friends, school and living environment

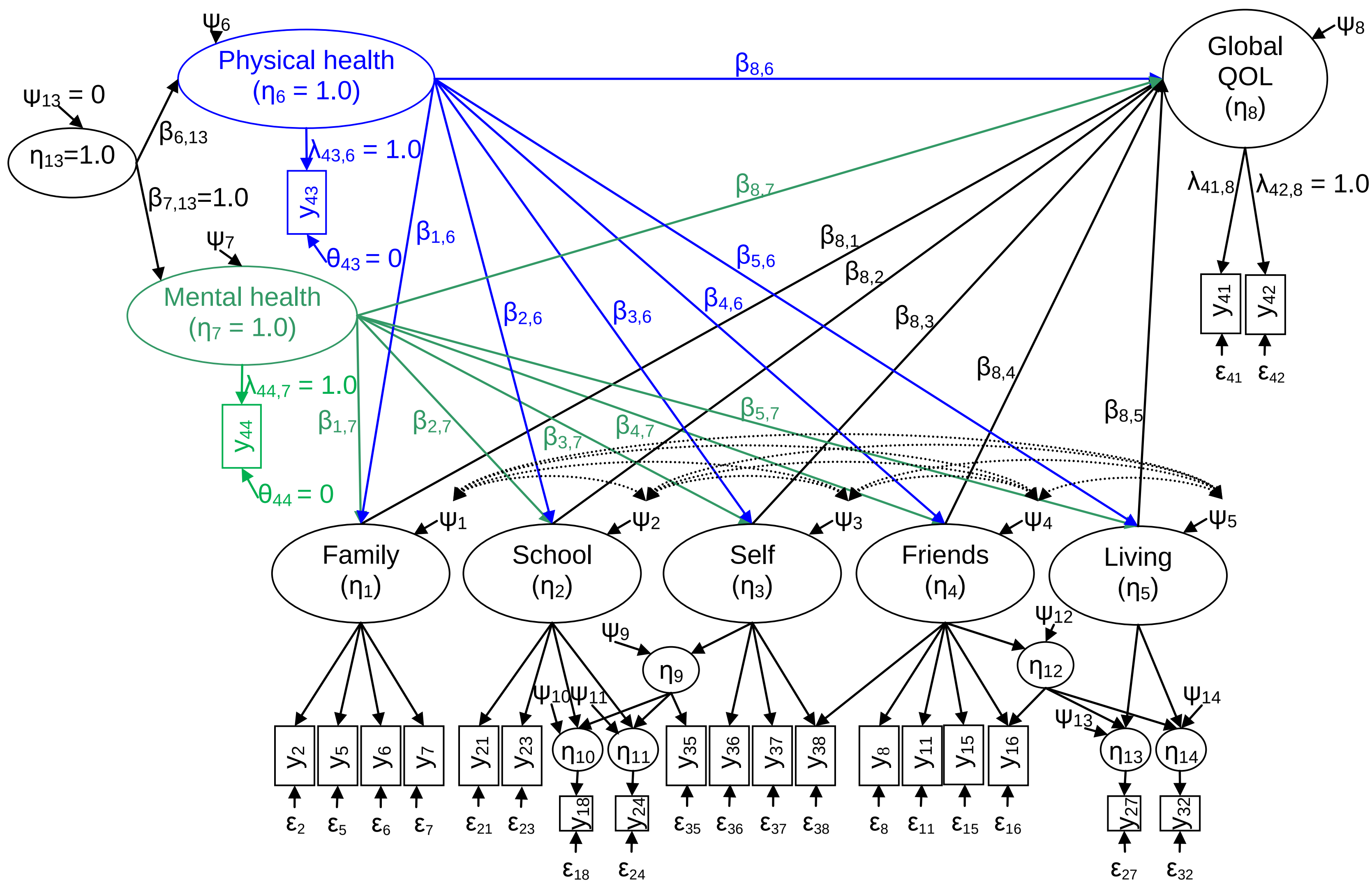
Variable	*B*	*SE B*	β	*r*	*d*
Explaining satisfaction with friends (R^2 = 39.9%)					
Self	0.59	0.02	0.46	0.60	69%
Family	0.31	0.02	0.24	0.51	31%
Explaining satisfaction with school (R^2 = 23.9%)					
Self	0.17	0.02	0.15	0.38	24%
Family	0.44	0.02	0.39	0.77	76%
Explaining satisfaction with living environment (R^2 = 49.8%)					
Self	0.27	0.03	0.19	0.53	20%
Family	0.82	0.03	0.58	0.69	80%

Notes: *r* = estimated correlation with the latent global quality of life variables, *d* = Pratt index. R^2 = 0.68. Analyses were based on the single imputation for subsample of respondents who completed both global quality of life variables and had no more than one missing MSLSS value (*N* = 6,163).

3.5 Relationships with health status and QOL

The final step in our analyses involved an examination of the degree to which perceived mental and physical health status were related to global QOL, and whether these relationships were mediated by the adolescents' perceptions of their family life, friends, school, living environment and self. We adopted the spurious indicator structure to specify the relationships between the life satisfaction dimensions and global QOL. The mental and physical health status variables were specified as correlated exogenous variables that explained the variance in global QOL, and the life satisfaction dimensions were specified as mediators of these relationships as shown in Figure 22. The results reported here were based on multiple imputation of values for those who provided a response to at least one of the

Figure 22 Model specification for the relationships between health status and QOL



Notes pertaining to Figure 22: Constraints for the measurement structure of η1 to η5 are presented in Figure 20. We included η_6 and η_7 to specify perceived mental and physical health status as ordinal variables in the model, and we specified a latent factor, η_{13}, to account for the correlation between η_6 and η_7. The variances and factor loadings for η_6 and η_7 were fixed at 1.0 for identification, and a theta matrix was used to fix the residual variances of the observed ordinal variables for perceived physical and mental health status at zero (see Appendix D for the corresponding Mplus (B. Muthén & L. K. Muthén, 2006) syntax).

MSLSS questions, both health status questions and both global QOL questions (*N* = 6,932). These results were compared with those based on the full multiple imputation subsample (*N* = 8,174), and to those based on the single imputation subsample of respondents who had a missing value for only one MSLSS item (*N* = 6,072) (see Table 23).

Table 23 Structural equation model results

Missing data technique	*N*	WSLMV χ^2	*df*	CFI	RMSEA	SRMR	Residual correlations	
							range	% > |0.1|
Single EM imputation[1]	6,072	1,812.73	116* 178**	0.953	0.049	0.025	-0.07 to 0.07	0.0%
Multiple imputation[2]	6,932	2,083.22 2,010.02	n/a	0.951	0.049	0.025	-0.06 to 0.07	0.0%
Multiple imputation[3]	8,174	2,436.96 2,300.99	n/a	0.948	0.048	0.025	-0.06 to 0.07	0.0%

Notes: * *df* based on WLSMV estimation. ** *df* based on number of free parameters.

[1] based on single imputation subsample of those who had a missing value for only one MSLSS item and complete data for the other variables.

[2] based on the multiple imputation subsample of those who answered both global QOL items, the mental or physical health status items, and at least one of the MSLSS items.

[3] based on the full multiple imputation subsample.

The specified model with the life satisfaction dimensions operating as mediators of the relationships between perceived mental and physical health status and global QOL resulted in acceptable overall fit (see Table 23). Taken together, satisfaction with family, friends, school and self, and perceived mental and physical health status explained 76.1% of the variance in the global QOL. These results indicated that global QOL was significantly explained by the variables in our model.

Before drawing further conclusions, it is important to verify the measurement structure of global QOL. This factor was measured by only two indicators and it was therefore not possible to examine its structure independently of the full model. The standardized factor loadings in the current model were 0.74 for the QOL-ladder (R^2 = 0.55) and 0.75 for the satisfaction with quality of life variable (R^2 = 0.56). Thus, the variance of the global QOL factor was fairly evenly distributed across these two indicators.

With respect to the relationships between latent factors, we found that all but one of the path coefficients for the relationships between the six explanatory variables and global QOL were statistically significant. Again, the smallest and statistically non-significant regression coefficient was for the relationships between satisfaction with friends and global QOL (β = -0.02) and for the relationship between school and global QOL (β = 0.02). We used the standardized regression coefficients and estimated the correlations for the relationships between global QOL and the six explanatory variables to calculate the Pratt index of relative importance (see Table 24). Although all of the variables significantly correlated with global QOL, satisfaction with friends and school together accounted for less than 2% of the explained variance relative to the other variables in the model. Global QOL was mostly explained by satisfaction with self (*d* = 42%), mental health (*d* = 30%), and satisfaction with family (*d* = 20%).

Table 24 Relative importance of variables explaining global QOL

Variable	*B*	*SE B*	β	*r*	*d*
Family	0.29	0.03	0.23	0.66	20%
Friends	-0.02	0.02	-0.02	0.51	-1%
Living environment	0.05	0.02	0.05	0.56	4%
School	0.02	0.01	0.02	0.40	1%
Self	0.62	0.03	0.41	0.78	42%
Physical health	0.04	0.01	0.05	0.49	3%
Mental health	0.26	0.01	0.33	0.70	30%

Notes: *r* = estimated correlation with the latent global quality of life variables, *d* = Pratt index. R^2 = 0.76. Analyses were based on the multiple imputation subsample of respondents who answered both global QOL items, the mental or physical health status items, and at least one of the MSLSS items (*N* = 6,932).

Another part of the model focused on the degree to which each of the life satisfaction dimensions was explained by perceived mental and physical health status. The parameters for this part of the model revealed that mental and physical health status predominantly explained satisfaction with self (33.0%), and, to a lesser extent, satisfaction with family (16.9%), friends (11.3%), and living environment (14.2%) (see Table 25). Only 7.9% of the variance in satisfaction with school was explained by perceived mental and physical health status. The Pratt index values indicated that most of the variance in each of the life satisfaction dimensions was explained by mental health status relative to physical health status.

Table 25 Relative importance of variables explaining the dimensions of life satisfaction

Variable	*B*	*SE B*	β	*r*	*d*
Explaining satisfaction with self (R^2 = 33.0%)					
Physical health	0.11	0.01	0.22	0.45	30%
Mental health	0.22	0.01	0.43	0.54	70%
Explaining satisfaction with family (R^2 = 16.9%)					
Physical health	0.05	0.01	0.08	0.27	13%
Mental health	0.23	0.01	0.36	0.41	87%
Explaining satisfaction with friends (R^2 = 11.3%)					
Physical health	0.07	0.02	0.09	0.24	19%
Mental health	0.24	0.01	0.28	0.33	81%
Explaining satisfaction with living environment (R^2 = 14.2%)					
Physical health	0.07	0.01	0.09	0.26	16%
Mental health	0.28	0.02	0.32	0.37	84%
Explaining satisfaction with school (R^2 = 7.9%)					
Physical health	0.09	0.01	0.11	0.22	32%
Mental health	0.17	0.01	0.21	0.27	69%

Notes: Analyses were based on the multiple imputation subsample of respondents who answered both global QOL items, the mental or physical health status items, and at least one of the MSLSS items (*N* = 6,932).

The parameters for the relationships between mental and physical health status, each of the life satisfaction dimensions, and global QOL could be used to determine the magnitude of the total and the indirect relationships between physical and mental health status and global QOL as mediated by each of the dimensions of life satisfaction (see Table 26). The total relationship between perceived health status and global QOL was larger for perceived

Table 26 Mediation effects for physical and mental health status and global QOL

	Effect of perceived physical health status on global quality of life			Effect of perceived mental health status on global quality of life		
Mediating variable	$B_{indirect}$	*SE B*	% mediation	$B_{indirect}$	*SE B*	% mediation
Family[1]	0.01	0.00	10.8%	0.07	0.01	13.7%
Friends[1]	-0.00	0.00	-1.0%	-0.00	0.00	-0.8%
Living[1]	0.00	0.00	2.8%	0.01	0.01	2.8%
School[1]	0.00	0.00	1.2%	0.00	0.00	0.6%
Self[1]	0.07	0.01	54.0%	0.14	0.01	29.1%
Total indirect effects[2]	0.08		67.8%	0.22		45.4%

Notes: Degree of mediation attributed to each satisfaction variable was calculated as the indirect effect for that variable divided by the total effect for physical or mental health status. Analyses were based on the multiple imputation subsample of respondents who answered both global QOL items, the mental or physical health status items, and at least one of the MSLSS items (N = 6,932).

[1] Indirect effect of physical or mental health status on global quality of life as mediated by one of the satisfaction variables.

[2] Sum of all indirect effects for physical and mental health status explaining global quality of life.

mental health status (β = 0.61), while adjusting for perceived physical health status, than for perceived physical health status (β = 0.17) while adjusting for perceived mental health status.[32] These relationships were partially mediated by the dimensions of life satisfaction (67.8% mediation for physical health and 54.4% mediation for mental health status). The

[32] The total effects were calculated as the sum of the standardized coefficients of the direct and all indirect effects for both physical and mental health status in relation to global QOL.

relationships between the health status variables and global QOL were primarily mediated by satisfaction with self (54.0% mediation for perceived physical health and 29.1% mediation for perceived mental health) and, to a lesser extent, by satisfaction with family (10.8% mediation for perceived physical health and 13.7% mediation for perceived mental health).

Taken together, these results suggested that global QOL was significantly explained by perceived physical and mental health status and satisfaction with self and family. Satisfaction with friends and school did not significantly explain the variance in global QOL when the other variables in the model were controlled. And, the relationships between perceived physical health and mental health status and global QOL were predominantly mediated by satisfaction with self and, to a lesser extent, by satisfaction with family. This was not surprising considering our earlier finding that the relationships between global QOL and satisfaction with friends, school and living environment were almost entirely confounded by satisfaction with self and family.

We had not determined, however, whether the relationships between the health status variables and satisfaction with friends, school and living environment were mediated by satisfaction with self and family. To examine this, we specified another model by replacing some of the correlations among the five life satisfaction dimensions with the relationships that were shown in Figure 21 (see page 127). In this model, the effects of mental and physical health on global QOL were mediated by satisfaction with self and family which, in turn, explained satisfaction with friends, school and living environment. Similar to the model in Figure 21, the relationships between global QOL and satisfaction with friends, school and living environment were constrained to be zero (we had already shown that these relationships were spurious with respect to satisfaction with self and family). The resulting model fit well (WLSMV $\chi^2_{(118)} = 1792.23$, RMSEA = 0.048, CFI 0.953, $N = 6{,}072$) and the difference in chi-square was not statistically significant (ΔWLSMV $\chi^2_{(3)} = 7.97$, $p = 0.047$).[33] The model parameters and their corresponding Pratt indices suggested that the relationships between mental and physical health status and satisfaction with friends, self and living environment were almost entirely mediated by satisfaction with self and family (see Table 27).

3.6 Summary of findings

The results of our analyses indicated that the indirect reflective measurement structure of the 40-item MSLSS did not fit well. The negatively worded items were identified as a predominant source of model misfit. In addition, the factor mixture analyses revealed that the adolescents in our sample did not respond to the positively worded items in a consistent manner. Better model fit for each subscale was obtained when only the four most consistent items for each subscale were retained. The correlated five-factor model for the abridged MSLSS resulted in good fit after excluding two of the living environment items and after allowing for a few theoretically defensible modifications. The corresponding indirect reflective model (second-order factor model) did not fit as well, and we therefore proceeded with our analyses based on a model with five correlated latent factors for the life satisfaction dimensions.

---

[33] This chi-square difference was based on a comparison with the first model in Table 23 (see page 131) based on single imputation subsample of those who had a missing value for only one MSLSS item and complete data for the other variables ($N = 6{,}072$). The findings were similar with multiple imputation but the chi-square differences were not computed because: (a) the chi-squares differed for each multiple imputation sample and (b) WLSMV may have produced modest differences in the estimated degrees of freedom (see Table 30).

Table 27 The relative importance of variables explaining satisfaction with friends, school and living environment

Variable	*B*	*SE B*	β	*r*	*d*[*]
Explaining satisfaction with friends (R^2 = 41.8%)					
Self	0.79	0.04	0.49	0.62	72%
Family	0.35	0.03	0.26	0.53	32%
Physical health	-0.03	0.01	-0.04	0.33	-3%
Mental health	-0.02	0.02	-0.03	0.24	-1%
Explaining satisfaction with school (R^2 = 49.7%)					
Self	0.17	0.03	0.11	0.37	16%
Family	0.50	0.03	0.38	0.47	74%
Physical health	0.04	0.01	0.06	0.22	5%
Mental health	0.03	0.01	0.03	0.27	3%
Explaining satisfaction with living environment (R^2 = 23.7%)					
Self	0.24	0.04	0.15	0.52	15%
Family	0.82	0.04	0.57	0.69	79%
Physical health	0.01	0.01	0.01	0.26	0%
Mental health	0.06	0.02	0.07	0.37	5%

Notes: Analyses were based on the multiple imputation subsample of respondents who answered both global QOL items, the mental or physical health status items, and at least one of the MSLSS items (*N* = 6,932). Multiple imputation WLSMV χ^2 = 1994.69 to 2066.30, RMSEA = 0.048, CFI = 0.952.

* Pratt index values for some variables do not sum exactly to 100% due to rounding error.

We subsequently examined the degree to which the life satisfaction dimensions contributed to global QOL by testing a spurious indicator model. This model fit well, and the parameter estimates indicated that global QOL was predominantly explained by satisfaction with self and family. However, the unexplained variance in global QOL revealed that global QOL also was explained by factors that were not accounted for in this model. Thus, the life satisfaction dimensions were not equivalent to, or comprehensively deterministic of, global QOL. We also specified an equivalent model so as to explicitly determine whether the relationships between global QOL and satisfaction with friends, school and living environment were explained by satisfaction with self and friends. The results of this model confirmed that this was indeed the case.

In the last set of analyses we sought to examine whether perceived physical and mental health status explained global QOL when controlling for each of the life satisfaction dimensions. We also examined whether the relationships between global QOL and perceived physical and mental health status were mediated by satisfaction with family, friends, school, living environment and self. The results revealed that physical and mental health status substantially contributed to global QOL. The relationships between global QOL and perceived physical and mental health status were primarily mediated by satisfaction with self and family. In addition, although the physical and mental health status variables were significantly associated with all the life satisfaction dimensions, we found that their relationships with satisfaction with friends, school and living environment were mediated by satisfaction with self and, to a lesser extent, satisfaction with family.

The findings as a whole indicated that global QOL was predominantly explained by the adolescents' perceptions of their selves, their families, and their mental health status.

Satisfaction with friends, living environment, and school, and physical health status also were associated with global QOL, but these relationships were almost entirely accounted for by the adolescents' satisfaction with self, family and perceived mental health status.

4 DISCUSSION AND CONCLUSIONS

The results of our analyses have several theoretical and methodological implications pertaining to the measurement of various dimensions of life satisfaction and global QOL and their relationships with perceived physical and mental health status in adolescents. We first discuss the implications with respect to the following three objectives of our analyses: (a) to test the assumptions underlying the putative indirect reflective measurement structure of the MSLSS, (b) to determine the degree to which the dimensions of life satisfaction explain global QOL, and (c) to examine whether perceived mental health status, perceived physical health status, or both contribute to global QOL, and whether the dimensions of life satisfaction mediate these relationships. Several methodological recommendations pertaining to the use of CFA and FMA for examining the reliability and validity of measures based on reflective measurement structures are also discussed. Recommendations for further study are highlighted throughout. We conclude with a brief overview of some limitations of our analyses.

4.1 Implications for the measurement structure of the MSLSS

Our first analytical objective was to test the assumptions of the putative measurement structure of the Multidimensional Students' Life Satisfaction Scale (MSLSS) (Huebner, 1994) with the goal of assessing the reliability and validity of this instrument with respect to the measurement of adolescents' satisfaction with their family, friends, living environment, school and self, and their general life satisfaction. The findings revealed several concerns about this instrument for the measurement of general life satisfaction and the dimensions of life satisfaction. First, the CFA results of the indirect reflective measurement structure indicated that this specified structure did not accurately account for the covariances among the 40 MSLSS items. Thus, the 40 items did not provide a valid measure of general life satisfaction. Second, the CFA results of the correlated five-factor structure indicated that the negatively worded items did not co-vary in a consistent manner with the positively worded items, although such consistency was implied by the measurement structure. In addition, the FMA results for each of the subscales suggested that the adolescents did not respond to some of the positively worded items in each subscale in a consistent manner. Thus, the items for each subscale were not reliable with respect to the measurement of satisfaction with family, friends, living environment, school, and self. Third, good model fit for each of the subscales was obtained when we excluded the negatively worded items and when we retained only the four items with the most invariant model parameters across latent classes. We also obtained good model fit for the correlated five-factor model when specifying a few theoretically defensible modifications to account for the otherwise unexplained correlations among some of the items. These findings provide preliminary support for an abridged version of the MSLSS with five correlated first-order factors.

Our findings pertaining to the lack of fit of the indirect reflective measurement structure of the original 40-item version of the MSLSS is somewhat surprising considering the results of other published CFAs that purportedly provided support for a measurement structure with five correlated latent factors corresponding to the dimensions of life satisfaction (Gilman, 1999; Gilman et al., 2000; Greenspoon & Saklofske, 1998; Huebner et al., 1998; Park, 2000; Park et al., 2004), as well as a measurement structure with a second-order factor that purportedly represented general life satisfaction (Gilman, 1999; Gilman et al., 2000; Huebner et al., 1998). We therefore set out to examine some possible explanations for these differences in study findings by comparing our analyses strategies with those utilized by other researchers who reported CFA results of the MSLSS. We also compared findings pertaining to the correlations among dimensions of life satisfaction across published studies so as to further examine whether these dimensions respond in a consistent manner to a common source, which Huebner (1998) labeled "general life satisfaction". If this were the case, then we would expect the correlations among the dimensions of life satisfaction to be fairly consistent across studies. We continue our discussion of objective one with some suggested explanations for our finding that the adolescents did not respond to some of the MSLSS items in a consistent manner, and we conclude with some recommendations for the use of the MSLSS in future studies.

4.1.1 Comparison of our analysis strategies with those utilized in other published CFAs of the MSLSS

There were substantial differences in the analytical strategies that we used to examine the factor structure of the MSLSS in comparison with those used by other researchers. In the following paragraphs, we demonstrate how these different analytical strategies may explain the different results in our study as compared with the results in other published CFAs of the MSLSS. We specifically address three methodological considerations: (a) would we have obtained different findings if we based our analyses on the assumption that the data were continuous and interval-based, as was done in published CFAs of the MSLSS? (b) to what degree might the use of item parceling methods in other published CFAs have concealed potentially important areas of model misfit? and (c) is it possible that we would have arrived at different conclusions had we solely relied on the global model fit indices while ignoring the residual correlations?

The first consideration is based on the fact that, to our knowledge, all published CFAs of the MSLSS assumed normal distributions based on continuous, interval-based measurement for each of the observed indicators. Although this assumption might be justifiable for some ordinal variables in general, we found that the MSLSS variables were clearly not normally distributed in our study. Our concern was that ignoring this assumption would result in downwardly biased Pearson correlations for the relationships among the observed indicators. Thus, the observed Pearson correlation matrix would not have accurately represented the data, and the subsequent estimation of model fit may not have resulted in trustworthy parameter estimates. For example, if we were to determine the difference between the *observed* Pearson correlations and the *observed* polychoric correlations for the matrix of 40 MSLSS variables, we would have found that the differences ranged from -0.15 to 0.01. Thus, some of the Pearson correlations would clearly have been underestimated. We also would have found that 2.8% of the correlations differed by more than 0.10, and 38.1% by more than 0.05. In addition, we would have found that most of the differences involved items of the satisfaction with friends subscale.[34] This occurred because most of the adolescents provided relatively high ratings of their satisfaction with their friends (i.e., the satisfaction with friends subscale variables were very skewed in their distributions, as shown in Figure 8). Although we would have found better global fit had we used maximum likelihood estimation to fit the correlated five-factor model using Pearson correlations (i.e., assuming continuous variables) (MLM $\chi^2_{(125)}$ = 15,138.17, RMSEA =

[34] Thus, not only are most Pearson correlations smaller in magnitude than their corresponding polychoric correlations, the two correlation matrices also differ with respect to the pattern of the 780 observed correlations for the 40 MSLSS variables. The distributions of some MSLSS variables deviated more from normality than others (see Figure 8).

0.056, CFI 0.84), our model fit would not have been based on correlations that accurately represented the distributions of the observed variables. In addition, we would have found that the residual correlations ranged from -0.33 to 0.66 with 32.3% of the residual correlations having absolute values larger than 0.10. Instead, we used probit regressions and polychoric correlations to account for the ordinality of the MSLSS variables, which may explain some of the differences between the results found in our study and those found in published CFAs of the MSLSS.

A second consideration is that most of the published CFAs were based on the analysis of item parcels (Gilman, 1999; Gilman et al., 2000; Huebner et al., 1998; Park, 2000; Park et al., 2004),[35] whereas the models in our study were based on observed variable scores. Little et al. (2002) offered the following definition of an item parcel: "A parcel can be defined as an aggregate-level indicator comprised of the sum (or average) of two or more items, responses, or behaviors" (p. 152). When using item parcels, the model is estimated based on the correlations (or covariances) among the item parcels rather than the observed variables. Item parceling can therefore be used to reduce the shared systematic and random error by averaging the effects of these errors across the items that share the same parcel. Generally, the use of item parceling methods results in better fitting models (Bandalos, 2002; Little, Cunningham, Shahar, & Widaman, 2002). However, an important assumption of item parceling is that the factor structure is known to be unidimensional. Item parceling can lead

[35] To our knowledge, the CFA published by Greenspoon and Saklofske (1998) was the only CFA of the MSLSS that did not involve the use of item parcels. Although they concluded that the model fit indices were indicative of "adequate fit" (p. 968) for the model with five correlated first-order factors, it is noteworthy that the fit was much less favorable than those that were reported by Gilman (Gilman, 1999; Gilman et al., 2000) and Huebner et al. (1998), who based their analyses on item parcels. Indeed, we concur with Shevlin, Miles, and Lewis (2000) who suggested that the findings by Greenspoon and Saklofske may not be indicative of adequate model fit.

to inaccurate parameter estimates and misleading conclusions about model fit, if this assumption is not met (Bandalos, 2002; Little et al., 2002; Nasser, 2003). A particular concern is that item parcels can hide areas of model misspecification and, thereby, lead to an increased chance of a Type II error. Our concern is, therefore, that the use of item parcels in previously published CFAs of the MSLSS may have led to overly favorable model fit results (i.e., Type II error).

To illustrate this concern, we compared the findings of our study to those that we would have obtained had we used item parcels. We applied both the parceling criteria that were used by Gilman (1999) and by Huebner et al. (1998).[36] As expected, the model's fit was much improved in both cases. For example, if we were to apply Gilman's criteria to create 18 parcels, we would find acceptable overall fit for a correlated five-factor model (MLM $\chi^2_{(125)}$ = 2,629.37, RMSEA = 0.056, CFI 0.95).[37] Similar results would be obtained if we were to apply the same criteria to replicate the 20 item parcels in the study by Huebner et al. (MLM $\chi^2_{(160)}$ = 3,544.56, RMSEA = 0.058, CFI 0.94). Thus, we may have concluded that this model fit the data reasonably well.

The areas of misfit identified in our analyses would not have been revealed if we had relied on item parceling methods. In particular, our analyses revealed that a unidimensional

[36] Gilman's (1999; Gilman, Huebner, & Laughlin, 2000) parceling criteria were: (a) "an individual item comprising each parcel was not paired with the item that it most strongly correlated with" and (b) "adjacent items in the MSLSS that measured the same construct were not combined" (p. 60). The items for which these criteria held were subsequently randomly assigned to 3 or 4 parcels for each of the MSLSS subscales (3 parcels for the family and self subscales, and 4 parcels for the other subscales). The same procedure was followed by Park (2000; Park et al., 2004). However, Huebner et al. (1998) created four item parcels for each of the subscales. We therefore created two sets of items parcels to replicate both Gilman's and Huebner's approaches.

[37] We used the mean adjusted maximum likelihood (MLM) estimator in Mplus 4.2 (B. Muthén & L. K. Muthén, 2006), which is analogous to the Satorra-Bentler chi-square (see L. K. Muthén & B. Muthén, 2006).

structure for each of the subscales was not supported because the adolescents did not respond to the negatively worded items, and some of the positively worded items, in a consistent manner. The use of item parceling methods was not justifiable in our study. We therefore conclude that some of the differences between our CFA results and the published CFA results likely can be attributed to the use of item parceling methods in previously published CFAs of the MSLSS.

A third consideration is that we emphasized the importance of examining the residual correlations as the basis for assessing model fit. We found that only relying on global fit indices to evaluate our models could lead to misleading conclusions. Unfortunately, the residual correlations were not mentioned in any of the published CFAs of the MSLSS. However, Gilman (1999) and Huebner (1998) did report the observed correlations among the item parcels. When we specified a correlated five-factor model, using the correlation matrix provided by Gilman (1999), we found that, although the obtained fit indices were similar to those reported by Gilman, the residual correlations among the item parcels ranged from -0.13 to 0.23, and 18 (11.8%) of the 153 correlations had absolute values that were greater than 0.10. Similar findings were obtained when we used the correlation matrix provided by Huebner et al. (1998) to estimate the model. Thus, based on an inspection of the residual correlation matrices, we concluded that the correlated five-factor model did not fit their data well.

We therefore conclude that the differences between our findings and published CFAs of the MSLSS could be attributed, at least in part, to methodological differences. When we accounted for the ordinal nature of the observed variables and based the CFA on the original item responses (not item parcels), we found that the indirect reflective measurement structure for the MSLSS did not fit the data we obtained from our sample of adolescents. In addition, the correlated five-factor model did not fit well, and independent CFAs of each of the MSLSS subscales did not result in good model fit. Thus, the specified measurement structure of general life satisfaction and satisfaction with family, friends, living environment, school, and self was not valid with respect to the 40 items evaluated in our study.

4.1.2 An explanation of the adolescents' inconsistent responses to some items

The next step in our analysis was to find a solution to our concerns about the lack of reliability and construct validity evidence for the MSLSS. We set out to identify patterns of potential misfit in the measurement structure and found that lack of fit could be partially attributed to inconsistencies in the responses to the negatively worded items. The results of the FMAs suggested that the adolescents also did not respond to some of the positively worded items in a consistent manner. We continued by excluding items with inconsistent response patterns from our analyses. The question remains, however, as to why the adolescents may not have responded to some of the positively worded items in a consistent manner. Although this question was not explicitly addressed in our study, we offer a few possible explanations and suggestions for further study.

To address the question about the inconsistencies in the item responses it is useful to consider the actual wording of the items. One can readily observe that almost all of the items include the words *I, me,* or *my*. These items required the adolescents to reflect on how they viewed themselves in various social contexts. For example, although the items, "I like spending time with my parents" and "I enjoy being at home with my family" were part of the family subscale, they may have conflated the adolescents' perceptions of their families with how they viewed *themselves*, at least in the context of their families. Similarly, the items, "I have a bad time with my friends" and "I wish I had different friends" in the friends subscale may also relate to how the adolescents viewed *themselves* in relation to others.

The above examples are provided to illustrate that several MSLSS items may have conflated the adolescents' evaluations of their selves with their satisfaction with various other aspects of their lives. The concern is that adolescents typically view themselves in fragmented and seemingly incongruent ways (Harter, 1999). For example, adolescents may describe themselves differently and inconsistently in different relational contexts (e.g., they may describe themselves very differently in relation to their peers than in relation to their friends). It is therefore plausible that the adolescents may not have responded to some items that involved a degree of self evaluation in a consistent manner. This might be particularly concerning for those items where the social context may have been variously understood. For example, the terms "family" or "home" may have evoked a wide variety of social contexts and may therefore have resulted in response patterns that were inconsistent with respect to the other items of the family subscale. Similarly, the term "neighborhood" may have taken on very different meanings for adolescents in urban regions than for those who lived in rural regions (Greenspoon & Saklofske, 1997).

Although the above considerations do not provide definitive answers to our questions, they do provide some guidance for further research to determine what adolescents think about when they respond to the MSLSS items. For example, cognitive interviewing or talk-aloud protocols could be used to obtain qualitative data for this purpose (Drennan, 2003). It also is possible that the differences in how adolescents respond to some of the items may be explained by differences associated with culture, personality factors, or other variables that affect interpretation of the items. The FMAs in this study could be elaborated upon by regressing the latent classes on various psychological, social, and demographic variables so as to explain the latent class membership. This was not the purpose of our analyses and, despite the inconsistencies in some item responses, we were able to identify a subset of items for each subscale to which the adolescents responded in a relatively consistent manner.

4.1.3 A discussion of whether a second-order factor structure is warranted

We now turn to the question of whether a second-order factor model would be justifiable for the purposes of measuring *general* life satisfaction. Good fit was obtained for the model with five correlated first-order latent factors, after specifying a few theoretically defensible model modifications to account for the otherwise unexplained correlations among some of the observed indicators of the abridged MSLSS. However, the second-order factor structure (i.e., the indirect reflective model) did not fit the data as well, although the fit indices for the second-order factor model approximated the suggested criteria of good fit and some researchers may have concluded that the model fit was acceptable. The question, therefore, remains as to what these findings mean with respect to the use of the indirect reflective measurement structure to obtain a total score of general life satisfaction, as was suggested by Huebner (1998). That is, do the dimensions of life satisfaction correlate in a consistent manner with respect to a common source?

To address this question, we compared the correlations among the five latent factors in our study with other researchers' reported correlations among MSLSS derived scores of satisfaction with family, friends, living environment, and self.[38] Based on assumption four,

[38] Another approach to address this question would have been to test a factor mixture model of the second-order factor structure. Inconsistency of the second-order factor loadings across two or more latent classes would have

we proposed that, if a second-order factor were to provide a reliable and valid measure of general life satisfaction, the patterns of correlations among the dimensions of life satisfaction would have to be fairly consistent across different studies.[39] If this were not the case, then we would have to conclude that the dimensions of life satisfaction could not be reliably represented by a second-order factor and that a total score for general life satisfaction would therefore not be warranted.

Based on a literature search of several databases (including CINAHL, PubMed, Embase, PsychINFO, Social Sciences Citation Index, Science Citation Index, Dissertation Abstracts International, and Health and Psychosocial Instruments), we found 15 studies that reported correlations among the dimensions of life satisfaction (see Table 30 in Appendix E).[40] We used meta-analysis techniques to estimate the average values and the magnitude of the between-studies variance for each of the correlations among the 5 factors.[41] We then compared the average correlations to those that were obtained in our study to determine whether these were, on average, similar in magnitude.

---

meant that these factor loadings were not invariant with respect to unobserved differences among the adolescents. Although we attempted to estimate such a model, we were not able to consistently replicate the log likelihood values. This was not surprising considering the large number of parameters that needed to be estimated. Thus, we were not confident that we were obtaining a maximum likelihood estimate (rather than a local maximum) and decided to not report these analyses (even though the results that were obtained suggested that there may be substantial differences in the second-order factor loadings across at least two latent classes).

[39] Assumption four was described as follows: the first-order latent factors consistently correlate because of a second-order latent factor (i.e., the second-order factor loadings are invariant) (see page 37).

[40] Huebner et al. (1998) did not report the correlations among the first-order factors. However, we were able to replicate their model by using the observed correlations among the item parcels as reported in their publication.

[41] The search strategy, meta-analysis methods and results are described in Appendix E. Although we recognize that it is not conventional to report additional analyses in a discussion section, it seemed appropriate to report the results of these analyses here as a formal approach to comparing our findings with those of other published studies.

Table 28 A comparison of correlations among dimensions of life satisfaction in our study with the average correlations found in other published studies

	Correlations in our study	Fixed effects model analysis			Random effects model analysis		
Relationships	*r* (95% CI)	Average *r* (95% CI)	*Q*	*p*	Average *r* (95% CI)	*Q*	*p*
family with friends	0.53 (0.51 - 0.55)	0.29 (0.26 - 0.32)	39.26	0.02	0.30 (0.26 - 0.34)	21.41	0.56
family with school	0.46 (0.43 - 0.48)	0.38 (0.35 - 0.41)	44.74	0.00	0.38 (0.33 - 0.42)	21.48	0.55
family with living environment	0.68 (0.66 - 0.70)	0.44 (0.41 - 0.46)	120.20	0.00	0.42 (0.35 - 0.49)	23.33	0.44
family with self	0.60 (0.58 - 0.62)	0.42 (0.4 - 0.45)	97.30	0.00	0.45 (0.39 - 0.51)	18.79	0.71
friends with school	0.28 (0.26 - 0.31)	0.19 (0.16 - 0.22)	88.81	0.00	0.21 (0.14 - 0.27)	22.65	0.48
friends with living environment	0.44 (0.40 - 0.47)	0.32 (0.29 - 0.35)	157.48	0.00	0.34 (0.25 - 0.42)	18.61	0.72
friends with self	0.62 (0.60 - 0.64)	0.42 (0.39 - 0.45)	249.21	0.00	0.46 (0.36 - 0.54)	16.07	0.85
school with living environment	0.45 (0.42 - 0.47)	0.32 (0.29 - 0.35)	71.86	0.00	0.32 (0.26 - 0.38)	28.33	0.20
school with self	0.36 (0.33 - 0.38)	0.31 (0.28 - 0.34)	48.86	0.00	0.31 (0.27 - 0.36)	21.42	0.56
self with living environment	0.53 (0.50 - 0.55)	0.31 (0.28 - 0.33)	86.79	0.00	0.32 (0.25 - 0.38)	24.98	0.35

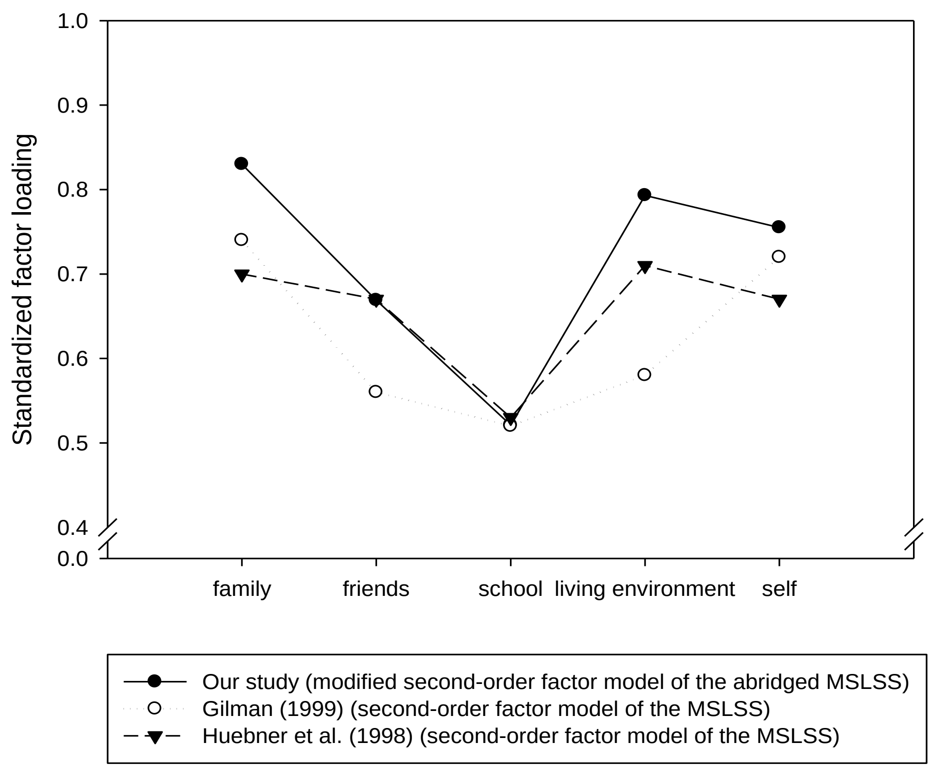
Notes: Average *r* = average of the 23 correlations in 15 published studies. *Q* = statistic for the homogeneity of correlations between studies (*df* = 22) (Lipsey & Wilson, 2001). *p* = statistical significance of the *Q* statistic.

The results of these comparisons revealed that the correlations for the 10 relationships (family with friends, family with school, family with living environment, … etc.) varied significantly across the correlations in the 23 published correlation matrices that we examined (i.e., the between-studies variance was statistically significant for each of the correlations) (see Figure 24 in Appendix E and Table 28 below). We also observed that several correlations in our study were, on average, larger than the correlations found in other published studies. This was not unexpected considering that we had improved the reliability of the measurement structure by excluding items with inconsistent response patterns.

With respect to the pattern of correlations, we observed that, although there were similarities between our findings and the average of the published correlations (e.g., the correlations between satisfaction with self and family, and satisfaction with self and friends were larger than most other correlations), there were some noteworthy differences. In particular, the four correlations involving satisfaction with family, as well as the correlations between satisfaction with self and living environment, were significantly larger in our study. The observed differences in the patterns of correlations found in other studies of the MSLSS suggest that the dimensions of life satisfaction may not co-vary in a consistent manner.

In addition to comparing the correlations among the dimensions of life satisfaction, we compared the standardized second-order factor loadings in our study (i.e., the dimensions regressed on the second-order factor of general life satisfaction) with those reported by Gilman (Gilman, 1999; Gilman et al., 2000) and Huebner et at. (1998). Again, these loadings should have been similar if inferences pertaining to the second-order factor were indeed to be comparable across different studies. As shown in Figure 23, there were some observable similarities and differences in the magnitude of the factor loadings. For example, the second-order factor loading for satisfaction with school was of the smallest magnitude in all the studies. The second-order factor loading for satisfaction with friends, however, was smaller in the study by Gilman than that found in the other studies. In addition, we found larger standardized second-order factor loadings for satisfaction with family, living environment, and self in comparison with those reported by Huebner et al. and Gilman.

Figure 23 Standardized second-order factor loadings for our study and published CFAs of the MSLSS.



There are, of course, numerous possible explanations for these differences. For example, the youth in Huebner et al.'s (1998) sample were younger (mean age = 10.9, *SD* = 2.0) than the participants in our study and Gilman's study (mean age = 16.1, *SD* = 1.1) (Gilman, 1999; Gilman et al., 2000). It is possible that the second-order general life satisfaction factor relates differently to the specific dimensions of life satisfaction for children or adolescents that differ in age. In addition, the abridged MSLSS in our study included only 18 of the 40 original MSLSS items. Thus, it could be argued that the second-order factor in our study was not comparable with that of the other two studies. Nevertheless, if the second-order factor structure were to be valid as a measurement of general life satisfaction, then the items should, to some degree, be exchangeable (Zumbo, 2007). That is, the measurement of a construct should not be dramatically affected by the inclusion or exclusion of items that reflect that construct. We would therefore expect the factor structure to be similar based on the premise that the items ought to be, to some degree, exchangeable.

Finally, it is noteworthy that the second-order factor accounted for less than 50% of the variance in three of the five first-order factors in the studies by Huebner et al. (1998) and Gilman (Gilman, 1999; Gilman et al., 2000), and in two of the first-order factors in our study. This means that the second-order factor accounted for a relatively small percentage of the variance in some of the observed indicators of the first-order factors. For example, in our study we found that the second-order factor accounted for less than 25% percent of the variance in 9 of 18 items of the abridged MSLSS. Similarly, in Gilman's study, the second-order factor accounted for less than 25% of the variance in 14 of the 19 item parcels. Although there are no criteria for how much variance a second-order factor should explain in its first-order factors and observed variables, the finding that the second-order factor explained less than 25% of the variance in one half of the observed variables and less than one half of the item parcels raises some question about what the second-order factor actually measured.

Although these observations do not necessarily disprove a second-order factor structure, they do warrant caution regarding the validity of a second-order factor for the measurement of general life satisfaction and the comparability of general life satisfaction scores across different samples (i.e., the loadings might not be consistent across different samples). Further studies are needed to determine the extent to which the second-order factor loadings are invariant. For example, multi-group CFAs could be conducted to compare second-order factor loadings across observed groups (e.g., age groups, gender, geographic, and ethnic groups), and FMAs could be used to examine the invariance of the second-order factor loadings across unobserved latent classes in very large samples. In the meantime, we suggest that the second-order factor model may not be the best approach for the measurement of general life satisfaction.

4.1.4 Recommendations for the use of the MSLSS

We conclude our discussion of the measurement structure of the MSLSS with a few recommendations pertaining to the use of this instrument for the measurement of dimensions of life satisfaction in adolescents. Our first recommendation relates to the use of the negatively worded items in the MSLSS. The findings suggested that the adolescents did not respond to the negatively worded items in a manner that was consistent with their responses to the positively worded items. Several researchers have raised similar concerns with the use of negative-worded items for the measurement of psychological variables in children and adolescents, such as those related to self-esteem and self-concept (Barnette, 2000; Borgers et al., 2004; Marsh, 1986, 1996). It has been noted that the use of negatively worded items, with children and adolescents, results in bias associated with differences in age, verbal ability (Marsh, 1986), and gender (Fletcher & Hattie, 2005), that the responses are inconsistent with responses to positively worded items (Borgers et al., 2004), and that the combination of negatively- and positively worded items results in lower internal consistency (Barnette, 2000).

One approach to addressing these concerns is to model the effect of negatively worded items by including a method factor. This has, for example, been recommended for the Rosenberg Self-Esteem Scale (Rosenberg, 1989) by some researchers (e.g., Marsh, 1996; Tomas & Oliver, 1999). We tried this approach and found that, although the model with the method factor resulted in better fit (see Figure 5 and Table 7), it did not sufficiently account for the patterns of misfit involving the negatively worded items. Another approach is to exclude the negatively worded items altogether (e.g., Barnette, 2000; Fletcher & Hattie, 2005; Marsh, 1996). With respect to this approach, we found that the exclusion of negatively worded items resolved many patterns of misfit in the residual correlation matrix (see Figure 15 on page 101).

Based on these findings, and considering other findings pertaining to the use of negatively worded items in studies involving adolescents in general, we suggest that it is justifiable to use only the positively worded items for the measurement of the dimensions of life satisfaction in adolescents. However, since the purpose of negatively worded items is to address potential concerns with acquiescence bias, we concur with Marsh (1996) who suggested that it may still be valuable to include some negatively worded items in the questionnaire so as to assess the possibility of acquiescence bias.[42] These items, however, should not be used for scoring purposes.[43]Another approach to addressing the potential

---

[42] We found evidence of acquiescence bias when we identified 72 respondents that provided the same response for all positively- and negatively worded MSLSS items. We believed that their responses were implausible and therefore excluded them from our analyses.

[43] Although we used latent factors for the measurement of the dimensions of life satisfaction, the same concerns regarding the inconsistencies in item responses arise when one uses the subscale totals.

concern of acquiescence bias or response-set bias is to use only the positively worded items while reversing the response options for some of the questions (e.g., ranging from strongly disagree to strongly agree for some questions and from strongly agree to strongly disagree for other questions) (Barnette, 2000). This approach, however, was not tested in our study.

Our second recommendation relates to the use of the abridged version of the MSLSS. Although the abridged version of the MSLSS requires further validation in different samples, the findings provide preliminary support for its use for the measurement of the dimensions of life satisfaction in adolescents. We found good fit for a unidimensional structure for the abridged four-item version of each subscale. However, we also found that the adolescents may not have responded to the four items in the living environment subscale in a consistent manner. We therefore retained only two of the living environment items in our subsequent CFAs of a modified measurement structure with five correlated latent factors. Although we recommend replicating our findings in another sample before drawing further conclusions, we suggest that it may be necessary to develop new items for the measurement of adolescents' satisfaction with their living environment.

Our third recommendation is that researchers ought to be cautious in the use of the MSLSS as a measure of general life satisfaction whether they use total scores or a second-order factor. Even though a second-order factor structure may result in acceptable fit for the abridged version of the MSLSS, a comparison of the correlations among the dimensions of life satisfaction found in other studies suggested that the dimensions of life satisfaction may not correlate in a consistent manner and, therefore, may not respond in a consistent manner to a common second-order factor. This concern was further highlighted when we compared the second-order factor loadings that were found in three independent CFAs. Further studies are needed to examine whether the second-order factor loadings are indeed invariant with respect to observed and unobserved differences in adolescents before assuming that the five first-order factors reflect a second-order factor such that it provides a valid approach for the measurement of general life satisfaction.

4.2 Explaining global quality of life

In the second part of our study, we examined the degree to which the dimensions of life satisfaction explained global QOL. Instead of using the dimensions of life satisfaction for the measurement of general life satisfaction, we proposed that it would be more theoretically plausible to model the dimensions of life satisfaction as factors that may contribute to global QOL. To do so, we regressed global QOL on each of the dimensions of life satisfaction in a latent variable model (see Figure 20 on page 125). We considered this model to be more theoretically plausible because the dimensions of life satisfaction were not constrained to co-vary consistently with respect to a common latent factor. Indeed, why would one expect adolescents' satisfaction with their families, friends, schools, living environments, and selves to co-vary consistently? We found that our model, as specified, fit reasonably well, and that global QOL was explained predominantly by the adolescents' satisfaction with their selves and their families.

We were surprised that the other dimensions of life satisfaction barely contributed to global QOL after accounting for satisfaction with self and family. Our test of a subsequent model (see Figure 21 on page 127) confirmed that the relationships between the adolescents' satisfaction with their friends, schools, living environments, and their global QOL were almost entirely spurious with respect to their satisfaction with their selves and families. More specifically, the findings suggested that the relationship between the adolescents' satisfaction with their friends and their global QOL was predominantly explained by their satisfaction with their selves. In contrast, the relationships between the adolescents' global QOL and their evaluations of their experiences at school and in their living environment were predominantly explained by their satisfaction with their families.

We compared our results to those reported by other researchers who also examined the relationships between the dimensions of life satisfaction, as measured by the MSLSS, and global QOL. We identified 6 studies where global QOL or global life satisfaction, as measured by the Students' Life Satisfaction Scale (SLSS) (Huebner, 1991),[44] was regressed on each of the dimensions of life satisfaction. Of these studies, 4 also included the correlations between these variables, which enabled us to calculate Pratt index values even though these were not reported in the publications (see Table 29). The findings of the 7 samples in these 4 studies confirmed that global QOL was predominantly explained by satisfaction with self and family. However, we also observed variation in the relative importance of each of the dimensions of life satisfaction with respect to global QOL across these studies in comparison with our findings. In particular, whereas the relative importance of satisfaction with self was greatest in our study (Pratt index = 66%), it was not greatest in 5 of the 7 samples included in our review. For example, McCullough (2003) reported that the regression of global QOL on satisfaction with self was statistically non-significant in

[44] Although these researchers referred to the measurement of "global life satisfaction," their conceptualization and measurement of this concept is comparable to how we conceptualized and measured global QOL. To avoid confusion and remain consistent, we will use the terms global QOL to refer to an appraisal of one's life as a whole (globally).

Table 29 The relative importance of the life satisfaction dimensions found in other published studies and in our study

		MSLSS subscale															
		Family			Friends			School			Living environment			Self			
Sample ID	N	*β*	*r*	*d*	*β*	*r*	*d*	*β*	*r*	*d*	*β*	*r*	*d*	*β*	*r*	*d*	R^2
18	80	0.60*	0.64*	**87%**	0.20	0.36*	16%	-0.02	0.24*	-1%	-0.02	0.36*	-2%	0.00	0.30*	0%	0.44
19	111	0.45*	0.67*	**54%**	0.21*	0.51*	19%	-0.03	0.38*	-2%	0.26*	0.53*	25%	0.04	0.46*	3%	0.56
1	61	0.26*	0.74*	25%	0.26*	0.46*	16%	0.30*	0.59*	23%	0.33*	0.74*	**32%**	0.04	0.62*	3%	0.77
2	61	0.27*	0.49*	31%	0.14	0.46*	15%	0.03	0.21	2%	0.16	0.39*	15%	0.30*	0.54*	**38%**	0.43
13	49	0.59*	0.77*	**78%**	-0.12	0.22*	-5%	0.02	0.33*	1%	0.17	0.54*	16%	0.16	0.58*	16%	0.58
14	49	0.27*	0.45*	30%	0.29*	0.54*	39%	0.02	0.54	3%	-0.06	0.32*	-5%	0.37*	0.55*	**51%**	0.40
4	321	0.34*	0.58*	**45%**	0.04	0.32*	2%	0.07	0.33*	5%	0.20*	0.46	29%	0.23*	0.49*	26%	0.44
Our study**	6,163	0.26*	0.67*	26%	-0.01	0.51*	-1%	0.03	0.40*	2%	0.08*	0.57*	7%	058*	0.78*	**66%**	0.67

Notes: Sample ID corresponds to the descriptions provided in Appendix E. *B* = standardized regression coefficient for global QOL on the life satisfaction dimensions. *r* = *Pearson correlation.* *d* = Pratt index based on the coefficients found in the published studies (largest values are shown in bold). * indicates statistical significance at $p < 0.05$. ** see Table 21 on page 126 adolescents (grades 9 to 12) with learning disabilities and “normally achieving” adolescents (Pratt index = 0.0% and 3.3%, respectively). Ash and Huebner (1998) similarly reported that this relationship was statistically non-significant in 61 “academically gifted” students in grades 3 to 8 (corresponding Pratt index = 3.3%). In contrast, they found that this relationship was significant in a matched group of 61 so-called “non-gifted” students (corresponding Pratt index = 37.7%). Clearly, the relative importance of satisfaction with self varied significantly across samples drawn from populations of adolescent.

Similarly, the relative importance of satisfaction with family was greatest in 4 of the 7 samples with Pratt index values ranging from 45% to 87%. In contrast, we found that the relative importance of satisfaction with family was secondary to that of satisfaction with self, and much smaller (Pratt index = 26%). Thus, there appears to be substantial variation in the relative importance of satisfaction with family in samples drawn from different populations. Variations were also observed for the relative importance of other dimensions of life satisfaction in the different samples. Significant regression coefficients for satisfaction with friends were found in three samples (the corresponding Pratt index values ranged from 15.5% to 39.2%), whereas this coefficient was found to be statistically non-significant in the four other samples (the corresponding Pratt index values ranged from -4.6% to 16.4%). The regression coefficient for satisfaction with school was found to be not statistically significant and very small in 6 of 7 samples (the corresponding Pratt index values ranged from -2.0% to 5.3% for the non-significant effect and was 23.0% in the one sample where this effect was found to be significant). The Pratt index values for satisfaction with living environment ranged from -4.8% to 31.7% with the corresponding regression coefficients being statistically significant in 3 of the 7 samples.

Our comparison of regression parameters and Pratt index values found in different samples indicates that the dimensions of life satisfaction do not relate to global QOL in a consistent manner. These observations suggest that global QOL may take on different meanings for adolescents in different samples and from different populations. At this stage, it is difficult to determine what kind of differences might explain these variations in the relative importance of dimensions of life satisfaction. This is clearly an area that is worthy of further investigation.

4.3 The relationships between perceived physical and mental health status and global QOL

Our third analytical objective was to examine whether perceived mental health status, perceived physical health status, or both contributed to global QOL, and whether the dimensions of life satisfaction mediated the relationship(s). The results indicated that perceived mental and physical health status significantly explained global QOL. However, the total effect of perceived physical health status was significantly smaller than that of perceived mental health status. We also found that the relationships between global QOL and perceived mental and physical health status were predominantly mediated by the adolescents' satisfaction with self and family, and that satisfaction with self and family mediated the relationships between perceived mental and physical health status and the other dimensions of life satisfaction.

These results indicate that perceived mental and physical health status do not relate to each of the dimensions of life satisfaction to the same degree. For example, we found that perceived mental and physical health status predominantly explained satisfaction with self (R^2 = 33%) and, to a much lesser extent, satisfaction with family (R^2 = 17%), living environment (R^2 = 14%), friends (R^2 = 11%), and school (R^2 = 8%). These differences in the degree to which perceived mental and physical health status explained the dimensions of life satisfaction would, obviously, not have been found had we only focused on explaining global QOL. An important observation from these findings is that, although perceived mental and physical health status explained a significant percentage of the variance in global QOL, these explanatory relationships were partially explained by the explanatory relationships between perceived mental and physical health status and the adolescents' satisfaction with self and family.

In our view, the most important theoretical conclusion to be drawn from these findings is that perceived mental and physical health status and the dimensions of life satisfaction can be viewed as factors that contribute to global QOL in adolescents. These contributing factors were specified as sources of global QOL in our models. Our model was consistent with the following theoretical propositions: (a) the concepts that are commonly included in multidimensional quality of life instruments constitute conditions that have the potential to *affect* quality of life, rather than factors that *measure* quality of life and (b) quality of life can be conceptualized as a unidimensional concept that is partially explained by satisfaction with various areas of life (otherwise referred to as domain satisfactions) (Campbell et al., 1976). In this sense, we conceptualized quality of life as being distinct from the multitude of contributing conditions that may affect it (Nordenfelt, 1993). We emphasize that the theoretical relationships implied by these propositions are fundamentally different from those that are implied by the indirect reflective measurement structures that have frequently been used for the measurement of quality of life or general life satisfaction. Whereas an indirect reflective structure for the measurement of quality of life is based on the theoretical proposition that the various factors (e.g., health status and dimensions of life satisfaction) respond to quality of life in a consistent manner, the model with latent factors that explain quality of life is based on the theoretical proposition that that these factors may contribute to quality of life and that they therefore do not necessarily relate to quality of life in a consistent manner.

The proposition that various factors may contribute to quality of life is compatible with several philosophical and theoretical ideas about quality of life. For example, in his theoretical scheme of "the four qualities of life," Veenhoven (2000) distinguished between "outer qualities" that refer to resources for living well that are external to the person (e.g., social and environmental resources) and "inner qualities" that refer to resources that are internal to the person (e.g., physical and mental health status). Both the outer and inner qualities may contribute to a person's quality of life. Nordenfelt (1993) similarly discussed the conditions for quality of life in terms of external and inner welfare. He used the term "external welfare" to refer to those "phenomena which surround us and continuously affect us" (p. 35), and "inner welfare" to refer to "that combination of inner properties which lead to or positively affect our wellbeing" (p. 37). These ideas are consistent with the findings of Michalos (2001) and Campbell et al. (1976) that showed that quality of life can be partially explained by satisfaction with various areas of life (i.e., domain satisfaction, health status, etc.). Although our intent here is not to defend a particular theoretical perspective, the above examples illustrate that it is entirely plausible that quality of life can be conceived of as a unidimensional concept that can be distinguished from a multitude of potential factors that may be contribute to it (Beckie & Hayduk, 1997, 2004).

4.4 Methodological recommendations

We now turn to a discussion of some methodological recommendations that can be drawn from the results of our work and the methodological challenges that we encountered. We specifically discuss a few recommendations pertaining to the use of CFA and FMA for the examination of the reliability and validity of the MSLSS. Although some of these recommendations could apply more generally, researchers should determine the applicability of these recommendations in relation to the particular analytical objectives of their study.

4.4.1 Methodological recommendations for CFA of the MSLSS

The first recommendation relates to the assumptions of multivariate normality and interval-based measurement as applied to ordinal variables in CFA. Although we recognize that some ordinal variables may sufficiently approximate a continuous and normal distribution, there is no theoretical reason to believe that the six-point Likert scale for the MSLSS variables ought to result in approximately normal distributions with equal distances between each of the response options. It is therefore imperative to test these assumptions. When we compared the results of the CFAs to those that we would have obtained if we had ignored the deviations from multivariate normality, we found that the latter approach led to very misleading conclusions about model fit and the magnitude of the parameters in the model. We also found that the correlations among the observed variables were negatively biased if we were to assume that the scale for the MSLSS items was continuous. We therefore recommend that the MSLSS items, with six-point Likert scales, be treated as ordinal variables in future CFAs and SEMs.

The second recommendation relates to the use of item parceling in CFA. As discussed earlier, our concern is that item parceling is not congruent with the theoretical purpose of

CFA, which is to test the validity of a measurement structure pertaining to the relationships among observed variables and one or more latent factor(s) (e.g., Schumacker & Lomax, 2004; Viswanathan, 2005). The incongruence lies in the fact that the results of CFAs based on item parcels are only valid if the items in the parcels are unidimensional, whereas the purpose of CFA involves testing the dimensionality among the items. We found that CFAs based on item parcels can hide potentially important areas of misfit and that the results can therefore be misleading. We therefore concur with Little et al. (2002) who recommended that item parceling methods should not be used for the purpose of testing the factor structure of an instrument.

The third recommendation relates to the examination of fit in CFA. Of course, fit in CFA ultimately involves an assessment of the degree to which the correlations (or covariances) among observed variables, as implied by the statistical model, approximate the observed correlations. The chi-square statistic, and the related fit indices, can be used to assess the degree of approximation. However, we found that such global evaluations of model fit can hide potentially important areas of misfit. We therefore recommend that it is of fundamental importance to also carefully examine the residual correlation matrices when assessing model fit, and to provide some information about the patterns in the residual correlation matrix when reporting CFA results. For this purpose, we presented the range of residual correlations as well as the percentage of residual correlations with absolute values larger than 0.1 in our tables. We also mapped the residual correlations so as to reveal potential areas of misfit. The findings in our study demonstrated that these types of approaches can be effectively used for the purpose of reporting the fit of a model in an informative manner.

4.4.2 Using FMA to examine sample heterogeneity

We now turn to the use of FMA for the assessment of the degree to which a sample is homogenous with respect to a particular measurement structure. Both CFA and FMA can be used to test claims about the validity of an instrument. However, whereas CFA relates to the examination of exchangeability between items (construct validity), FMA relates to the examination of exchangeability between people. In other words, CFA is an approach for testing the construct validity of an instrument in a particular sample. In the Draper-Lindely-de Finetti (DLD) framework of measurement validity (Zumbo, 2007), this is referred to as validity that pertains to “domain specific inferences” (p. 59). These types of measurement inferences are not necessarily generalizable beyond the particular sample in which they were tested. Obviously, the ideal is to find support for validating claims that can be applied to different samples. In the DLD framework, this is referred to as “general measurement inference” (p. 59).

One approach to examine the validity of a measurement structure in different samples is to use multi-group CFA to examine whether the structural relations are the same in different observed groups that are represented in the sample. However, with respect to the measurement of quality of life, it is plausible that there are various unobserved groups that may differ in how their members interpret and respond to questions about their quality of life. In these situations, FMA can be used to examine the degree to which a large sample of respondents with diverse characteristics is homogenous with respect to the particular measurement structure that is being tested (Lubke & Muthén, 2005). Obviously, the conclusions from such an analysis can only apply to the observed or unobserved differences that were represented in a particular sample. Nevertheless, the validity claims that could be drawn from such an analysis are a step beyond the claims that could be drawn if one were to only rely on the results of a CFA in a particular sample.

Based on the results of our analyses, we suggest that FMA can be a very useful addition to the toolbox of psychometric methods that can be used for the testing of validity claims pertaining to the measurement of life satisfaction or quality of life, particularly in large samples of individuals who may not have interpreted or responded to the items of an instrument in a consistent manner. However, we also caution that this exploratory approach essentially involves dividing the sample in such a way as to maximize the differences in model parameters across two or more latent classes. It is therefore not unlikely that such differences will indeed be found in almost any large sample of diverse individuals. With respect to the validity of the latent classes, it is therefore imperative to replicate these finding in other samples. In addition, one should seek to identify those differences that explain latent class membership if the purpose is to draw conclusions pertaining to the nature of the latent classes found.

This, however, was not our purpose. Instead, we used FMA to identify potentially unreliable items by comparing the structural parameters of those items across the latent classes that were identified in our analyses. In doing so, we were able to identify those items to which the adolescents responded in the most consistent manner. We used this approach to develop an abridged version of the MSLSS, which resulted in a remarkable improvement in model fit.

4.5 Limitations

It is conventional to alert the reader to some of the limitations that should be kept in mind when considering the results of a study. Several limitations have already been alluded to throughout our discussion. Here, we provide a brief summary of what we believe to be the most salient limitations.

We already discussed the issue of missing data. Although the approaches that we used to address this issue (i.e., single imputation using the EM algorithm, multiple imputation, and full information maximum likelihood) have less restrictive assumptions than some of the other approaches that we could have used (e.g., listwise deletion), the fundamental assumption remains that data were missing at random (MAR) after controlling for all the covariates that were included in our model. This assumption is much less restrictive than the assumption of data that are missing completely at random (MCAR). However, the MAR assumption cannot be formally tested and thus we ultimately do not know the degree to which the assumption was justified. Nevertheless, the consistency in results across the different missing data techniques applied suggests that the missing data mechanisms that may have been operative were unlikely to have substantially influenced our findings.

Another limitation in our study was that the analyses were entirely based on cross-sectional data. This limitation would be of significant concern if we sought to explicitly prove causality in the sense that a change in, for example, mental health status in a particular individual would lead to a corresponding change in, for example, global QOL in that same individual. This, however, was not our purpose. Rather, our purpose was to test the assumptions of a particular model for the measurement of the dimensions of life satisfaction, and to examine the explanatory relationships among these dimensions and perceived mental and physical health status and global QOL. In other words, although the findings were consistent with the relationships that were specified in the models, they did not prove these relationships to be causal in the sense described above. We therefore recommend additional studies to validate the inferences pertaining to the relationships among perceived mental and physical health status, the dimensions of life satisfaction, and global QOL.

Finally, related to the above limitation, we caution that the abridged version of the MSLSS needs to be validated in other studies. Although we are confident that the abridged version of the MSLSS will probably result in greater reliability and validity pertaining to inferences about the dimensions of life satisfaction in samples that are similar to the sample in this study, this proposition needs to be established.

4.6 Conclusions

The general purpose of our analyses was to test several assumptions underlying the use of indirect reflective models for the measurement of quality of life in health research, and to propose an alternative model for examining the relationships among satisfaction with various domains of life, global QOL, and physical and mental health status (see Figure 2). We used the MSLSS, two global quality of life measures, and two measures of perceived physical and mental health status to examine these relationships in adolescents. The results revealed several concerns pertaining to the measurement of adolescents' satisfaction with life generally and various life domains including: (a) adolescents may not respond in a consistent manner to questions about their satisfaction with various life domains and (b) combining domain satisfaction scores in a second-order factor model may not be a valid approach for the measurement of adolescents' general life satisfaction. Factor mixture analysis can be used effectively to identify items with inconsistent response patterns. We used this method to develop an abridged 18-item version of the MSLSS for the measurement of adolescents' satisfaction with their family, friends, school, living environment, and their perception of self. However, only two of the living environment items were retained, and we therefore suggest that it may be necessary to develop additional items for this subscale.

We caution that population health researchers who use instruments for the measurement of quality of life must critically examine whether individuals respond to the items in a consistent manner. Although we specifically observed inconsistent item response patterns in adolescents, we caution that similar concerns may arise in adult populations. Adults may differentially interpret questions about various life domains because of differences in age, culture or language, because of contextual differences such as different living environments, or because of different life experiences resulting from mental or physical health challenges or other challenging life circumstances. Health researchers must therefore test the assumption of consistent item responses to justify the generalizability of their inferences to the target population.

We also suggest that an indirect reflective measurement model may not be the best approach to combining measures of various health outcomes and life domains into a total score for the measurement of general life satisfaction or quality of life. Health researchers often derive a total quality of life score by combining measures of physical, mental, and social functioning, perceived health status, satisfaction with various domains of life, overall quality of life, and happiness. Factor analysis is generally used to validate this approach to quality of life measurement. However, the theoretical proposition that measures pertaining to so-called life domains arise from a common source is not plausible. Rather, we recommend that the measures of various life domains that are typically included in quality of life instruments should be viewed as factors that *contribute* to quality of life, and that future studies should focus on examining theoretically plausible relationships between the life domains rather than assuming that they constitute reflective measures of general quality of life.

The abridged MSLSS was used to examine whether perceived mental and physical health status significantly explained adolescents' global QOL and whether these relationships were mediated by adolescents' satisfaction with five important life domains. We found that mental health status and, to a lesser extent, physical health status were associated with significant differences in the adolescents' appraisals of their family, friends, living environment, school, self, and their global QOL. Global QOL was predominantly explained by mental health status (d = 30%) and by adolescents' satisfaction with self (d = 42%) and family (d = 20%). And, satisfaction with self and family were also the predominant mediating variables for the relationships between mental health status (45% total mediation) and physical health status (68% total mediation) and global QOL. These findings warrant more attention to questions pertaining to these important life domains in health assessments and in population health research so as to target appropriate supportive services for adolescents, particularly those with mental or physical health challenges.

Quality of life clearly has become an important consideration in health care. Knowledge about the impact of illness, disease, and medical or other therapeutic interventions on various aspects of people’s lives provides an essential theoretical foundation for the practice of nurses and other health-care professionals. This knowledge is used to inform health promotion activities, public health policy, and to develop appropriate educational and supportive services for people with chronic health challenges. However, theoretical developments pertaining to the relationship between health and quality of life obviously are contingent upon reliable and valid measures of people’s experiences in various life domains. Researchers cannot assume that people respond to these measures in a consistent manner without first carefully testing this assumption. Nor can they assume the validity of combining measures pertaining to diverse life domains into a total score for the measurement of quality of life. Further studies are needed to substantiate theoretically defensible propositions about the relationships among various life domains, health status, and global QOL so as to inform the practice of nurses and other health-care professionals.

REFERENCES

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.

Andrews, F. M., & Robinson, J. P. (1991). Measures of subjective well-being. In J. P. Robinson, P. R. Shaver, L. S. Wrightsman, F. M. Andrews & J. P. Robinson (Eds.), *Measures of personality and social psychological attitudes* (pp. xiv, 753 p.). San Diego: Academic Press.

Ash, C., & Huebner, E. S. (1998). Life satisfaction reports of gifted middle-school children. *School Psychology Quarterly, 13*, 310-321.

Asparouhov, T., & Muthén, B. (2006). Robust chi-Square difference testing with mean and variance adjusted test statistics. *Mplus Web Notes: No.10.* Retrieved February 13, 2007, from http://www.statmodel.com/download/webnotes/webnote10.pdf

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling, 9*, 78-102.

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*, 361-370.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13*, 186-203.

Beckie, T. M., & Hayduk, L. A. (1997). Measuring quality of life. *Social Indicators Research, 42*, 21-37.

Beckie, T. M., & Hayduk, L. A. (2004). Using perceived health to test the construct-related validity of global quality of life. *Social Indicators Research, 65*, 279-298.

Bergner, M., Bobbitt, R. A., Carter, W. B., & Gilson, B. S. (1981). The Sickness Impact Profile: Development and final revision of a health status measure. *Medical Care, 19*, 787-805.

Bircher, J. (2005). Towards a dynamic definition of health and disease. *Medicine, Health Care, and Philosophy, 8*, 335-341.

Blalock, H. M. (1974). *Measurement in the social sciences: Theories and strategies*. Chicago: Aldine.

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*, 305-314.

Borgers, N., Hox, J., & Sikkel, D. (2004). Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral mid-point. *Quality & Quantity, 38*, 17-33.

Borooah, V. K. (2002). *Logit and probit: Ordered and multinomial models*. Thousand Oaks, CA: Sage Publications.

Bowling, A. (2002). *Research methods in health: Investigating health and health services* (2nd ed.). Buckingham, Philadelphia: Open University Press.

Bowling, A. (2005). *Measuring health: A review of quality of life measurement scales* (3rd ed.). Maidenhead, England: Open University Press.

Bradford, R., Rutherford, D. L., & John, A. (2002). Quality of life in young people: Ratings and factor structure of the Quality of Life Profile-Adolescent Version. *Journal of Adolescence, 25*, 261-274.

Brantley, A., Huebner, E. S., & Nagle, R. J. (2002). Multidimensional life satisfaction reports of adolescents with mild mental disabilities. *Mental Retardation, 40*, 321-329.

Burckhardt, C. S. (1985). The impact of arthritis on quality of life. *Nursing Research, 34,* 11-16.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: L. Erlbaum.

Campbell, A., Converse, P., & Rodgers, W. L. (1976). *The quality of American life*. New York: Sage.

Cantril, H. (1966). *Pattern of human concerns*. Piscataway, NJ: Rutgers University Press.

Chan, G. W.-Y., McBride-Chang, C., Leung, P. W. L., Tsoi, K. W., Ho, C. S.-H., & Cheuk, C. S. M. (2003). Factors influencing anxiety/depression among adolescents migrating between Chinese societies: The case of adolescent new arrivals to Hong Kong. *Journal of Psychology in Chinese Societies, 4*, 121-140.

Chang, L., McBride-Chang, C., Stewart, S. M., & Au, E. (2003). Life satisfaction, self-concept, and family relations in Chinese adolescents and children. *International Journal of Behavioral Development, 27*, 182-189.

Chinn, P. L., Kramer, M. K., & Chin, P. L. (2004). *Integrated knowledge development in nursing* (6th ed.). St. Louis, MO: Mosby.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: L. Erlbaum.

Crandall, J. E. (1975). Scale for social interest. *Journal of Individual Psychology, 31*, 187-195.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64,* 391-418.

Dannerbeck, A., Casas, F., Sadurni, M., & Coenders, G. (Eds.). (2004). *Quality-of-life research on children and adolescents* (Vol. 23). Dordrecht, Netherlands: Kluwer.

Donaldson, G. (2005). Structural equation models for quality of life response shifts: Promises and pitfalls. *Quality of Life Research, 14*, 2345-2351.

Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing, 42,* 57-63.

Drotar, D. (Ed.). (1998). *Measuring health-related quality of life in children and adolescents: Implications for research and practice*. Mahwah, NJ: Lawrence Erlbaum.

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational Research Methods, 4*, 144-192.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5*, 155-174.

Edwards, T. C., Huebner, C. E., Connell, F. A., & Patrick, D. L. (2002). Adolescent quality of life, part I: Conceptual and measurement model. *Journal of Adolescence, 25*, 275-286.

Edwards, T. C., Patrick, D. L., & Topolski, T. D. (2003). Quality of life of adolescents with perceived disabilities. *Journal of Pediatric Psychology, 28,* 233-241.

Elkinton, J. R. (1966). Medicine and the quality of life. *Annals of Internal Medicine, 64,* 711-714.

Enders, C. K. (2006). Analyzing structural equation models with missing data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 313-342). Greenwich: Information Age Publishing.

Fabrigar, L. R., MacCallum, R. C., Wegener, D. T., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Fayers, P. (2004). Quality-of-life measurement in clinical trials: The impact of causal variables. *Journal of Biopharmaceutical Statistics, 14*, 155-176.

Fayers, P., & Hand, D. J. (2002). Causal variables, indicator variables and measurement scales: An example from quality of life. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 165*, 233-253.

Fayers, P., & Machin, D. (2007). *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes*. Chichester, West Sussex, England: John Wiley & Sons.

Feinstein, A. R. (1987). Clinimetric perspectives. *Journal of Chronic Diseases, 40*, 635-640.

Ferrans, C. E. (2005). Definitions and conceptual models of quality of life. In J. Lipscomb, C. C. Gotay & C. Snyder (Eds.), *Outcomes assessment in cancer: Measures, methods, and applications* (pp. 14-30). Cambridge, NY: Cambridge University Press.

Ferrell, B. R., Dow, K. H., Leigh, S., Ly, J., & Gulasekaram, P. (1995). Quality of life in long-term cancer survivors. *Oncology Nursing Forum, 22*, 915-922.

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269-314). Greenwich: Information Age Publishing.

Fletcher, R., & Hattie, J. (2005). Gender differences in physical self-concept: A multidimensional differential item functioning analysis. *Educational and Psychological Measurement, 65*, 657-667.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*, 466-491.

Gilligan, T. D., & Huebner, E. S. (2002). Multidimensional life satisfaction reports of adolescents: A multitrait-multimethod study. *Personality & Individual Differences, 32*, 1149-1155.

Gilman, R. (1999). *Validation of the Multidimensional Students' Life Satisfaction Scale with adolescents.* Dissertation Abstracts International, 60(04), 1901B.

Gilman, R. (2001). The relationship between life satisfaction, social interest, and frequency of extracurricular activities among adolescent students. *Journal of Youth & Adolescence, 30*, 749-767.

Gilman, R., & Ashby, J. S. (2003). A first study of perfectionism and multidimensional life satisfaction among adolescents. *Journal of Early Adolescence, 23*, 218-235.

Gilman, R., Ashby, J. S., Sverko, D., Florell, D., & Varjas, K. (2005). The relationship between perfectionism and multidimensional life satisfaction among Croatian and American youth. *Personality and Individual Differences, 39*, 155-166.

Gilman, R., Easterbrooks, S. R., & Frey, M. (2004). A preliminary study of multidimensional life satisfaction among deaf/hard of hearing youth across environmental settings. *Social Indicators Research, 66*, 143-164.

Gilman, R., & Handwerk, M. L. (2001). Changes in life satisfaction as a function of stay in a residential setting. *Residential Treatment for Children & Youth, 18*(4), 47-65.

Gilman, R., Huebner, E. S., & Laughlin, J. E. (2000). A first study of the Multidimensional Students' Life Satisfaction Scale with adolescents. *Social Indicators Research, 52*, 135-160.

Grant, M. M., & Dean, G. E. (2003). Evolution of quality of life in oncology and oncology nursing. In C. R. King & P. S. Hinds (Eds.), *Quality of life from nursing and patient perspectives: Theory, research, practice* (2nd ed., pp. 3-27). Sudbury, MA: Jones and Bartlett.

Greenspoon, P. J., & Saklofske, D. H. (1997). Validity and reliability of the multidimensional students' life satisfaction scale with Canadian children. *Journal of Psychoeducational Assessment, 15*, 138-155.

Greenspoon, P. J., & Saklofske, D. H. (1998). Confirmatory factor analysis of the Multidimensional Students' Life Satisfaction Scale. *Personality and Individual Differences, 25*, 965-971.

Greenspoon, P. J., & Saklofske, D. H. (2001). Toward an integration of subjective well-being and psychopathology. *Social Indicators Research, 54*, 81-108.

Griffin, M. D., & Huebner, E. S. (2000). Multidimensional life satisfaction reports of middle school students with serious emotional disturbance. *Journal of Psychoeducational Assessment, 18*, 111-124.

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis*. Cambridge, NY: Cambridge University Press.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: The Guilford Press.

Hershberger, S. L. (2006). The problem of equivalent structural models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 13-43). Greenwich: Information Age Publishing.

Hinds, P. S., & Haase, J. (2003). Quality of life in children and adolescents with cancer. In C. R. King & P. S. Hinds (Eds.), *Quality of life from nursing and patient perspectives: Theory, research, practice* (2nd ed., pp. xxvii, 533 p.). Sudbury, MA: Jones and Bartlett.

Horan, P. M., DiStefano, C., & Motl, R. W. (2003). Wording effects in self-esteem scales: Methodological artifact or response style? *Structural Equation Modeling, 10*, 435-445.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Huebner, E. S. (1991). Initial development of the Student's Life Satisfaction Scale. *School Psychology International, 12*, 231-240.

Huebner, E. S. (1994). Preliminary development and validation of a multidimensional life satisfaction scale for children. *Psychological Assessment, 6*, 149-158.

Huebner, E. S. (2001). *Manual for the Multidimensional Students' life Satisfaction Scale*. Columbia: University of South Carolina. Document Number)

Huebner, E. S. (2004). Research on assessment of life satisfaction of children and adolescents. *Social Indicators Research, 66*, 3-33.

Huebner, E. S. (Ed.). (1997). *Life satisfaction and happiness*. Silver Springs, MD: National Association of School Psychologists.

Huebner, E. S., Brantley, A., Nagle, R. J., & Valois, R. F. (2002). Correspondence between parent and adolescent ratings of life satisfaction for adolescents with and without mental disabilities. *Journal of Psychoeducational Assessment, 20*, 20-29.

Huebner, E. S., & Gilman, R. (2002). An introduction to the Multidimensional Students' Life Satisfaction Scale. *Social Indicators Research, 60*, 115-122.

Huebner, E. S., Laughlin, J. E., Ash, C., & Gilman, R. (1998). Further validation of the Multidimensional Students' Life Satisfaction Scale. *Journal of Psychoeducational Assessment, 16*, 118-134.

Huebner, E. S., Nagle, R. J., & Suldo, S. (2003). Quality of life assessment in child and adolescent health care: The Multidimensional Students' Life Satisfaction Scale (MSLSS). *Social Indicators Research Series: Advances in Quality-of-Life Theory and Research, 20*, 179-189.

Huebner, E. S., Valois, R. F., Suldo, S. M., Smith, L. C., McKnight, C. G., Seligson, J. L., et al. (2004). Perceived quality of life: A neglected component of adolescent health assessment and intervention. *Journal of Adolescent Health, 34*, 270-278.

Johnson, M., & North American Nursing Diagnosis Association. (2001). *Nursing diagnoses, outcomes, & interventions: NANDA, NOC, and NIC linkages*. Toronto: Mosby.

Johnson, V. E., & Albert, J. H. (2004). Ordinal regression models. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 151-174). Thousand Oaks, CA: Sage.

Kaplan, R. M. (1998). Implication of quality of life assessment in public policy for adolescent health. In D. Drotar (Ed.), *Measuring health-related quality of life in children and adolescents* (pp. 63-84). Mahwah, NJ: Lawrence Erlbaum.

King, C. R., & Hinds, P. S. (2003). *Quality of life from nursing and patient perspectives: Theory, research, practice* (2nd ed.). Sudbury, MA: Jones and Bartlett.

Koot, H. M., & Wallander, J. L. (2001). *Quality of life in child and adolescent illness: Concepts, methods and findings*. Hove, East Sussex: Brunner-Routledge.

Leung, C. Y. W., McBride-Chang, C., & Lai, B. P. Y. (2004). Relations among maternal parenting style, academic competence, and life satisfaction in Chinese early adolescents. *Journal of Early Adolescence, 24,* 113-143.

Lipscomb, J., Gotay, C. C., & Snyder, C. (2005). *Outcomes assessment in cancer: Measures, methods, and applications*. Cambridge: Cambridge University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9,* 151.

Long, P. H. (1960). On the quantity and quality of life. *Medical Times, 88,* 613-619.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*, 21-39.

Mackinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17,* 144-158.

Mackinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology, 58*, 593-614.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. xiii, 511). Thousand Oaks, CA: Sage.

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology, 22*, 37-49.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*, 810-819.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341.

Mccullough, G. C. (2003). *Life satisfaction of adolescents with learning disabilities.* Dissertation Abstracts International, 63(07), 3480B.

McCutcheon, A. L. (2002). Basic concepts and procedures in single- and multiple-group latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 56-84). Cambridge: Cambridge University Press.

Medved'ova, L. u. (1999). Je vyberova skola predpokladom pre lepsie zvladanie stresu? *Psychologia a Patopsychologia Dietata, 34*, 202-214. Abstract retrieved March 22, 2007, from PubMed database.

Michalos, A. C. (2004). Social indicators research and health-related quality of life research. *Social Indicators Research, 65*, 27.

Michalos, A. C., Hubley, A. M., Zumbo, B. D., & Hemingway, D. (2001). Health and other aspects of the quality of life of older people. *Social Indicators Research*, *54*, 239-274.

Michalos, A. C., Thommasen, H. V., Read, R., Anderson, N., & Zumbo, B. D. (2005). Determinants of health and the quality of life in the Bella Coola Valley. *Social Indicators Research, 72*, 1-50.

Michalos, A. C., Zumbo, B., D. , & Hubley, A. (2000). Health and the quality of life. *Social Indicators Research, 51*, 245-286.

Michalos, A. C., & Zumbo, B. D. (2002). Healthy days, health satisfaction and satisfaction with the overall quality of life. *Social Indicators Research*, *59*, 321-338.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, *39*, 479.

Moors, G. (2003). Diagnosing response style behavior by means of a latent-class factor approach. Socio-demographic correlates of gender role attitudes and perceptions of ethnic discrimination reexamined. *Quality and Quantity*, *37*, 277-302.

Mulaik, S. A. (2004). Objectivity in science and structural equation modeling. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 425-446). Thousand Oaks, CA: Sage.

Musschenga, A. W. (1997). The relation between concepts of quality-of-life, health and happiness. *Journal of Medicine and Philosophy, 22*, 11-28.

Muthén, B. (2004). Mplus technical appendices. Version 3. Retrieved February 13, 2007, from http://www.statmodel.com/download/techappen.pdf

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadraic estimating equations in latent variable modeling with categorical and continuous outcomes. Retrieved September 24, 2006, from http://www.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf

Muthén, B., & Muthén, L. K. (2006). MPlus (Version 4.2). Los Angeles: Statmodel.

Muthén, L. K., & Muthén, B. (2006). *Mplus user's guide* (4.1 ed.). Los Angeles: Muthen & Muthen.

Nasser, F. (2003). The effect of using item parcels on ad hoc goodness-of-fit indexes in confirmatory factor analysis: An example using Sarason's Reactions to Tests. *Applied Measurement in Education, Vol 16*, 75-97.

National Center for Biotechnology Information. (2005). PubMed. Retrieved October 8, 2005, from http://www.pubmed.com/

National Library of Medicine. (2005). Medical subject headings. Retrieved October 4, 2005, from http://www.nlm.nih.gov/mesh/meshhome.html

Neugarten, B. L., Havigurst, R. J., & Tobin, S. S. (1961). The measurement of life satisfaction. *Journal of Gerontology, 24*, 465-469.

Nickerson, A. B., & Nagle, R. J. (2004). The influence of parent and peer attachments on life satisfaction in middle childhood and early adolescence. *Social Indicators Research, 66*, 35-60.

Nordenfelt, L. (1993). *Quality of life, health and happiness*. Aldershot, England: Avebury.

Nylund, K. L., Asparouhov, T., & Muthén, B. (2006, August 31, 2006). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Manuscript submitted for publication. Retrieved August 31, 2006, from http://www.statmodel.com/download/LCA_tech11_nylund_v83.pdf

Padilla, G. V., Frank-Stromborg, M., & Setsuko, K. (2004). Single instruments for measuring quality of life. In M. Frank-Stromborg & S. J. Olsen (Eds.), *Instruments for clinical health-care research* (3rd ed., pp. 128-149). Sudbury, MA: Jones and Bartlett.

Padilla, G. V., & Grant, M. M. (1985). Quality of life as a cancer nursing outcome variable. *Advances in Nursing Science, 8*, 45-60.

Park, N. S. (2000). *Life satisfaction of school age children: Cross-cultural and cross-developmental comparisons.* Dissertation Abstracts International, 62(02), 1118B.

Park, N. S. (2005). Life satisfaction among Korean children and youth: A developmental perspective. *School Psychology International, 26*, 209-223.

Park, N. S., Huebner, E. S., Laughlin, J. E., Valois, R. F., & Gilman, R. (2004). A cross-cultural comparison of the dimensions of child and adolescent life satisfaction reports. *Social Indicators Research, 66*, 61-79.

Patrick, D. L., & Chiang, Y. P. (2000). Measurement of health outcomes in treatment effectiveness evaluations: Conceptual and methodological challenges. *Medical Care, 38*(9 Suppl), II14-25.

Patrick, D. L., Edwards, T. C., & Topolski, T. D. (2002). Adolescent quality of life, part II: Initial validation of a new instrument. *Journal of Adolescence, 25*, 287-300.

Polit, D. F., & Beck, C. T. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia: Lippincott Williams & Wilkins.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385-401.

Raphael, D. (1996). Determinants of health of North-American adolescents: Evolving definitions, recent findings, and proposed research agenda. *Journal of Adolescent Health, 19*, 6-16.

Raphael, D., Brown, I., Rukholm, E., & Hill-Bailey, P. (1996). Adolescent health: Moving from prevention to promotion through a quality of life approach. *Canadian Journal of Public Health, 87*, 81-83.

Rapley, M. (2003). *Quality of life research: A critical introduction*. Thousand Oaks, CA: Sage.

Richardson, C. G., Johnson, J. L., Ratner, P. A., Zumbo, B. D., Bottorff, J. L., Shoveller, J. A., et al. (in press). Validation of the dimensions of tobacco dependence scale for adolescents. *Addictive Behaviors*.

Rosenberg, M. (1989). *Society and the adolescent self-image* (Rev. ed.). Middletown, CT: Wesleyan University Press.

SAS Institute. (2005). Statistical Analysis Software (Version 9.1). Cary, NC: Author.

Schiff, M., Nebe, S., & Gilman, R. (2006). Life satisfaction among Israeli youth in residential treatment care. *British Journal of Social Work, 36*, 1325-1343.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Seligson, J. L., Huebner, E. S., & Valois, R. F. (2003). Preliminary validation of the Brief Multidimensional Students' Life Satisfaction Scale (BMSLSS). *Social Indicators Research, 61*, 121-145.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling* *multilevel, longitudinal, and structural equation models*. Retrieved August 30, 2006, from http://www.statsnetbase.com/ejournals/books/book_km.asp?id=1292

Smith, K. W., Avis, N. E., & Assmann, S. F. (1999). Distinguishing between quality of life and health status in quality of life research: A meta-analysis. *Quality of Life Research, 8*, 447-459.

Spieth, L. E., & Harris, C. V. (1996). Assessment of health-related quality of life in children and adolescents: An integrative review. *Journal of Pediatric Psychology, 21*, 175-193.

Spilker, B. (1996). *Quality of life and pharmacoeconomics in clinical trials* (2nd ed.). Philadelphia: Lippincott.

Statistics Canada. (2003). Ethnicity. Retrieved May 21, 2007, from http://www.statcan.ca/english/concepts/definitions/ethnicity.htm

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon.

Thomas, D. R., Hughes, E., & Zumbo, B. D. (1998). On variable importance in linear regression. *Social Indicators Research, 45*, 253-175.

Tian, L.-L., & Liu, W. (2005). Test of the Chinese version of the Multidimensional Students' Life Satisfaction Scale. *Chinese Mental Health Journal, 19*, 301-303. Abstract retrieved March 22, 2007, from PsycINFO Database.

Tomas, J. M., & Oliver, A. (1999). Rosenberg's Self-Esteem Scale: Two factors or method efects. *Structural Equation Modeling, 6*, 84.

Topolski, T. D., Edwards, T. C., & Patrick, D. L. (2004). Toward youth self-report of health and quality of life in population monitoring. *Ambulatory Pediatrics, 4*, 387-394.

Topolski, T. D., Patrick, D. L., Edwards, T. C., Huebner, C. E., Connell, F. A., & Mount, K. K. (2001). Quality of life and health-risk behaviors among adolescents. *Journal of Adolescent Health, 29*, 426-435.

Tu, A. W., Ratner, P. A., & Johnson, J. L. (in press). Gender differences in the correlates of adolescents' cannabis use. *Substance Use & Misuse*.

Vallerand, A. H., & Payne, J. K. (2003). Theories and conceptual models to guide quality of life research. In C. R. King & P. S. Hinds (Eds.), *Quality of life From nursing and patient perspectives: Theory, research, practice* (2nd ed., pp. 45-64). Sudbury, MA: Jones and Bartlett.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-69.

Veenhoven, R. (2000). The four qualities of life. *Journal of Happiness Studies, 1*, 1-39.

Vingilis, E. R., Wade, T. J., & Adlaf, E. (1998). What factors predict student self-rated physical health? *Journal of Adolescence, 21*, 83-97.

Vingilis, E. R., Wade, T. J., & Seeley, J. S. (2002). Predictors of adolescent self-rated health: Analysis of the National Population Health Survey. *Canadian Journal of Public Health, 93*, 193-197.

Viswanathan, M. (2005). *Measurement error and research design*. Thousand Oaks, CA: Sage.

Wade, T. J., Pevalin, D. J., & Vingilis, E. (2000). Revisiting student self-rated physical health. *Journal of Adolescence, 23*, 785-791.

Wade, T. J., & Vingilis, E. (1999). The development of self-rated health during adolescence: An exploration of inter- and intra-cohort effects. *Canadian Journal of Public Health, 90*, 90-94.

Wallander, J. L. (2001). Theoretical and developmental issues in quality of life for children and adolescents. In H. M. Koot & J. L. Wallander (Eds.), *Quality of life in child and adolescent illness: Concepts, methods and findings* (pp. 23-48). Hove, East Sussex: Brunner-Routledge.

Wallander, J. L., Schmitt, M., & Koot, H. M. (2001). Quality of life measurement in children and adolescents: Issues, instruments, and applications. *Journal of Clinical Psychology, 57*, 571-585.

Wang, J., Siegal, H. A., Falck, R. S., & Carlson, R. G. (2001). Factorial structure of Rosenberg's Self-Esteem Scale among crack-cocaine drug users. *Structural Equation Modeling, 8*, 275.

Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 health survey: Manual and interpretation guide*. Boston: The Health Institute, New England Medical Center.

Wilson, I. B., & Cleary, P. D. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Journal of the American Medical Association, 273*, 59-65.

World Health Association. (1946). *Official Records of the World Health Organization.* Retrieved March 15, 2007, from http://whqlibdoc.who.int/hist/official_records/2e.pdf

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.* Dissertation Abstracts International, 63(10), 3527B.

Zullig, K. J., Valois, R. F., Huebner, E. S., & Drane, J. W. (2005). Adolescent health-related quality of life and perceived satisfaction with life. *Quality of Life Research, 14*, 1573.

Zumbo, B. D. (2002). Investigating structure in the residuals: A method for studying non-fitting models. Unpublished working paper. Vancouver, BC: Edgeworth Laboratory at the University of British Columbia.

Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26: Psychometrics, pp. 45-79). Amsterdam: Elsevier Science.

APPENDICES

Appendix A UBC Research Ethics Board Certificate of Approval

The University of British Columbia
Office of Research Services and Administration
Behavioural Research Ethics Board

Certificate of Approval

PRINCIPAL INVESTIGATOR	DEPARTMENT	NUMBER
Johnson, J.L.	Nursing	**B03-0203**

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT

Public Schools ,

CO-INVESTIGATORS:

Bottorff, Joan, Nursing; Ratner, Pamela, Nursing; Sawatzky, Richard, Nursing; Shoveller, Jean Anne, Health Care/Epidemiology; Zumbo, Bruno, Educ & Couns Psych & Spec Educ

SPONSORING AGENCIES

Canadian Institutes of Health Research

TITLE :

Exploring Gender Differences in Tobacco Dependence Among Adolescents

APPROVAL RENEWED DATE	TERM (YEARS)
AUG - 2 2006	1

CERTIFICATION:

The request for continuing review of the above-named project has been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approved on behalf of the Behavioural Research Ethics Board
by one of the following:
Dr. Peter Suedfeld, Chair,
Dr. Susan Rowley, Associate Chair
Dr. Jim Rupert, Associate Chair
Dr. Arminee Kazanjian, Associate Chair

This Certificate of Approval is valid for the above term provided there is no change in the experimental procedures

Appendix B Sections of questionnaire including relevant analysis variables

E1: How old are you <u>today</u>?

I am ☐☐ **years old.**

E2: Are you male or female?

◯ Male ◯ Female

E3: What grade are you currently in?

◯ Grade 8 ◯ Grade 11

◯ Grade 9 ◯ Grade 12

◯ Grade 10 ◯ Other

If other, please describe: ____________________

E4: What language do you speak <u>most often</u> at the home you are living in? Please check only one response.

◯ English

◯ Other What language? ________________

E5: Do you speak other languages <u>at home</u> on a regular basis?

◯ Yes What language? ________________

◯ No

E6: Where were you born?

◯ Canada

◯ Other Specify country: ________________

E7: How long have you lived in Canada?

◯ I was born in Canada.

OR

◯ I have lived in Canada for

☐☐ years ☐☐ months.

E8: What is your 6 digit postal code? (i.e. V6T – 2B5)

☐☐☐ ☐☐☐

E9: How would you describe yourself? <u>Please mark all that apply</u>. (These categories come from the 2001 Census.)

◯ Aboriginal / First Nation (e.g., North American Indian, Metis, Eskimo)

◯ Arab

◯ Black (e.g., African, Haitian, Jamacian, Somali)

◯ Chinese

◯ Filipino

◯ Japanese

◯ Korean

◯ Latin American

◯ South East Asian (e.g., Cambodian, Indonesian, Vietnamese, Laotian)

◯ South Asian (e.g., East Indian, Pakistani, Punjabi, Sri Lankan)

◯ West Asian (e.g., Afghan, Iranian)

◯ White / Caucasian

◯ Other Specify:

____

E10: To which ethnic or cultural group(s) did your ancestors belong?

(e.g., Canadian, French, Chinese, English, Italian, Scottish, Polish, Jewish, Greek, Chilean, Micmac, Cree, etc.)

Please describe:

____

____

E11: Are you a member of an Indian Band/ First Nation?

◯ No

◯ Yes Specify the Band: ________________

____

Section F: Your Health

F1: How would you rate your physical health?

- ◯ excellent
- ◯ very good
- ◯ good
- ◯ fair
- ◯ poor

F5: How would you rate your emotional or mental health?

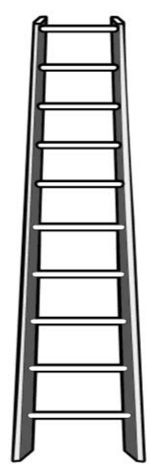
- ◯ excellent
- ◯ very good
- ◯ good
- ◯ fair
- ◯ poor

F14: Please indicate how strongly you agree or disagree with the following statements...

	strongly disagree	disagree	agree	strongly agree
f. I am satisfied with my quality of life.	◯	◯	◯	◯

Assume that this ladder is a way of picturing your life.The top of the ladder represents the best possible <u>life for you</u>. The bottom rung of the ladder represents the worst <u>possible life for you.</u>

F15: Indicate where on the ladder you feel you personally stand right now by marking the circle.



◯ **Best possible life**
◯
◯
◯
◯
◯
◯
◯ **Worst possible life**

F10: How often have you felt or behaved in the following manner in the past week (7 days)?

		rarely or none of the time (less than 1 day)	some or a little of the time (1-2 days)	occasionally or a moderate amount of the time (3-4 days)	most or all of the time (5-7 days)
a.	I was bothered by things that usually don't bother me.	○	○	○	○
b.	I felt I could not shake off the blues even with help from family or friends.	○	○	○	○
c.	I felt that I was just as good as other people.	○	○	○	○
d.	I felt depressed.	○	○	○	○
e.	I could not get going.	○	○	○	○
f.	I felt hopeful about the future.	○	○	○	○
g.	I thought my life had been a failure.	○	○	○	○
h.	I was happy.	○	○	○	○
i.	I felt lonely.	○	○	○	○
j.	I enjoyed life.	○	○	○	○
k.	I felt sad.	○	○	○	○
l.	I felt that people disliked me.	○	○	○	○

Section I: Life Satisfaction

We would like to know what thoughts about life you've had during the past two weeks. Think about how you spend each day and night and then think about how your life has been during most of this time. The following questions ask you to indicate your satisfaction with life. Circle the number (from 1 to 6) next to each statement that indicates the extent to which you agree or disagree with each statement. It is important to know what you REALLY think, so please answer the question the way you really feel, not how you think you should.

	strongly disagree	disagree	mildly disagree	mildly agree	agree	strongly agree
1. My friends are nice to me.	1	2	3	4	5	6
2. I am fun to be around.	1	2	3	4	5	6
3. I feel bad at school.	1	2	3	4	5	6
4. I have a bad time with my friends.	1	2	3	4	5	6
5. There are lots of things I can do well.	1	2	3	4	5	6
6. I learn a lot at school.	1	2	3	4	5	6
7. I like spending time with my parents.	1	2	3	4	5	6
8. My family is better than most.	1	2	3	4	5	6
9. There are many things about school I don't like.	1	2	3	4	5	6
10. I think I am good looking.	1	2	3	4	5	6
11. My friends are great.	1	2	3	4	5	6
12. My friends will help me if I need it.	1	2	3	4	5	6
13. I wish I didn't have to go to school.	1	2	3	4	5	6
14. I like myself.	1	2	3	4	5	6
15. There are lots of fun things to do where I live.	1	2	3	4	5	6
16. My friends treat me well.	1	2	3	4	5	6
17. Most people like me.	1	2	3	4	5	6
18. I enjoy being at home with my family.	1	2	3	4	5	6
19. My family gets along well together.	1	2	3	4	5	6
20. I look forward to going to school.	1	2	3	4	5	6
21. My parents treat me fairly.	1	2	3	4	5	6
22. I like being in school.	1	2	3	4	5	6
23. My friends are mean to me.	1	2	3	4	5	6
24. I wish I had different friends.	1	2	3	4	5	6
25. School is interesting.	1	2	3	4	5	6
26. I enjoy school activities.	1	2	3	4	5	6
27. I wish I lived in a different house.	1	2	3	4	5	6
28. Members of my family talk nicely to one another.	1	2	3	4	5	6

29. I have a lot of fun with my friends.	1	2	3	4	5	6
30. My parents and I do fun things together.	1	2	3	4	5	6
31. I like my neighbourhood.	1	2	3	4	5	6
32. I wish I lived somewhere else.	1	2	3	4	5	6
33. I am a nice person.	1	2	3	4	5	6
34. This town/city is filled with mean people.	1	2	3	4	5	6
35. I like to try new things.	1	2	3	4	5	6
36. My family's house is nice.	1	2	3	4	5	6
37. I like my neighbours.	1	2	3	4	5	6
38. I have enough friends.	1	2	3	4	5	6
39. I wish there were different people in my neighbourhood.	1	2	3	4	5	6
40. I like where I live.	1	2	3	4	5	6

Appendix C CFAs of MSLSS subscales based on multiple imputation

Model	WLSMV χ^2	# free parameters	CFI	RMSEA	SRMR
Friends (all items)[1]	3,740.69 – 3,861.30	9	0.895	0.174	0.057
Friends (no neg.)[2]	610.26 – 663.44	6	0.980	0.105	0.025
Friends (abridged)[3]	67.65 – 78.63	4	0.994	0.071	0.015
Living env. (all items)[1]	3,644.27 – 3,814.64	9	0.844	0.171	0.079
Living env. (no neg.)[2]	439.96 – 498.75	5	0.968	0.113	0.032
Living env. (abridged)[3]	31.86 – 40.32	4	0.997	0.049	0.012
School (all items)[1]	2,827.73 – 2,922.69	8	0.914	0.162	0.054
School (no neg.)[2]	1,287.92 – 1,353.67	5	0.957	0.213	0.042
School (abridged)[3]	3.63 – 8.46	4	1.000	0.017	0.004
Family[1]	2,933.99 – 3,042.71	7	0.912	0.192	0.048
Family (abridged)[3]	7.32 – 12.10	4	0.999	0.023	0.005
Self[1]	1,011.42 – 1,091.11	7	0.939	0.105	0.040
Self (abridged)[3]	11.18 – 18.80	4	0.999	0.031	0.007

Notes: Analyses based on multiple imputation for those who completed at least one MSLSS items (*N* = 7,305). The Chi-square and the fit indices are based on mean and variance adjusted weighted least squares (WLSMV). RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index. The estimated degrees of freedom based on WLSMV estimation varied slightly across the different multiple imputation samples. The number of free parameters is therefore provided instead.

[1] CFA based on all subscale items.

[2] CFA excluding negatively worded items.

[3] CFA of abridged subscales (4 items).

Appendix D MPlus 4.2 syntax and output

The following output pertains to the final mediation model using multiple imputation. Only a selection of the relevant output is included here.

```
Mplus VERSION 4.2
MUTHÉN & MUTHÉN
02/17/2007  11:57 AM

INPUT INSTRUCTIONS
      DATA: file is mi.txt;
      TYPE = IMPUTATION;

      !NOTE: PHYHLTH MENTHLTH ETHGRP6A grade5 lifepicc F14 MISSMSLSS
      !ARE ORIGINAL NOT-IMPUTED VARIABLES;
      !ALL OTHER VARIABLES ARE IMPUTED;
      !mi VARIABLE IS NUMBER OF IMPUTED DATASET;

      VARIABLE:  NAMES ARE
      mi id SEX PHYHLTH MENTHLTH F14SAT PEERCOMP SRESPCT1
      CESD12IT ETHGRP6A y1-y40
      missmslss grade5 ageb lifepicc phyhlthr menhlthr
      lifepic8 F14
      ;

      IDVARIABLE = id;
      MISSING ARE ALL(99999);
       USEVARIABLES ARE
       y2 y5 y6 y7
       y8 y11 y15 y16
       y18 y21 y23 y24
       Y27 Y32
       y35 y36 y37 y38
       lifepic8 F14SAT
       phyhlthr menhlthr;

       CATEGORICAL ARE    !ALL VARIABLES ARE ORDINAL;
       y2 y5 y6 y7
       y8 y11 y15 y16
       y18 y21 y23 y24
       Y27 Y32
       y35 y36 y37 y38
       lifepic8 F14SAT
       phyhlthr menhlthr;

      !SPECIFY MISSING DATA SUBSAMPLE;
      !ONLY INCLUDE THOSE WHO HAVE AT LEAST ONE VALUE
      !FOR THE MH OR PH, AND LIFEPICC OR F14SAT AND ONE OF THE MSLSS ITEMS.
      USEOBSERVATIONS ARE
      (MENTHLTH NE 99999 OR PHYHLTH NE 99999) AND
      (LIFEPICC NE 99999 OR F14 NE 99999) AND
      MISSMSLSS <=39;

      ANALYSIS: ESTIMATOR = WLSMV;
      PARAMETERIZATION = THETA;
```

```
MODEL:
   !MSLSS MEASUREMENT STRUCTURE;

   !CREATE LATENT VARIABLES FOR Y35 AND Y16;
       Fy35 by y35@1;
       y35@0;   !--> RESIDUAL OF Y35 = 0 IN THETA MATRIX;
       Fy16 by y16@1;
       y16@0;   !--> RESIDUAL OF Y16 = 0 IN THETA MATRIX;

       family by y2@1 y5 y6 y7;
       friends BY y8@1 y11 y15 Fy16 y38; !--> NOTE CROSS LOADING FOR Y38;
       living BY y27@1 y32;
       school BY y21@1 y18  y23 y24;
       self BY y36@1 Fy35 y37 y38;
       family;
       friends;
       school;
       living;
       self;

       y18 on Fy35;
       y24 on Fy35;

       y27 on Fy16;
       y32 on Fy16;

   !ALLOW LATENT MSLSS FACTORS TO CORRELATE;
       family with friends living school self;
       friends with living school self;
       living with school self;
       school with self;

   !OVERALL QOL MEASUREMENT STRUCTURE;
       QOL BY lifepic8@1 F14SAT;
       QOL;

   !LATENT VARIABLES FOR MH AND PH;
       MH BY menhlthr@1;
       MH@1;
       menhlthr@0;   !--> RESIDUAL = 0 IN THETA;
       PH BY phyhlthr@1;
       PH@1;
       phyhlthr@0;   !--> RESIDUAL = 0 IN THETA;

   !LATENT FACTOR TO ACCOUNT FOR CORRELATION MH WITH PH;
   !SO THAT MH AND PH BECOME ENDOGENOUS;
       f by MH@1 PH;
       f@1;

   !REGRESSION MODEL;
            QOL on family friends living school self;
            QOL on MH PH;
            family on MH PH;
            friends on MH PH;
            living on MH PH;
            school on MH PH;
            self on MH PH;
```

```
OUTPUT: SAMPSTAT STANDARDIZED TECH1;

SUMMARY OF ANALYSIS

Number of groups                                                 1
Average number of observations                                6932

Number of replications
    Requested                                                   10
    Completed                                                   10

Number of dependent variables                                   22
Number of independent variables                                  0
Number of continuous latent variables                           11

Observed dependent variables

  Binary and ordered categorical (ordinal)
   Y2          Y5          Y6          Y7          Y8          Y11
   Y15         Y16         Y18         Y21         Y23         Y24
   Y27         Y32         Y35         Y36         Y37         Y38
   LIFEPIC8    F14SAT      PHYHLTHR    MENHLTHR

Continuous latent variables
   FY35        FY16        FAMILY      FRIENDS     LIVING      SCHOOL
   SELF        QOL         MH          PH          F

Variables with special functions

  ID variable           ID

Estimator                                                    WLSMV
Maximum number of iterations                                  1000
Convergence criterion                                    0.500D-04
Maximum number of steepest descent iterations                   20
Parameterization                                             THETA

Input data file(s)
  Multiple data files from
    mi.txt
Input data format  FREE


           CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)
              Y2            Y5            Y6            Y7            Y8
              ________      ________      ________      ________      ________
 Y2
 Y5             0.497
 Y6             0.465         0.560
 Y7             0.526         0.602         0.597
 Y8             0.245         0.304         0.299         0.305
 Y11            0.250         0.308         0.262         0.290         0.614
 Y15            0.240         0.336         0.342         0.388         0.595
 Y16            0.211         0.261         0.273         0.284         0.411
 Y18            0.312         0.327         0.273         0.351         0.273
 Y21            0.213         0.280         0.256         0.309         0.194
 Y23            0.230         0.287         0.245         0.303         0.157
 Y24            0.226         0.282         0.255         0.347         0.195
 Y27            0.326         0.404         0.388         0.513         0.298
```

Y32	0.366	0.409	0.383	0.478	0.304
Y35	0.340	0.329	0.291	0.363	0.349
Y36	0.281	0.211	0.216	0.234	0.281
Y37	0.354	0.409	0.383	0.403	0.415
Y38	0.299	0.345	0.328	0.354	0.485
LIFEPIC8	0.355	0.370	0.337	0.394	0.292
F14SAT	0.335	0.382	0.353	0.386	0.323
PHYHLTHR	0.200	0.188	0.189	0.229	0.195
MENHLTHR	0.275	0.310	0.298	0.316	0.274

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	Y11	Y15	Y16	Y18	Y21
Y15	0.637				
Y16	0.411	0.482			
Y18	0.263	0.234	0.185		
Y21	0.170	0.168	0.147	0.473	
Y23	0.145	0.146	0.110	0.593	0.641
Y24	0.201	0.231	0.157	0.478	0.526
Y27	0.285	0.366	0.315	0.321	0.273
Y32	0.282	0.340	0.360	0.319	0.273
Y35	0.315	0.331	0.298	0.449	0.245
Y36	0.252	0.234	0.241	0.208	0.197
Y37	0.387	0.415	0.357	0.329	0.260
Y38	0.469	0.502	0.419	0.293	0.276
LIFEPIC8	0.233	0.268	0.269	0.299	0.249
F14SAT	0.284	0.311	0.290	0.307	0.235
PHYHLTHR	0.146	0.181	0.174	0.202	0.181
MENHLTHR	0.190	0.254	0.246	0.267	0.223

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	Y23	Y24	Y27	Y32	Y35
Y24	0.636				
Y27	0.282	0.323			
Y32	0.295	0.300	0.638		
Y35	0.265	0.326	0.306	0.345	
Y36	0.145	0.169	0.221	0.231	0.442
Y37	0.252	0.288	0.373	0.406	0.561
Y38	0.212	0.287	0.358	0.362	0.462
LIFEPIC8	0.223	0.258	0.319	0.351	0.400
F14SAT	0.238	0.256	0.339	0.382	0.416
PHYHLTHR	0.151	0.249	0.205	0.229	0.365
MENHLTHR	0.197	0.252	0.289	0.327	0.396

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	Y36	Y37	Y38	LIFEPIC8	F14SAT
Y37	0.578				
Y38	0.453	0.547			
LIFEPIC8	0.350	0.508	0.367		
F14SAT	0.309	0.528	0.349	0.553	
PHYHLTHR	0.295	0.319	0.296	0.394	0.333
MENHLTHR	0.297	0.479	0.355	0.532	0.502

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	PHYHLTHR	MENHLTHR
MENHLTHR	0.536	

TESTS OF MODEL FIT

Number of Free Parameters 53

Chi-Square Test of Model Fit

Number of successful computations 10

Proportions		Percentiles	
Expected	Observed	Expected	Observed
0.990	1.000	86.074	2010.017
0.980	1.000	89.500	2010.017
0.950	1.000	94.811	2010.017
0.900	1.000	99.707	2010.017
0.800	1.000	105.860	2010.017
0.700	1.000	110.453	2052.339
0.500	1.000	118.334	2056.786
0.300	1.000	126.582	2082.135
0.200	1.000	131.752	2082.858
0.100	1.000	139.149	2083.223
0.050	1.000	145.461	2083.223
0.020	1.000	152.785	2083.223
0.010	1.000	157.800	2083.223

CFI/TLI

CFI

Mean 0.951
Std Dev 0.001
Number of successful computations 10

Proportions		Percentiles	
Expected	Observed	Expected	Observed
0.990	1.000	0.950	0.950
0.980	1.000	0.950	0.950
0.950	1.000	0.950	0.950
0.900	0.900	0.950	0.950
0.800	0.800	0.951	0.950
0.700	0.600	0.951	0.951
0.500	0.500	0.951	0.951
0.300	0.300	0.952	0.951
0.200	0.200	0.952	0.952
0.100	0.100	0.952	0.952
0.050	0.100	0.952	0.952
0.020	0.100	0.952	0.952
0.010	0.000	0.953	0.952

TLI

Mean 0.987
Std Dev 0.000
Number of successful computations 10

Proportions Percentiles

Expected	Observed	Expected	Observed
0.990	1.000	0.987	0.987
0.980	1.000	0.987	0.987
0.950	1.000	0.987	0.987
0.900	0.900	0.987	0.987
0.800	0.800	0.987	0.987
0.700	0.600	0.987	0.987
0.500	0.500	0.987	0.987
0.300	0.300	0.987	0.987
0.200	0.200	0.987	0.987
0.100	0.100	0.987	0.987
0.050	0.100	0.988	0.987
0.020	0.100	0.988	0.987
0.010	0.000	0.988	0.987

RMSEA (Root Mean Square Error Of Approximation)

Mean	0.049
Std Dev	0.000
Number of successful computations	10

Proportions		Percentiles	
Expected	Observed	Expected	Observed
0.990	1.000	0.048	0.048
0.980	0.900	0.048	0.048
0.950	0.900	0.048	0.048
0.900	0.900	0.048	0.048
0.800	0.800	0.048	0.048
0.700	0.800	0.048	0.048
0.500	0.500	0.049	0.048
0.300	0.500	0.049	0.049
0.200	0.100	0.049	0.049
0.100	0.100	0.049	0.049
0.050	0.000	0.049	0.049
0.020	0.000	0.049	0.049
0.010	0.000	0.049	0.049

SRMR (Standardized Root Mean Square Residual)

Mean	0.025
Std Dev	0.000
Number of successful computations	10

Proportions		Percentiles	
Expected	Observed	Expected	Observed
0.990	1.000	0.025	0.025
0.980	1.000	0.025	0.025
0.950	0.900	0.025	0.025
0.900	0.900	0.025	0.025
0.800	0.800	0.025	0.025
0.700	0.700	0.025	0.025
0.500	0.500	0.025	0.025
0.300	0.400	0.025	0.025
0.200	0.300	0.025	0.025
0.100	0.100	0.025	0.025
0.050	0.000	0.025	0.025
0.020	0.000	0.026	0.025
0.010	0.000	0.026	0.025

WRMR (Weighted Root Mean Square Residual)

```
     Mean                               2.504
     Std Dev                            0.015
     Number of successful computations     10

          Proportions                Percentiles
     Expected    Observed         Expected       Observed
        0.990       1.000            2.468          2.471
        0.980       0.900            2.472          2.471
        0.950       0.900            2.478          2.471
        0.900       0.900            2.484          2.471
        0.800       0.800            2.491          2.471
        0.700       0.700            2.496          2.495
        0.500       0.500            2.504          2.499
        0.300       0.500            2.512          2.513
        0.200       0.100            2.517          2.516
        0.100       0.100            2.524          2.517
        0.050       0.000            2.529          2.517
        0.020       0.000            2.536          2.517
        0.010       0.000            2.540          2.517


MODEL RESULTS

                   Estimates     S.E.  Est./S.E.    Std     StdYX

 FY35     BY
    Y35                1.000    0.000      0.000    1.037    1.000

 FY16     BY
    Y16                1.000    0.000      0.000    1.914    1.000

 FAMILY   BY
    Y2                 1.000    0.000      0.000    0.875    0.658
    Y5                 1.309    0.042     30.889    1.145    0.753
    Y6                 1.168    0.038     30.724    1.021    0.714
    Y7                 1.639    0.053     30.821    1.433    0.820

 FRIENDS  BY
    Y8                 1.000    0.000      0.000    1.192    0.766
    Y11                0.936    0.032     28.930    1.116    0.745
    Y15                1.128    0.041     27.372    1.345    0.802
    Y38                0.455    0.021     21.587    0.542    0.374

 LIVING   BY
    Y27                1.000    0.000      0.000    1.242    0.754
    Y32                1.013    0.049     20.728    1.258    0.747

 SCHOOL   BY
    Y21                1.000    0.000      0.000    1.168    0.760
    Y18                0.747    0.022     33.267    0.872    0.597
    Y23                1.331    0.051     26.068    1.555    0.841
    Y24                0.873    0.023     37.556    1.019    0.688

 SELF     BY
    Y36                1.000    0.000      0.000    0.742    0.596
    Y37                2.287    0.084     27.212    1.696    0.861
    Y38                0.846    0.034     25.185    0.627    0.432

 QOL      BY
```

```
    LIFEPIC8           1.000      0.000      0.000      1.105      0.741
    F14SAT             1.014      0.037     27.778      1.121      0.746

 MH       BY
    MENHLTHR           1.000      0.000      0.000      1.414      1.000

 PH       BY
    PHYHLTHR           1.000      0.000      0.000      1.532      1.000

 FRIENDS  BY
    FY16               1.000      0.000      0.000      0.623      0.623

 SELF     BY
    FY35               1.000      0.000      0.000      0.715      0.715

 F        BY
    MH                 1.000      0.000      0.000      0.707      0.707
    PH                 1.160      0.045     25.805      0.757      0.757

 QOL      ON
    FAMILY             0.290      0.026     11.233      0.229      0.229
    FRIENDS           -0.017      0.016     -1.066     -0.018     -0.018
    LIVING             0.047      0.017      2.812      0.053      0.053
    SCHOOL             0.017      0.012      1.364      0.018      0.018
    SELF               0.617      0.034     18.267      0.414      0.414
    MH                 0.259      0.012     22.037      0.332      0.332
    PH                 0.039      0.010      3.915      0.054      0.054

 FAMILY   ON
    MH                 0.225      0.011     20.107      0.364      0.364
    PH                 0.045      0.009      4.859      0.078      0.078

 FRIENDS  ON
    MH                 0.236      0.015     15.611      0.280      0.280
    PH                 0.070      0.013      5.312      0.090      0.090

 LIVING   ON
    MH                 0.284      0.017     16.633      0.323      0.323
    PH                 0.071      0.014      4.950      0.088      0.088

 SCHOOL   ON
    MH                 0.169      0.014     12.303      0.205      0.205
    PH                 0.086      0.013      6.870      0.113      0.113

 SELF     ON
    MH                 0.224      0.009     24.370      0.427      0.427
    PH                 0.106      0.008     13.598      0.219      0.219

 Y18      ON
    FY35               0.407      0.021     19.649      0.422      0.289

 Y24      ON
    FY35               0.202      0.018     11.330      0.209      0.141

 Y27      ON
    FY16               0.105      0.013      8.097      0.200      0.121

 Y32      ON
    FY16               0.139      0.014     10.176      0.267      0.158
```

FAMILY WITH					
FRIENDS	0.408	0.020	20.320	0.391	0.391
LIVING	0.576	0.025	23.233	0.530	0.530
SCHOOL	0.359	0.017	21.037	0.352	0.352
SELF	0.238	0.011	21.033	0.367	0.367
FRIENDS WITH					
LIVING	0.447	0.034	13.169	0.302	0.302
SCHOOL	0.273	0.021	13.158	0.196	0.196
SELF	0.373	0.017	21.530	0.422	0.422
LIVING WITH					
SCHOOL	0.499	0.027	18.534	0.344	0.344
SELF	0.283	0.016	17.694	0.307	0.307
SCHOOL WITH					
SELF	0.180	0.013	14.071	0.208	0.208
Variances					
F	1.000	0.000	0.000	1.000	1.000
Residual Variances					
Y16	0.000	0.000	0.000	0.000	0.000
Y35	0.000	0.000	0.000	0.000	0.000
PHYHLTHR	0.000	0.000	0.000	0.000	0.000
MENHLTHR	0.000	0.000	0.000	0.000	0.000
FY35	0.526	0.034	15.313	0.489	0.489
FY16	2.241	0.167	13.421	0.612	0.612
FAMILY	0.636	0.031	20.778	0.831	0.831
FRIENDS	1.260	0.066	19.085	0.887	0.887
LIVING	1.322	0.078	17.011	0.858	0.858
SCHOOL	1.255	0.056	22.476	0.921	0.921
SELF	0.368	0.017	21.049	0.670	0.670
QOL	0.292	0.021	14.038	0.239	0.239
MH	1.000	0.000	0.000	0.500	0.500
PH	1.000	0.000	0.000	0.426	0.426

R-SQUARE

Observed Variable	Scale Factors	R-Square
Y2	0.753	0.433
Y5	0.658	0.567
Y6	0.700	0.510
Y7	0.572	0.673
Y8	0.643	0.587
Y11	0.667	0.555
Y15	0.597	0.644
Y16	0.523	1.000
Y18	0.685	0.531
Y21	0.650	0.577
Y23	0.541	0.707
Y24	0.675	0.544
Y27	0.607	0.632
Y32	0.594	0.647
Y35	0.964	1.000
Y36	0.803	0.355
Y37	0.508	0.742
Y38	0.689	0.525

LIFEPIC8	0.671	0.550
F14SAT	0.666	0.557
PHYHLTHR	0.653	1.000
MENHLTHR	0.707	1.000

Latent Variable	R-Square
FY35	0.511
FY16	0.388
FAMILY	0.169
FRIENDS	0.113
LIVING	0.142
SCHOOL	0.079
SELF	0.330
QOL	0.761
MH	0.500
PH	0.574

QUALITY OF NUMERICAL RESULTS

Average Condition Number for the Information Matrix 0.321E-03
(ratio of smallest to largest eigenvalue)

TECHNICAL 1 OUTPUT

PARAMETER SPECIFICATION

LAMBDA

	FY35	FY16	FAMILY	FRIENDS	LIVING
Y2	0	0	0	0	0
Y5	0	0	1	0	0
Y6	0	0	2	0	0
Y7	0	0	3	0	0
Y8	0	0	0	0	0
Y11	0	0	0	4	0
Y15	0	0	0	5	0
Y16	0	0	0	0	0
Y18	0	0	0	0	0
Y21	0	0	0	0	0
Y23	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0
Y35	0	0	0	0	0
Y36	0	0	0	0	0
Y37	0	0	0	0	0
Y38	0	0	0	8	0
LIFEPIC8	0	0	0	0	0
F14SAT	0	0	0	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

LAMBDA

SCHOOL	SELF	QOL	MH	PH

Y2	0	0	0	0	0
Y5	0	0	0	0	0
Y6	0	0	0	0	0
Y7	0	0	0	0	0
Y8	0	0	0	0	0
Y11	0	0	0	0	0
Y15	0	0	0	0	0
Y16	0	0	0	0	0
Y18	0	0	0	0	0
Y21	0	0	0	0	0
Y23	6	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0
Y35	0	0	0	0	0
Y36	0	0	0	0	0
Y37	0	7	0	0	0
Y38	0	9	0	0	0
LIFEPIC8	0	0	0	0	0
F14SAT	0	0	10	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

LAMBDA

	F	Y18	Y24	Y27	Y32
Y2	0	0	0	0	0
Y5	0	0	0	0	0
Y6	0	0	0	0	0
Y7	0	0	0	0	0
Y8	0	0	0	0	0
Y11	0	0	0	0	0
Y15	0	0	0	0	0
Y16	0	0	0	0	0
Y18	0	0	0	0	0
Y21	0	0	0	0	0
Y23	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0
Y35	0	0	0	0	0
Y36	0	0	0	0	0
Y37	0	0	0	0	0
Y38	0	0	0	0	0
LIFEPIC8	0	0	0	0	0
F14SAT	0	0	0	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

THETA

	Y2	Y5	Y6	Y7	Y8
Y2	0				
Y5	0	0			
Y6	0	0	0		
Y7	0	0	0	0	
Y8	0	0	0	0	0
Y11	0	0	0	0	0
Y15	0	0	0	0	0

Y16	0	0	0	0	0
Y18	0	0	0	0	0
Y21	0	0	0	0	0
Y23	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0
Y35	0	0	0	0	0
Y36	0	0	0	0	0
Y37	0	0	0	0	0
Y38	0	0	0	0	0
LIFEPIC8	0	0	0	0	0
F14SAT	0	0	0	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

THETA

	Y11	Y15	Y16	Y18	Y21
Y11	0				
Y15	0	0			
Y16	0	0	0		
Y18	0	0	0	0	
Y21	0	0	0	0	0
Y23	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0
Y35	0	0	0	0	0
Y36	0	0	0	0	0
Y37	0	0	0	0	0
Y38	0	0	0	0	0
LIFEPIC8	0	0	0	0	0
F14SAT	0	0	0	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

THETA

	Y23	Y24	Y27	Y32	Y35
Y23	0				
Y24	0	0			
Y27	0	0	0		
Y32	0	0	0	0	
Y35	0	0	0	0	0
Y36	0	0	0	0	0
Y37	0	0	0	0	0
Y38	0	0	0	0	0
LIFEPIC8	0	0	0	0	0
F14SAT	0	0	0	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

THETA

	Y36	Y37	Y38	LIFEPIC8	F14SAT
Y36	0				

Y37	0	0			
Y38	0	0	0		
LIFEPIC8	0	0	0	0	
F14SAT	0	0	0	0	0
PHYHLTHR	0	0	0	0	0
MENHLTHR	0	0	0	0	0

THETA

	PHYHLTHR	MENHLTHR
PHYHLTHR	0	
MENHLTHR	0	0

BETA

	FY35	FY16	FAMILY	FRIENDS	LIVING
FY35	0	0	0	0	0
FY16	0	0	0	0	0
FAMILY	0	0	0	0	0
FRIENDS	0	0	0	0	0
LIVING	0	0	0	0	0
SCHOOL	0	0	0	0	0
SELF	0	0	0	0	0
QOL	0	0	21	22	23
MH	0	0	0	0	0
PH	0	0	0	0	0
F	0	0	0	0	0
Y18	29	0	0	0	0
Y24	31	0	0	0	0
Y27	0	33	0	0	0
Y32	0	34	0	0	35

BETA

	SCHOOL	SELF	QOL	MH	PH
FY35	0	0	0	0	0
FY16	0	0	0	0	0
FAMILY	0	0	0	11	12
FRIENDS	0	0	0	13	14
LIVING	0	0	0	15	16
SCHOOL	0	0	0	17	18
SELF	0	0	0	19	20
QOL	24	25	0	26	27
MH	0	0	0	0	0
PH	0	0	0	0	0
F	0	0	0	0	0
Y18	30	0	0	0	0
Y24	32	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0

BETA

	F	Y18	Y24	Y27	Y32
FY35	0	0	0	0	0
FY16	0	0	0	0	0

FAMILY	0	0	0	0	0
FRIENDS	0	0	0	0	0
LIVING	0	0	0	0	0
SCHOOL	0	0	0	0	0
SELF	0	0	0	0	0
QOL	0	0	0	0	0
MH	0	0	0	0	0
PH	28	0	0	0	0
F	0	0	0	0	0
Y18	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0

PSI

	FY35	FY16	FAMILY	FRIENDS	LIVING
FY35	36				
FY16	0	37			
FAMILY	0	0	38		
FRIENDS	0	0	39	40	
LIVING	0	0	41	42	43
SCHOOL	0	0	44	45	46
SELF	0	0	48	49	50
QOL	0	0	0	0	0
MH	0	0	0	0	0
PH	0	0	0	0	0
F	0	0	0	0	0
Y18	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0

PSI

	SCHOOL	SELF	QOL	MH	PH
SCHOOL	47				
SELF	51	52			
QOL	0	0	53		
MH	0	0	0	0	
PH	0	0	0	0	0
F	0	0	0	0	0
Y18	0	0	0	0	0
Y24	0	0	0	0	0
Y27	0	0	0	0	0
Y32	0	0	0	0	0

PSI

	F	Y18	Y24	Y27	Y32
F	0				
Y18	0	0			
Y24	0	0	0		
Y27	0	0	0	0	
Y32	0	0	0	0	0

```
     Beginning Time:  11:57:07
        Ending Time:  11:59:02
       Elapsed Time:  00:01:55
```

MUTHÉN & MUTHÉN
3463 Stoner Ave.
Los Angeles, CA 90066

Tel: (310) 391-9971
Fax: (310) 391-8971
Web: www.StatModel.com
Support: Support@StatModel.com

Appendix E Comparison of correlations among dimensions of life satisfaction across studies

In this Appendix, we provide a brief overview of our literature search for studies that used the MSLSS and that reported the correlations among the five dimensions of life satisfaction. We also discuss the methods that were used to: (a) estimate the average correlation for each relationship across studies and (b) calculate the between-studies variance in the distribution of the correlations across studies. The findings of these analyses were used to compare the correlations in our study with those found in other studies.

Of the 39 studies that we identified in our literature review (see discussion of the literature review on p. 28), we identified 15 studies in English that reported correlations between the MSLSS subscales. Several studies had more than one subsample for which correlations matrices were provided. In all, we obtained 23 independent correlation matrices. A brief description of the sample for each of the correlations is provided in Table 30.

The values of each of the 23 correlations are shown in Figure 24. We calculated confidence intervals by transforming the Pearson correlations to standardized Fisher correlations and by subsequently determining the variance for each correlation recommended by Lipsey and Wilson (2001). The Fisher correlations and the corresponding 95% confidence intervals were then transformed back to Pearson correlations prior to creating the graph.

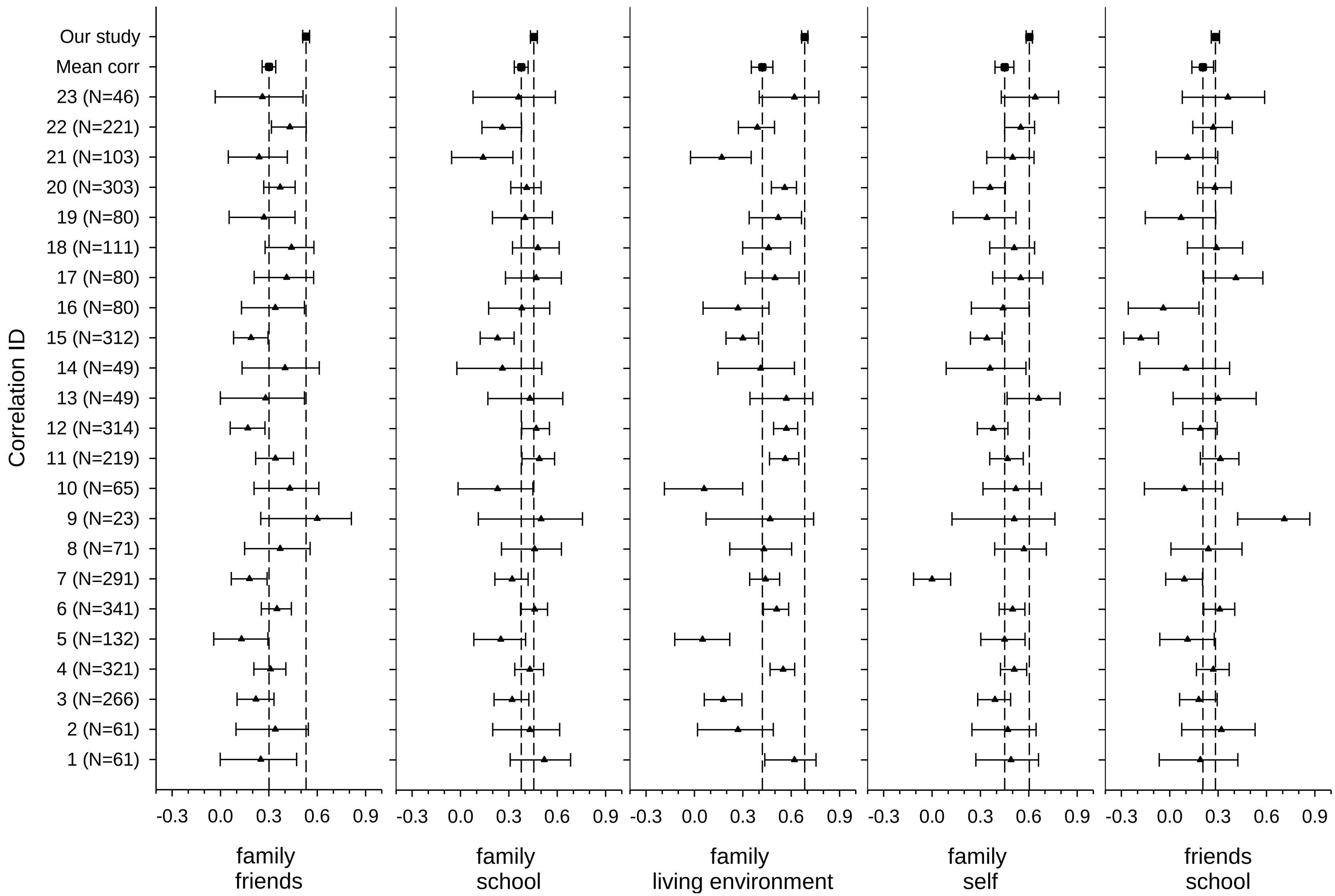
To calculate the mean of the correlations for each relationship across the 23 correlation matrices, we first estimated a fixed effect model for each correlation. We then calculated the between-studies variances, and used the Q-statistic to determine whether the differences between the studies were statistically significant (Lipsey & Wilson, 2001). Having found statistical significance, we then proceeded by estimating the mean correlations based on a random effects model. Of course, the variances between the studies were not likely to be entirely random. However, a detailed analysis of the differences between studies (as in a mixed effects analysis) was not desired for our purposes.

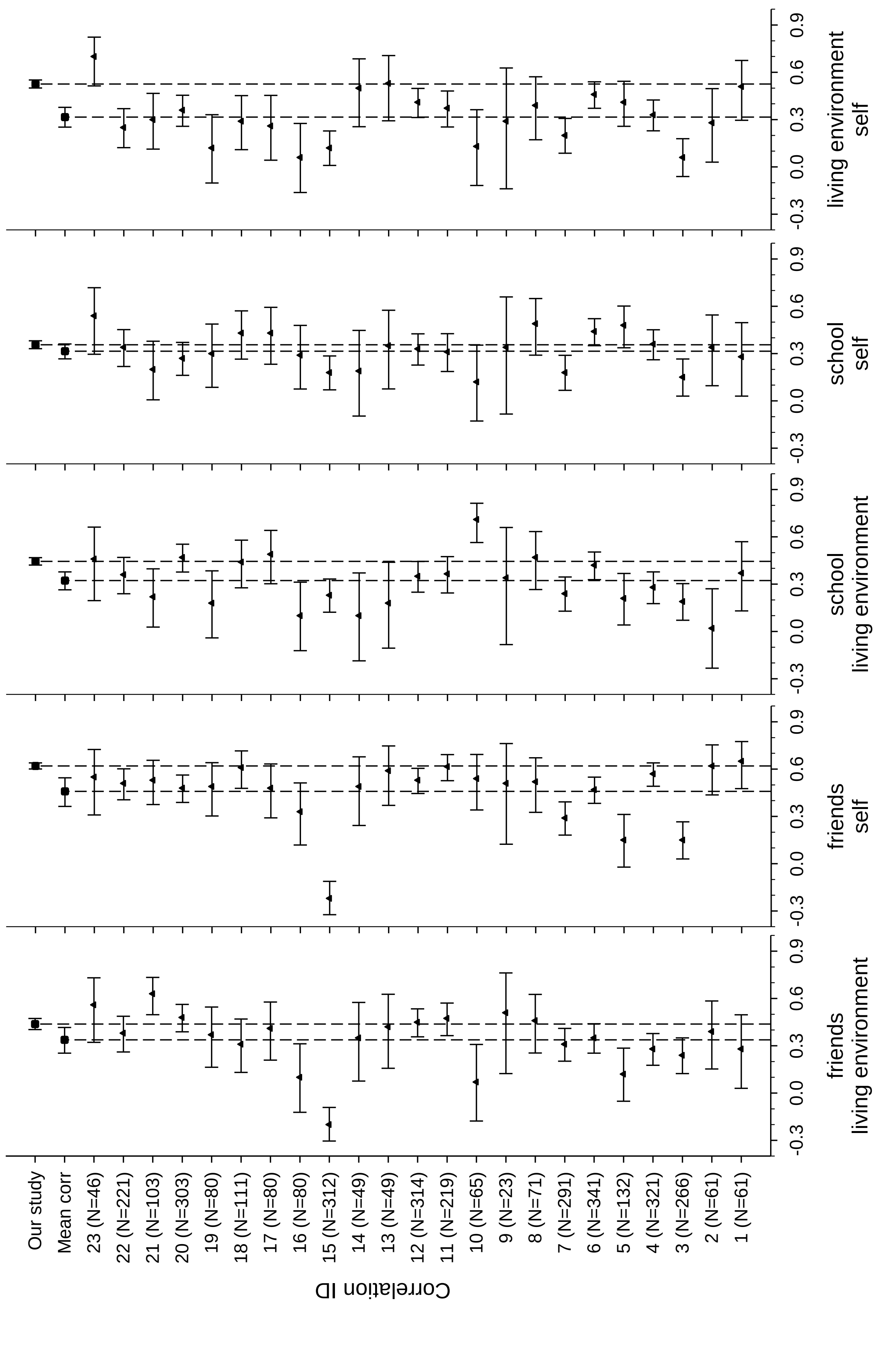
Table 30 Sample descriptions of studies that produced correlations among the dimensions of life satisfaction

ID	Reference	Sample description
1 (*N*=61)	Ash & Huebner (1998)	"Academically gifted" students, grades 6 to 8, from 1 urban school in a Southeastern US state
2 (*N*=61)	Ash & Huebner (1998)	"Not academically gifted" students, grades 6 to 8, from 1 urban school in a Southeastern US state
3 (*N*=266)	Gilligan & Huebner (2002)	Students, grades 9 to 12, from 2 high schools in a Southeastern US state
4 (*N*=321)	Gilman (1999)	Students, grades 9 to 12, from 2 high schools in a Southeastern US state
5 (*N*=132)	Gilman & Ashby (2003)	Students, grades 6 to 8, from 6 schools from one school district in a Southeastern US state
6 (*N*=341)	Gilman et al. (2005)	High school students (mean age (*SD*) = 14.6 (2.1) from 1 school district in the US
7 (*N*=291)	Gilman et al. (2005)	High school students (mean age (*SD*) = 15.1 (1.6) from 4 schools across 3 cities in Croatia
8 (*N*=71)	Gilman et al. (2004)	High school students (mean age (*SD*) = 14.1 (2.0) in the US
9 (*N*=23)	Gilman et al. (2004)	"Deaf/hard-of-hearing" students (mean age (*SD*) = 13.2 (2.5) from 1 segregated day school in a Southeastern US state.
10 (*N*=65)	Gilman et al. (2004)	"Deaf/hard-of-hearing" students (mean age (*SD*) = 15.1 (2.5) from 2 residential settings in a Midwestern and Northeastern US state.
11 (*N*=219)	Huebner et al. (1998)	Students, grades 6 to 8, from 2 urban middle schools from 2 metropolitan school districts in a Southeastern US state.
12 (*N*=314)	Greenspoon & Saklofske (1998)	Students, grades 3 to 8, from Western Canadian schools.

ID	Reference	Sample description
13 (*N*=49)	Griffin & Huebner (2000)	Students, grades 6 to 8, from 11 schools within a metropolitan area in a Southeastern US city.
14 (*N*=49)	Griffin & Huebner (2000)	Special education students, grades 6 to 8, from 11 schools within a metropolitan area in a Southeastern US city.
15 (*N*=312)	Huebner (1994)	Students, grades 3 to 8, in schools within urban, suburban and rural school districts in a Southeastern US state.
16 (*N*=80)	Huebner et al. (2002)	“Normally achieving” students, grades 9 to 12, from rural and urban school districts in South Carolina and Georgia (US).
17 (*N*=80)	Huebner et al. (2002)	“Mild mentally disabled” students, grades 9 to 12, from rural and urban school districts in South Carolina and Georgia (US).
18 (*N*=111)	McCullough (2003)	“Normally achieving” students, grades 9 to 12, from 1 high school in South Carolina (US).
19 (*N*=80)	McCullough (2003)	“Learning disabled” students, grades 9 to 12, from 2 high schools in South Carolina (US).
20 (*N*=303)	Nickerson & Nagle (2004)	Students, grades 4 to 8, from 3 elementary and 3 middle schools in a Southeastern US city.
21 (*N*=103)	Schiff et al. (2006)	Children and adolescents, ages 10 to 18, in a residential treatment setting (i.e., not living with their parents) in Israel.
22 (*N*=221)	Seligson et al. (2003)	Students, grades 6 to 8, from a school in a Southeastern US city.
23 (*N*=46)	Seligson et al. (2003)	Students (mean age (*SD*) = 15.7 (1.0) from a high school in a Southeastern US city.

Figure 24 A comparison of correlations among dimensions of life satisfaction across studies




Correlation ID
Our study
Mean corr
23 (N=46)
22 (N=221)
21 (N=103)
20 (N=303)
19 (N=80)
18 (N=111)
17 (N=80)
16 (N=80)
15 (N=312)
14 (N=49)
13 (N=49)
12 (N=314)
11 (N=219)
10 (N=65)
9 (N=23)
8 (N=71)
7 (N=291)
6 (N=341)
5 (N=132)
4 (N=321)
3 (N=266)
2 (N=61)
1 (N=61)
-0.3 0.0 0.3 0.6 0.9
friends
living environment
friends
self
school
living environment
school
self
living environment
self