

Validation of a Measure of Weight-Related Quality of Life in a Community Sample
of Normal Weight, Overweight, and Obese 4th and 5th Grade Students

By

Christopher C. Cushing
Master of Science, Missouri State University, 2007

Submitted to the graduate degree program in Clinical Child Psychology and the
Graduate Faculty of the University of Kansas in partial fulfillment of the
requirements for the degree of Doctor of Philosophy.

Chairperson Ric G. Steele, Ph.D., ABPP

Michael C. Roberts, Ph.D., ABPP

Ann M. Davis, Ph.D., MPH, ABPP

Nancy A. Hamilton, Ph.D.

Jeffrey A. Hall, Ph.D.

Date Defended: June 16, 2011

The Dissertation Committee for Christopher C. Cushing
certifies that this is the approved version of the following dissertation:

Validation of a Measure of Weight-Related Quality of Life in a Community Sample
of Normal Weight, Overweight, and Obese 4th and 5th Grade Students

Chairperson Ric G. Steele, Ph.D., ABPP

Date approved: June 21, 2011

Abstract

The current study extends the quality of life assessment literature by examining the reliability and validity of a disease-specific instrument in a sample of nontreatment-seeking school aged children with overweight and obesity. Participants were 4th and 5th grade students recruited from six Kansas elementary schools. Results of the current study were consistent with the initial evaluation of *Sizing Me Up* and revealed a five-factor first-order factor structure for the 22-item measure with one second-order factor representing a total score. Consistent with study hypotheses and the available literature, factorial invariance could not be established between a sample of children with healthy weight ($n = 168$) and the primary sample ($n = 134$) of children with overweight and obesity. Good evidence for convergent validity within *Sizing Me Up* factors as well as with similar constructs measured by a general quality of life instrument were revealed. The *Sizing Me Up* also demonstrated evidence for criterion-related validity with BMI%ile. The current study also advances the quality of life assessment literature by empirically testing the assumption that disease-specific measures assess different constructs than general quality of life measures. Study hypotheses that *Sizing Me Up* assesses weight-related quality of life constructs were supported. Finally, reliabilities for the five-factor *Sizing Me Up* factor structure were acceptable for research purposes. However, the scales are unacceptable for clinical use and only the total score should be used with individual children.

Acknowledgements

This project was made possible by funding from the general Clinical Child Psychology Program research fund, Pioneer Class Dissertation Award, Graduate College Dissertation Fund, and the Pediatric Health Promotion and Maintenance Lab. I want to thank my dissertation committee: Michael Roberts, who brings a keen eye for the real world significance of empirical work; Ann Davis, whose expertise and experience ensured that the current project would produce a result valuable to the field; Nancy Hamilton, who challenged me to consider several helpful additions to my proposed hypotheses, and Jeffrey Hall, who asked fundamental questions regarding the logical assumptions in my arguments. To my chairperson and advisor Ric Steele, Ric—you are truly an outstanding mentor. Losing the opportunity to work closely with you is the only downside to progressing beyond this point in my graduate training. I am certain that I would remain content, intellectually stimulated, and remarkably happy were I to remain under your close guidance for years to come. You are a patient, supportive, and intelligent mentor who makes me forget the major hurdles I have to overcome. Working with you is freeing. I want to acknowledge the love, support, and sacrifice of my family. To my parents William and Mary Cushing, I love you for the support that has brought me to this accomplishment; but I love you more because you would have been there in equal measure had I chosen a different pursuit. To my loving wife Angela Cushing, I am certain that I would not have accomplished this milestone without you. Your love, selflessness, and faith in me made my accomplishments possible—made them worthwhile.

Table of Contents

Abstract	iii
Acknowledgements.....	iv
Table of Contents	v
List of Tables	vii
List of Figures	viii
Overview	1
Consequences of Childhood Obesity	2
Health-Related QOL	4
QOL Assessment.....	5
Study Aims	13
Method	16
Participants	16
Sampling Procedure and Questionnaire Administration.....	17
Measures.....	19
Results.....	23
Data Screening	23
Aim 1	27
Nine-factor alternative model.....	33
Aim 2.....	39

Transformed Means and Standard Deviations	40
Reliability of First and Second Order Factors and Minimal Clinically Important Difference.....	43
Discussion.....	44
Construct Validity and Invariance.....	44
Clinical Implications	49
Limitations	51
Future Directions.....	51
Conclusion.....	53
Appendix A.....	66

List of Tables

Table 1: <i>Demographic Characteristics</i>	17
Table 2: <i>Descriptive Statistics for Sizing Me Up and the PedsQL</i>	24
Table 3: <i>Estimated and Standardized Factor Loadings, Residuals, and R² Values for Each Sizing Me Up and PedsQL Indicator</i>	31
Table 4: <i>Estimated and Standardized Factor Loadings, Residuals, and R² Values for Each Sizing Me Up and PedsQL Indicator</i>	35
Table 5: <i>Scaled Means and Standard Deviations</i>	41
Table 6: <i>Scaled Means and Standard Deviations by Weight Category</i>	42
Table 7: <i>Sizing Me Up Reliability and MCID Statistics</i>	43
Table 8: <i>PedsQL Reliability and MCID Statistics</i>	44

List of Figures

Figure 1: <i>Latent Correlations and Regression Paths for Aim 1 Alternative Model ...</i>	30
Figure 2: <i>Latent Intercorrelations of the Aim 2 Nine-Factor Model</i>	34
Figure 3: <i>Second-Order Factor Structure.....</i>	39

**Validation of a Measure of Weight-Related Quality of Life in a Community
Sample of Normal Weight, Overweight, and Obese 4th and 5th Grade Students**

Overview

Childhood obesity negatively affects the physical and psychosocial functioning of a significant number of youths in the United States. Several efforts to reduce the negative impact of weight-related problems have demonstrated success. Bariatric surgery, pharmaceutical interventions, group family-based, and individual inpatient and outpatient treatments all represent viable options for reducing weight in children with overweight and obesity (Braet, Tanghe, Decaluwé, Moens, & Rosseel, 2004; Kitzmann et al., 2010; Lawson et al., 2006; McGovern et al., 2008). However, to address only weight without consideration of the whole person is too insular and must be expanded to take the child's experience of their condition and treatment into account. Doing so allows researchers and clinicians the opportunity to understand the unique challenges faced by children at the outset of treatment as well as the ability to monitor and understand what changes in functioning are relevant from the patient's perspective. These pieces of information can be critical to determining what may be motivating or inhibiting to a particular child participating in a weight-management treatment.

Below, weight-related Health-Related Quality of Life (referred to as weight-related QOL for clarity) will be discussed as a sub-construct of general QOL that holds promise for achieving a patient-centered perspective of the overweight and

obesity experience, and if measured properly, may provide a proximal indicator of treatment success that is likely to demonstrate clinically meaningful change before weight-related outcomes. The pediatric QOL literature will be reviewed with the goal of elucidating the need for a well-validated measure of weight-related QOL in nontreatment-seeking school-aged children.

Consequences of Childhood Obesity

Physical health consequences of obesity. Childhood overweight and obesity affects approximately one third of children and adolescents in the United States (Ogden, Carroll, Curtin, Lamb, & Flegal, 2010). It is well-established that children with overweight and obesity are at significant risk for childhood and adult diseases as well as early mortality. Early warning signs of diseases previously thought to only occur in adults have been documented in children with obesity. Specifically, atherosclerotic vascular disease and coronary artery disease warning signs have appeared in children with obesity as young as 3 and 8-years-old, respectively (Freedman, 2002). In addition, the prevalence of endocrine disorders such as type 2 diabetes and menstrual abnormalities in females has increased dramatically in children and adolescents, further evidencing the downward shift of severe adult health conditions (Remsberg, Demerath, Schubert, Chumela, Sun, & Siervogel, 2005; Young, Dean, Flett, & Wood-Steiman, 2000). Children with obesity are at greater risk for pulmonary problems such as asthma and sleep disorders that in extreme cases may put them at further risk for impaired memory functioning and learning disorders (i.e.,

> 200% of ideal bodyweight; Gilliland et al., 2003; Mallory, Fiser, & Jackson, 1989; Rhodes et al., 1995).

Individual level psychosocial consequences of obesity. In addition to physical sequela, children and adolescents with overweight and obesity also experience significant consequences to their psychosocial functioning. A number of studies indicate that pediatric obesity is associated with internalizing symptoms (i.e., depressive/anxiety symptoms) in both clinical and non-clinical samples (see Zeller & Modi, 2008 for review); although, it is believed that internalizing symptomatology only rises to the level of clinical significance in approximately 11% of children and adolescents with obesity meaning that rates of clinical elevations are similar to those observed in the general population (Zeller & Modi, 2006). However, the absence of psychiatric diagnosis does not mean that the child with overweight or obesity does not experience clinically significant impairments in important psychosocial domains. For example, much has been learned about the day-to-day functioning of children and adolescents with overweight or obesity through the dramatic expansion of QOL assessment over the past decade. It has been observed that children with overweight and obesity are at significant risk for physical, emotional, and social QOL impairments with some studies reporting impairments similar to children with cancer (Friedlander, Larkin, Rosen, Palermo, & Redline, 2003; Schwimmer, Burwinkle, & Varni, 2003; Williams, Wake, Hesketh, Maher, & Waters, 2005). Given the severity of the QOL problems attendant to pediatric overweight and obesity, a nuanced

understanding of the QOL experience of children with overweight and obesity seems worthwhile.

Health-Related QOL

Historically, the emergence of QOL as a construct marked an important shift away from the view that health was simply the absence of infirmity (World Health Organization, 1947). Multidimensional QOL consists of an individual's physical health status, psychological and social functioning, and emotional well-being (Eiser & Morse, 2001). The multidimensionality of QOL is widely accepted and dates back to the original definition proposed by the World Health Organization, which has been called the "cornerstone" of other QOL definitions (Spieth & Harris, 1996). The World Health Organization (1947) specified that QOL is comprised of an individual's perspective of his or her own physical, mental, and social well-being. General QOL includes *health-related* and *environmental* QOL (Spilker & Revicki, 1999). *Health-related* QOL is a multidimensional construct that refers to an individual's functioning as directly affected by an illness or its treatment (Spieth & Harris, 1996). *Health-related* QOL is distinct from *environmental* QOL in that it refers to a subjective experience that is modulated by a disease or by the application of a treatment to an individual rather than the impact of one's environment on QOL (Jaschke, Singer, & Guyatt, 1989; Speith & Harris, 1996). Therefore, health-related QOL can include both positive and negative experiences associated with a health condition or its treatment.

QOL Assessment

QOL can be reliably and validly measured both broadly and narrowly within particular illness groups as well as across a spectrum of ill and healthy groups (Limbers, Newman, & Varni, 2008b; Palermo, Long, Lewandowski, Drotar, Quittner, & Walker, 2008). A strength of QOL is the reliance on the subjective experience of an individual to ensure that the patient's perspective is heard and valued during treatment (Guyatt, Feeny, & Patrick, 1993; Spieth & Harris, 1996). Assessing QOL requires the measurement of the individual's perception of her/his own physical health status, psychological and social functioning, and emotional well-being (Eiser & Morse, 2001; Guyatt et al., 1993; Palermo et al., 2008). In children, special care and consideration is given to assessing QOL in a way that accounts for age, reading ability, and emotional maturity (Turner, Quittner, Parsuraman, & Cleeland, 2007). When these considerations are taken into account, QOL can be measured equivalently across age-groups (Limbers, Newman, & Varni, 2008a). Children as young as 5-years of age are considered to be accurate reporters of QOL (Varni, Limbers, & Burwinkle, 2007). Consequently, QOL assessment in children should rely on self-report whenever possible and efforts should be made to ensure that developmentally appropriate measures are available for all age-groups.

QOL is assessed either by using broad and general or disease-specific QOL instruments depending on the scope of the health domains being assessed and the need to compare individual scores against normative data. This range of QOL assessment foci highlights the implicit assumption that, while QOL is universally

experienced, it can also be uniquely impacted by a disease or condition. For example, although asthma and diabetes might both be expected to negatively impact QOL (Varni et al., 2003), each is assumed to be associated with stressors or impairments that are unique to each specific illness (Ingersoll & Marrero, 1991; Juniper, Guyatt, Feeny, Ferrie, Griffith, & Townsend, 1996). Accepting this theoretical assumption leads to the conclusion that measurement instruments can be developed to assess theoretically distinct disease-specific constructs that may have greater sensitivity within a disease population than general measures. However, the assumption that disease specific measures contribute unique information to QOL assessment by tapping into a disease-specific QOL construct frequently goes untested in the assessment literature. This means that recommendations for assessment are based largely on assumptions rather than empirical evidence.

Broad and general QOL assessment. Broad and general instruments are typically multidimensional measures of physical, psychological, social, school, and family functioning (Landgraf, Abetz, & Ware, 1996; Varni et al., 2001). These measures are typically designed to be highly generalizable to allow for comparisons of QOL across groups including healthy controls. This use allows for epidemiological investigations of impairment relative to children experiencing another chronic-illness condition or to children without illness (Varni, Burwinkle, & Lane, 2005). For example, epidemiological studies have examined differences across children with diabetes, gastrointestinal disorders, cardiovascular problems, asthma, obesity, end stage renal disease, psychiatric diagnoses, cancer, rheumatoid arthritis,

and cystic fibrosis (Friedlander et al., 2003; Varni et al., 2007). These studies help to extend the understanding of the pediatric disease experience beyond just the frequency and intensity of the child's symptoms and may allow for the compilation and comparison of normative data across health conditions. Broad and general measures should also have the ability to discriminate between varying degrees of illness such that greater impairment translates to lower QOL scores on the instrument (e.g., Varni et al., 2007). Broad and general measures should measure theoretical constructs consistently across healthy and ill groups (Palermo et al., 2008; Spieth & Harris, 1996). A number of broad and general measures of QOL have demonstrated well-established reliability and validity in the pediatric psychology literature (Palermo et al., 2008) including the Child Health and Illness Profile (Starfield, Riley, & Green, 1999), the Child Health Questionnaire (Landgraf et al., 1996), the Pediatric Quality of Life Inventory (*PedsQL*TM; Varni et al., 1999), and the Youth Quality of Life (Edwards, Hubner, Connell, & Patrick, 2002). Recently a study of the *PedsQL* demonstrated that the items are interpreted similarly by both healthy and ill children, lending credibility to observed epidemiological differences and allowing for confident use of the instrument in large heterogeneous disease populations (Limbers et al., 2008a). To date, the *PedsQL* is the most widely researched of these instruments and has the most clear evidence for reliability and validity. Finally, the advantages of broad and general measures are amplified by the fact that QOL assessment is inexpensive, the questionnaires are brief, and allow for multiple reporters (Palermo et al., 2008; Varni et al., 2003).

A trend in the larger QOL literature is to calculate the standard error of measurement to develop minimal clinically important difference (MCID) estimates in addition to cut-off scores for impairment (Varni, Burwinkle, Seid, & Skarr, 2003). MCIDs provide information as to the smallest amount of change on a clinical tool that would mandate a change in an individual's treatment in the absence of excessive costs or negative side effects (Jaschke et al., 1989). This feature has clear advantages to both clinicians and researchers in defining clinical improvement as opposed to clinical impairment. The MCID is most commonly calculated for broad and general measures, but if an empirically derived MCID is acceptable then the MCID is not limited to broad and general measures.

Disease-specific HRQOL assessment. As stated previously, disease-specific measures carry with them an implicit assumption that the experience of a particular illness conveys an impact to the individual's functioning that is both specific to the disease and unmeasured by broad and general instruments. Disease-specific measures use item phrasing that is intended to put the respondent in mind of their disease experience in order to measure QOL limitations specific to a particular clinical condition (e.g., "...found it hard to keep up with other kids because of your size"; Zeller & Modi, 2009). Advantages to this type of instrument are its ability to detect specific changes in QOL and increased clinical relevance to families.

The Food and Drug Administration has recently recognized the utility of disease-specific measures in clinical trials as evidenced by the use of a cystic fibrosis-specific measure of QOL as the primary endpoint in a phase III clinical trial of an

inhaled antibiotic designed to improve pulmonary functioning (Palermo et al., 2008; Retsch-Bogart et al., 2009). Disease-specific measures retain the multidimensionality of broad and general QOL measures while losing the comparability to large banks of normative data (Connolly & Johnson, 1999; Matza, Swensen, Flood, Secnik, & Leidy, 2004; Varni et al., 2005). Disease-specific measures exist for a host of clinical conditions including cancer, cystic fibrosis, juvenile rheumatoid arthritis, asthma, and pain (Goodwin, Boggs, & Graham-Pole, 1994; Juniper et al., 1996; Palermo, Witherspoon, Valenzuela, & Drotar, 2004; Quittner et al., 2005; Singh, Athreya, Fries, & Goldsmith, 1994).

Recommendations for QOL assessment. General QOL measures have clear strengths for gathering epidemiological data; but often lack the specificity to detect small changes in QOL that may be specific to the disease experience for a particular illness population (Palermo et al., 2008; Zeller & Modi, 2008). Conversely, disease-specific measures are thought to target a particular illness experience so narrowly that they can not be validly administered to children who do not have a shared illness experience. Therefore, assessing QOL using both general and disease-specific measures appears to be the ideal solution for retaining the ability to characterize a sample against normative data as well as monitor and address specific clinical improvements related to disease-specific symptoms (Revicki et al., 2000)

Weight-related QOL assessment. Relative to broad assessments of QOL, validation of disease-specific measures is still a relatively new area of study. In fact, there are currently only two self-report measures of weight-related QOL for children

and adolescents, and only one validation study has been devoted to each instrument. The Impact of Weight on Quality of Life-Kids (IWQOL-Kids[®]; Kolotkin et al., 2006) was designed for adolescents aged 11-19. The IWQOL-Kids is comprised of four scales: physical comfort ($\alpha = .91$), body esteem ($\alpha = .95$), social life ($\alpha = .92$), and family relations ($\alpha = .88$). The IWQOL-Kids represented the first attempt to develop a measure of weight-related QOL, and demonstrated good internal consistency and convergent validity with the *PedsQL*. As further evidence of validity, the IWQOL-Kids was inversely correlated with z-BMI across all scales, discriminated between children of differing weight status, and was sensitive to changes in z-BMI among children participating in a summer camp treatment program (Kolotkin et al., 2006; Quinlan, Kolotkin, Fuemmeler, & Costanzo, 2009).

A second weight-related QOL instrument, the *Sizing Me Up* questionnaire was developed by Zeller and Modi (2009) with the stated purpose of assessing weight-related QOL in school-aged children (5 to 13 years old). Zeller and Modi (2009) developed *Sizing Me Up* by examining the literature for information about QOL as a construct in pediatric populations and soliciting advice from experts in the field of pediatric obesity. Ultimately, 30 developmentally appropriate items thought to address physical functioning and discomfort, emotional functioning, peer relations and victimization, and social withdrawal were agreed upon and administered to 5-13 year-old children. Each item was phrased to guide children to consider how much a statement was true "...because of my size." The authors reported that children 10-years-old and younger were administered the questionnaire in interview format while

older children read and completed the items independently. The instrument demonstrated acceptable test-retest (intraclass correlation coefficient = .53-.78) and internal consistency statistics in a sample of 141 treatment-seeking 5-13 year old children with obesity. Results from an exploratory factor analysis indicated that a three to six factor solution was appropriate given the observed data. Using individual factor loadings and conceptual content of the items, the authors arrived at a 5-factor solution made up of 22-items consisting of emotion ($\alpha = .85$), physical ($\alpha = .76$), social avoidance ($\alpha = .70$), positive social attributes ($\alpha = .68$), and teasing/marginalization ($\alpha = .71$) scales. Correlations between the *Sizing Me Up* and the *PedsQL* were significant (Total QOL, $r = .52$; Maximum estimated correlation, $r = .85$).

Limitations of the current weight-related QOL assessment literature. As noted above, the literature on assessing weight-related QOL in children is small and a number of significant gaps exist in the current literature. First, the initial *Sizing Me Up* validation study was conducted exclusively in a treatment-seeking sample consequentially limiting the generalizability of the measure. As indicated by Zeller and Modi (2009), *Sizing Me Up* requires application to a nontreatment-seeking sample in order to establish its generalizability.

Second, Zeller and Modi (2009) used exploratory factor analysis for the initial validation study. Exploratory factor analysis is a data-driven approach in which all items are allowed to load on all constructs; as a result, constructs are formed by observing high loadings of individual items on a scale and relatively low loadings of

the same item on other scales (Tabachnick & Fidel, 2001). While exploratory factor analysis may be appropriate early in the measure development process, it does not provide a test of *a priori* hypothesized factor loadings of items on theoretically meaningful latent constructs. Therefore, confirmatory factor analytic techniques are considered superior to exploratory analyses because confirmatory factor analysis is a theory driven technique that involves constraining the data to fit a specified model (Brown, 2006). This approach is considered a more strenuous test of a measure's factor structure because individual items are forced to load on theoretically derived latent constructs and the total model is evaluated based on its fit to observed patterns in the data (Brown, 2006).

Third, the reliability estimates for the *Sizing Me Up* were low for some scales in the initial validation study. Perhaps contributing to this issue, a number of items evidenced significant cross-loadings in the exploratory analysis. For example, the item "Chose not to participate in gym because of your size" demonstrated a loading of .62 on the Social Avoidance dimension and a loading of .46 on the Physical dimension. These cross-loadings may indicate that some children do not make the distinction between "chose not to participate" and "could not participate" in gym class. It is possible that a confirmatory analysis of the measure could reveal an alternative factor structure that might yield improved reliability. Structural Equation Modeling (SEM) provides an appropriate framework for such a test.

Fourth, the initial evaluation of *Sizing Me Up* did not test the theoretical assumption that a weight-related QOL measure adds additional information to a QOL

assessment beyond what is available from a broad and general measure. In order to provide recommendations about when to use a disease-specific instrument, test developers should test their assumptions before disseminating a measure. Another benefit of a CFA within a SEM framework is that it allows for such a test.

Study Aims

Aim 1. The first aim of the current study was to examine the construct validity of *Sizing Me Up* in a community sample of 4th and 5th grade children with overweight and obesity using a confirmatory factor analysis in an SEM framework. As noted above, the literature is currently limited to an examination of the *Sizing Me Up* in treatment-seeking children. Specification of the factor structure in a community sample allows future studies to apply the *Sizing Me Up* measure in both prevention interventions and school-based interventions for children with overweight and obesity as well as a clinical tool at the individual level. This aim partially answers the question, “*Are the scoring conventions established in a treatment seeking sample appropriate for nontreatment-seeking samples?*” Thus, the first hypothesis of the current study was that: a) a five-factor structure with a single second-order construct consistent with Zeller and Modi underlies *Sizing Me Up* in a nontreatment-seeking sample; and b) the factor structure of *Sizing Me Up* is not appropriate for use in healthy weight children.

The criterion-related validity of *Sizing Me Up* was assessed by specifying predictable associations between the five factors from hypothesis one and BMI percentile, and modeling the association between BMI percentile and each construct

measured. The associations were entered into the model as regression paths using BMI percentile to predict each latent construct of the *Sizing Me Up* measure. It was hypothesized that BMI%ile was associated with poorer weight-related QOL as indicated by significant and positive associations between BMI%ile and the *Sizing Me Up* physical scale; emotion scale; social scale; and the teasing and marginalization scale; and a significant negative association between BMI%ile and the positive social attributes scale. Overall, it was hypothesized that BMI%ile was significantly and positively associated with the *Sizing Me Up* single second-order factor (i.e., total score).

To further examine construct validity, the current study tested the pattern of convergent validity of *Sizing Me Up* by examining significant associations between latent *Sizing Me Up* factors identified by Zeller and Modi (2009), and, among the intercorrelations of the latent factors underlying *Sizing Me Up* and the *PedsQL*. By using a CFA framework to conduct this test the current study advances the literature by evaluating the associations of theoretically similar constructs modeled without measurement error, thereby providing a purer estimate of the true intercorrelation between the constructs (Brown, 2006). Based upon the correlations reported by Zeller and Modi (2009), it was hypothesized that all of the *Sizing Me Up* factors are significantly moderately correlated except for positive social attributes, which should only be correlated with social avoidance. Additionally, it was hypothesized that the *Sizing Me Up* scales demonstrate good convergent validity with the *PedsQL* scales as evidenced by small to moderate positive associations between the physical *Sizing Me*

Up scale and the physical *PedsQL* scale, the emotion *Sizing Me Up* scale and *PedsQL* emotion scale, the social avoidance *Sizing Me Up* scale and the social *PedsQL* scale, the teasing and marginalization *Sizing Me Up* scale and the social *PedsQL*, and the total scores of the *PedsQL* and *Sizing Me Up*.

Aim 2. With aim one achieved, it was possible to test the implicit theoretical assumption that the experience of overweight creates unique QOL experiences independent from general QOL. Analyses addressing aim two answers the question, “*Are QOL and weight-related QOL different theoretical constructs that clinicians and researchers should measure independently among community samples of overweight and obese children?*” If weight-related QOL is a different construct than general QOL, then introducing model constraints to fix the latent correlation between weight-related QOL and general QOL physical, emotional, and social constructs to 1.0 should result in significant model misfit. It was hypothesized that the following scale pairs do not measure unitary QOL constructs; the *Sizing Me Up* physical scale and the *PedsQL* physical scale; the *Sizing Me Up* emotion scale and the *PedsQL* emotion scale; the *Sizing Me Up* social avoidance scale and the *PedsQL* social scale; the *Sizing Me Up* teasing and marginalization scale and the *PedsQL* social scale; the second-order weight-related QOL and general QOL scales that underlie *Sizing Me Up* and the *PedsQL*, respectively.

Method

Participants

Participants were convenience sample of 4th and 5th grade students enrolled in one of six Lawrence Public Schools. Parental consent and child assent were obtained from 307 participant families. However, height and weight data were not available for five children. The final sample for analysis included 302 participants. For the purposes of the current study participants were categorized into overweight and obese and healthy weight groups¹. The healthy weight group was comprised of 168 participants with a mean Body Mass Index percentile (BMI%ile) of 50.6 (SD = 23.6). Mean age of participants in this group was 10.34 (SD = .76). The healthy weight group was approximately evenly divided between males and females (56.5% female and 43.5% male). The group was predominantly Caucasian with 66.6% identifying as White not Hispanic, 3.6% Black not Hispanic, 6.5% Hispanic, 4.2% Asian, 6.5% American Indian, 8.9% other, and 3.7% who chose not to report their race/ethnicity. The overweight and obese group was comprised of 134 participants with a mean BMI%ile of 94.4 (SD = 4.3). Mean age of participants in this group was 10.33 (SD = .69). Again, participants were approximately evenly divided between males and females (56.0% male and 44.0% female). The group was predominantly Caucasian with 54.4% identifying as White not Hispanic, 7.5% Black not Hispanic, 6.7%

¹ The healthy weight group in the current sample did have 9 participants (i.e., 3% of the sample) who would be classified as underweight (i.e., BMI%ile < 5th percentile). These participants were retained in the sample to maximize statistical variance.

Hispanic, 8.2% Asian, 4.5 American Indian, 14.9% other, and 3.8% who chose not to report their race/ethnicity.

Table 1

Demographic Characteristics

Demographics	BMI%ile \geq 85 ($n = 134$)	BMI%ile $<$ 85 ($n = 168$)
Age	10.22 ($SD = .69$)	10.34 ($SD = .76$)
Male	56.0%	43.5%
Female	44.0%	56.5%
BMI%ile	94.4 ($SD = 4.3$)	50.6 ($SD = 23.6$)
White not Hispanic	54.4%	66.6%
Black not Hispanic	7.5%	3.6%
Hispanic	6.7%	6.5%
Asian	8.2%	4.2%
American Indian	4.5%	6.5%
Other	14.9%	8.9%
Did not report	3.8%	3.7%

Sampling Procedure and Questionnaire Administration

Following receipt of approval from the University of Kansas Institutional Review Board, Unified School District 497, and respective building principals and classroom teachers, 4th and 5th grade students were recruited from six elementary schools in Lawrence, Kansas. Students interested in participating were given a

consent form that was to be completed by the child's parent before the child could participate in the proposed study. In addition, children were informed that they were not required to participate in the study, and were given the opportunity to provide assent. Of those approached, 88.7% ($n = 307$) provided parental consent and participant assent to participate. Participants were gathered in a cafeteria or classroom during a convenient time determined by school personnel. Survey packets containing study questionnaires and other instruments part of a larger evaluation of self-esteem, physical activity, and body image were distributed. Assent scripts were read to the students, who then indicated assent by circling "yes" on the form. After assenting, participants were asked to write their name on a page that was later removed from the rest of the packet (this page was used to link BMI data with questionnaire responses). Following this step, only a unique study identification number was used to identify participants. Students were then asked to complete the study measures. Research assistants were available to read measures to self-identified students requiring this accommodation. As part of the Unified School District 497 annual health assessment, height and weight measurements of each child were collected by a school nurse; these data were obtained and used to calculate BMI percentile.

Power analysis. An *a priori* power analysis was calculated using a power calculator developed by Preacher and Coffman (2006) based on the formula for determining good model fit proposed by McCallum, Brown, and Sugawara (1996). The proposed study tested a number of different models to determine the best fitting

model for each instrument. Therefore, three power calculations are presented using the most constrained proposed test for each model (i.e., lowest degrees of freedom). All of the following power calculations are based on a power estimate of .80, which is generally considered to be sufficient power to detect a statistical effect in SEM (Muthén & Muthén, 2002) and alpha levels of .05. Power analysis for the initial confirmatory factor analysis of the *Sizing Me Up* was based on 199 degrees of freedom where $df = [p(p + 1)/2] - q$ where p = manifest variables and q = unknown parameter estimates. Results of the power analysis for the *Sizing Me Up* measurement model indicate that 86 participants were required to achieve a close fit to the data if close fit between the model and the data were achievable. Results of the power analysis for the *PedsQL* measurement model ($df = 220$) indicated that 80 participants would be required to achieve a close fit to the data if close fit were achievable given the data. Finally, results of the power analysis for the measurement model examining the intercorrelations between the *Sizing Me Up* and the *PedsQL* measurement model ($df = 900$) indicate that 35 participants would be required to achieve a close fit to the data if close fit was present in the data. Thus the current sample of 134 children with overweight and obesity and 168 healthy weight children was sufficient for a robust test of the stated hypotheses.

Measures

Anthropometric data. Overweight and obesity are labels for adiposity; however, adiposity is impractical to directly measure in children and Body Mass Index (BMI) is considered an acceptable proxy (Barlow, 2007). BMI is expressed as

body weight in kilograms divided by height in meters squared (kg/m^2). BMI does not increase linearly across sex and age throughout childhood and adolescence.

Therefore, normative data are used to standardize individual scores before categorizing children as overweight or obese. In order to calculate BMI%ile, the U.S. Center for Disease Control (2007) growth charts are used to plot each child's weight and height to determine their BMI%ile score. A BMI%ile score can then be used to classify children as underweight (i.e., $<5^{\text{th}}$ percentile), healthy weight (5^{th} - 84^{th} percentile), overweight (85^{th} - 94^{th} percentile), or obese ($\geq 95^{\text{th}}$ percentile) as recommended by the American Academy of Pediatrics (Barlow, 2007). BMI categories are a reliable and valid predictor of current and subsequent health problems in children (Dietz & Bellizzi, 1999). In the current study, children's date of birth and height and weight were collected from school records and used to compute a BMI%ile score.

Weight-related QOL. Weight-related QOL was assessed using the 22-item *Sizing Me Up* self-report questionnaire (see Appendix A) designed for use with 5-13 year old children (Zeller & Modi, 2009). The instrument is made up of items that orient the participant to the weight-related component of the assessment by asking how much the item is true during the past month "...because of your size." Participants respond to the questionnaire using a ordinal scale with anchors of none of the time (1), a little (2), a lot (3), and all the time (4). As noted above *Sizing Me Up* has acceptable reliability estimates and evidence of convergent validity with the *PedsQL* in a treatment seeking sample. Reliability statistics for the five factor

solution are: physical $\alpha = .76$, emotion $\alpha = .85$, social avoidance $\alpha = .70$, positive social attributes $\alpha = .68$, teasing and marginalization $\alpha = .71$, and total score $\alpha = .82$. The *Sizing Me Up* measure has a Flesch-Kincaid readability index score of 2.1.

Broad and General QOL. The *PedsQL* (Varni et al., 2001) is a 23-item questionnaire that measures self-reported QOL using questions designed to assess how much each item has been a problem for the child in the last month. The *PedsQL* uses an ordinal scale with anchors of never, almost never, sometimes, often, and almost always. Previous studies of the *PedsQL* have found evidence for four- and five-factor solutions (Varni, Limbers, Newman, & Seid, 2008; Varni et al., 2001). The four factor solution is comprised of Physical ($\alpha = .80$); Emotional ($\alpha = .73$); Social ($\alpha = .71$); and School ($\alpha = .68$) QOL scales, and the five-factor solution takes two items from the School scale to create a Medical-School scale (Varni et al., 2008). The measure has demonstrated good internal consistency and appears to discriminate appropriately between well and ill groups (Varni et al., 2001; Varni et al., 2003).

Missing data. The proctored administration format described above reduces the likelihood of missing data compared to questionnaires administered via mail or proctored by a teacher or other school official. However, some children accidentally omitted items or had difficulty keeping up with the pace of the administration leading to randomly missing data. As stated above, all of the analyses in the proposed project were conducted in the SEM framework. Traditional *ad hoc* methods of dealing with missing data such as listwise deletion and mean replacement are known to produce biased parameter estimates as evidenced by the results of simulation studies (Graham,

Hofer, & MacKinnon, 1996; Schafer & Graham, 2002); on the other hand, imputation procedures are known to produce more accurate estimates when missing values are Missing at Random (MAR) or Missing Completely at Random (Schafer & Graham, 2002).

Rubin (1976) described the condition of MAR as a special case of missingness in which the value of variable *Y* may be dependent on variable *X* but not on other values of variable *Y*. The most likely missing data pattern in the proposed study was Missing Completely at Random (MCAR). MCAR is a case of missingness where variable *Y* is not related to any other measured variable in the dataset or other values of variable *Y* (Rubin, 1976). For example, if children accidentally skip items or fail to answer a page of items because they fail to notice the items, these missing values would be MCAR. A very small number of values were missing from the final dataset (i.e., 1.18%). The observed variables provided by *Sizing Me Up* and the *PedsQL* are ordinal; and, a complete dataset is necessary to output the polychoric correlation matrix for SEM analysis. Therefore, the Expectation Maximization (EM) algorithm was employed in the PRELIS program bundled with Lisrel 8.8 (Jöreskog & Sörbom, 2006) to achieve a single complete dataset. Imputation using the EM algorithm is a method of stochastic imputation considered to be consistent with the best statistical practices in applied psychology (Schlomer, Bauman, & Card, 2010). Due to the small amount of missing data and the assumption of MCAR only a single imputation using the EM algorithm was necessary (Schafer 1999).

Results

Data Screening

Initial data screening revealed that the variables included in both *Sizing Me Up* and the *PedsQL* were not normally distributed (see Table 1). Specifically, several items were significantly positively skewed. Therefore, the data were modeled using Robust Maximum Likelihood (RML) and evaluated using the Satorra-Bentler χ^2 scaled test of model fit. Additionally, data screening revealed that the *PedsQL* included one item with no variance to the measure. Specifically, responses to the item “It is hard for me to take a bath or shower by myself” were uniformly “never.” Because ordinal data require the use of a polychoric correlation matrix, it is not possible to produce estimates for variables with fewer than two response choices (i.e., singularly zero). Therefore, this item was eliminated from all analyses using the *PedsQL*.

Table 2

Descriptive Statistics for Sizing Me Up and the PedsQL

<i>Sizing Me Up</i>			<i>PedsQL</i>		
SMU Item	<i>M</i>	<i>SD</i>	<i>PedsQL</i> Item	<i>M</i>	<i>SD</i>
SMU 1	1.24	0.48	<i>PedsQL</i> 1	0.29	0.75
SMU 2	1.57	0.84	<i>PedsQL</i> 2	0.38	0.67
SMU 3	2.70	0.93	<i>PedsQL</i> 3	0.30	0.64
SMU 4	1.47	0.86	<i>PedsQL</i> 4	0.67	0.94
SMU 5	1.18	0.52	<i>PedsQL</i> 5*	0.00	0.00
SMU 6	1.13	0.43	<i>PedsQL</i> 6	0.22	0.59
SMU 7	2.68	1.04	<i>PedsQL</i> 7	0.68	0.96
SMU 8	1.99	1.04	<i>PedsQL</i> 8	0.49	0.82
SMU 9	1.32	0.63	<i>PedsQL</i> 9	0.69	0.95
SMU 10	1.41	0.72	<i>PedsQL</i> 10	0.75	0.96
SMU 11	1.10	0.42	<i>PedsQL</i> 11	0.92	1.14
SMU 12	1.15	0.51	<i>PedsQL</i> 12	0.84	1.22
SMU 13	2.63	1.02	<i>PedsQL</i> 13	0.93	1.26
SMU 14	1.90	0.95	<i>PedsQL</i> 14	0.58	0.94
SMU 15	1.19	0.53	<i>PedsQL</i> 15	0.61	0.96
SMU 16	2.70	0.95	<i>PedsQL</i> 16	0.59	0.94
SMU 17	1.17	0.54	<i>PedsQL</i> 17	0.47	0.84
SMU 18	1.22	0.65	<i>PedsQL</i> 18	0.38	0.76

SMU 19	1.22	0.56	<i>PedsQL</i> 19	0.78	1.07
SMU 20	1.40	0.67	<i>PedsQL</i> 20	1.22	1.20
SMU 21	1.57	0.72	<i>PedsQL</i> 21	0.63	1.03
SMU 22	1.04	0.21	<i>PedsQL</i> 22	0.97	1.05
			<i>PedsQL</i> 23	0.75	1.00

Note. The *Sizing Me Up* is on a 1-4 point scale while the *PedsQL* is on a 0-4 point scale. * Item deleted for all analyses.

Overview of analyses. The null and alternative models were specified using LISREL 8.8 (Jöreskog & Sörbom, 2006). Since the manifest data collected in the current study were ordinal and skewed, the polychoric correlation matrix with an asymptotic covariance matrix was analyzed in all structural models. SEM has a number of advantages over other factor analytic techniques for conducting CFA. As noted above, SEM provides the researcher with a flexible platform for handling missing data. Additionally, SEM benefits from the ability to control for measurement error, estimate latent constructs, and apply measurement constraints to test the equivalence of factor structures across different groups (Brown, 2006). All models were evaluated by examining the Sartorra-Bentler chi-square test of significance, comparative fit index (CFI), non-normed fit index (NNFI), and the root mean squared error of approximation (RMSEA). Model fit was considered to be acceptable if the CFI and NNFI were above .90 and the RMSEA was below .1. For nested model comparisons, chi-square change tests were considered significant at the $p < .05$ level. An *a priori* decision was made to include all *Sizing Me Up* items in the final models

even if factor loadings were not significant. This decision was made because the current study was focused on questions of generalizability and theoretical significance. Measure revision was not a goal of the current project.

All reliability statistics were taken from the final model with nine first-order and two second-order factors and are calculated using the formula $\rho = (\sum\lambda_i)^2 / [(\sum\lambda_i)^2 + (\sum\theta_i)^2]$ (where λ = the unstandardized factor loadings and θ = the unstandardized error terms) as this method provides an estimate of the true scale reliability in a CFA framework (Raykov, 2004). In the case of the second-order factors, error terms are replaced with the error variance of the first-order factors and the factor loadings are replaced with the disattenuated loadings of first-order factors on the second-order factor (Ping, 2004). This method of reliability estimation is free from many of the biases that are present in traditional methods of estimating scale reliability such as Chronbach's α . That is, all reliability estimates attempt to approximate true score reliability. However, Chronbach's α treats all item level covariance as true score covariance. This is not appropriate because, in fact, observed an observed covariance contains both true score variability and random variability leading to biased estimates. Another problem with the traditional Chronbach's α is that a scale with a small number of items will produce biased reliability estimates. The ρ reliability estimate solves this problem by including only true score information in the numerator and ignoring scale size altogether.

Finally, a Minimal Clinically Important Difference (MCID) score was calculated to give clinicians a guideline for the magnitude of change in a given scale

that would indicate that a change in treatment was clinically relevant (Jaeschke et al., 1989). Consistent with other studies of QOL assessment in the literature, the MCID was calculated by taking the square root of the result of one minus the internal consistency (i.e., ρ) times the standard deviation (i.e., Standard Error of Measurement; Varni et al., 2003; Wyrwich, Tierney, & Wolinsky, 1999).

Aim 1

As noted above, Aim 1 of the current study was to evaluate the construct validity of *Sizing Me Up* in a community sample of 4th and 5th grade children with overweight and obesity using a CFA in an SEM framework. Specifically, this aim addresses convergent validity both within *Sizing Me Up* factors and between *Sizing Me Up* factors and similar factors on the *PedsQL*. Additionally, criterion validity was assessed by examining the association between BMI%ile and *Sizing Me Up* factors.

Null model. The null model was specified as all 22 *Sizing Me Up* items with error terms freely estimated and factor loadings fixed to 0.0. BMI percentile and gender were included as exogenous variables specified by a single indicator to allow for regression tests using these variables in the alternative model. This is the mathematical equivalent of the statement, “*Sizing Me Up* items do not measure latent constructs and capture only error.” The null model evidenced poor fit to the data, $\chi^2(265, n = 134) = 1890.00, p < 0.001$. The Aim 1 alternative model was assessed using this null model.

Alternative model. The five factor model of *Sizing Me Up* was estimated using the RML estimator with a ridge constraint of 1.0. As stated previously, the

polychoric correlation matrix was used to account for the violation of the assumption of continuous data. The asymptotic covariance matrix was used to account for non-normality in the observed data. Results of the five factor solution revealed close fit to the data, $\chi^2(236, n = 134) = 311.58, p < 0.001$, RMSEA = .049, CFI = .96, NNFI = .95. The Satorra-Bentler χ^2 improved significantly compared to the null model, $\Delta\chi^2 = 1578.42, p < .05$. All of the lambda loadings were significant except for item number 8 “Stood up for or helped other kids because of your size.” Overall, results indicated that Zeller and Modi’s (2009) factor structure is adequate for use in community samples of children with overweight and obesity (see Table 3 for loadings and errors).

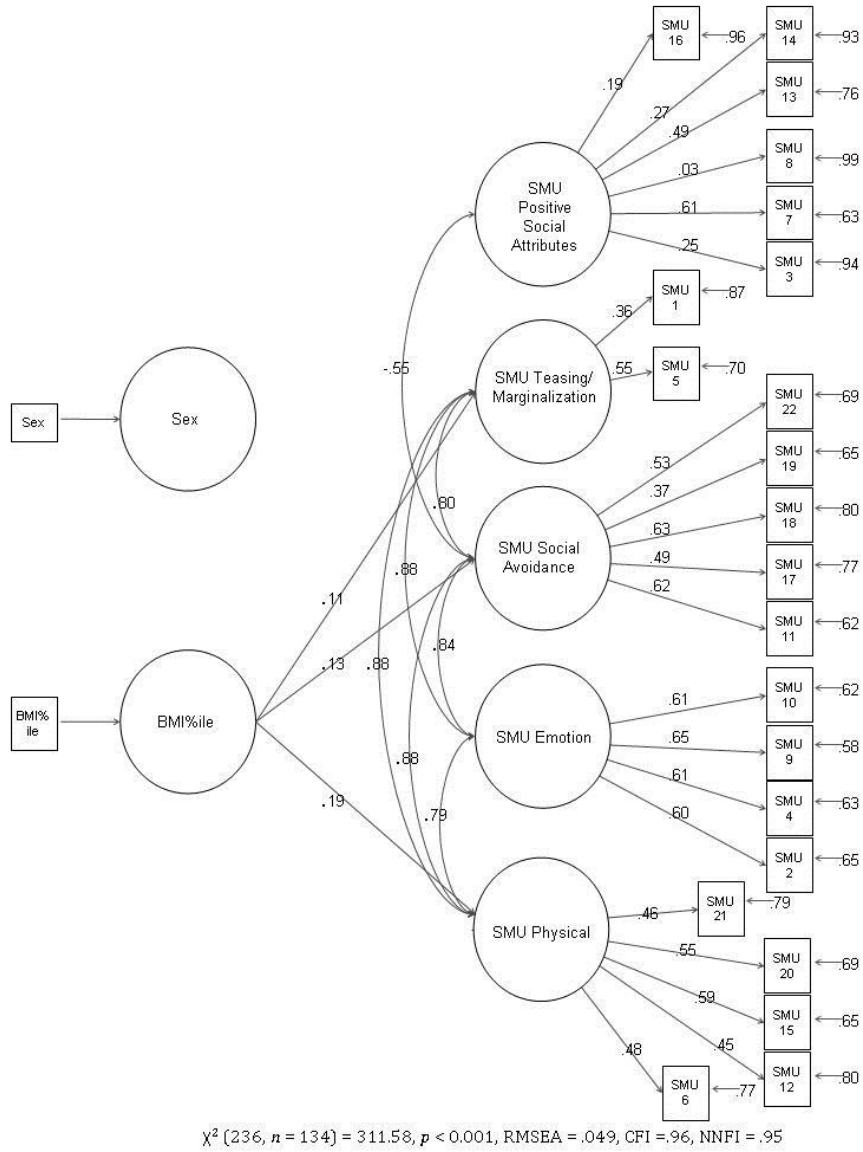
Using the same null and alternative model, responses from the 168 healthy weight subjects were entered as an independent group to test the assumption of factorial invariance across healthy and unhealthy weight categories. The loadings and intercepts were constrained across healthy weight and overweight groups in sequential steps. Based on the RMSEA test (i.e., does the 90% RMSEA confidence interval overlap with the 90% RMSEA confidence interval of the alternative model; Little, 1997) both constraints were untenable. These tests indicate that neither the factor loadings nor intercepts of *Sizing Me Up* are meaningfully similar across healthy weight and overweight groups. The results above provide support for hypothesis one, and suggest that scores on the *Sizing Me Up* should not be compared across children with healthy weight and children with overweight and obesity.

Gender and BMI percentile were included in the model as independent exogenous latent variables predicting all five of the previously specified endogenous

latent variables. None of the paths from gender to any of the other latent constructs were significant. Partial support for hypothesis three was observed as: a) the regression path from BMI percentile to the physical *Sizing Me Up* scale was small but significant ($\gamma = .19$); b) the path from BMI percentile to the *Sizing Me Up* social scale was small but significant ($\gamma = .13$); and c) the path from BMI percentile to the *Sizing Me Up* teasing and marginalization scale was small but significant ($\gamma = .11$). Therefore, criterion validity was established with three of the five first-order *Sizing Me Up* scales. The nonsignificant regression paths identified in this analysis were dropped from subsequent models to allow for more degrees of freedom (see Figure 1).

Figure 1

Latent Correlations and Regression Paths for Aim 1 Alternative Model



Note. BMI%ile and sex are measured using a single item. Therefore, loadings and errors are fixed to 1.0 and 0.0, respectively.

Table 3

Estimated and Standardized Factor Loadings, Residuals, and R² Values for Each Sizing Me Up Indicator

Indicator	Estimated	Standardized		
	Loading (SE)	Loading	Theta	R ²
Physical				
SMU 6	.66 (.13)	0.48	0.77	0.23
SMU 12	.62 (.14)	0.45	0.80	0.20
SMU 15	.82 (.08)	0.59	0.65	0.35
SMU 20	.77 (.08)	0.55	0.69	0.31
SMU 21	.64 (.11)	0.46	0.79	0.21
Emotion				
SMU 2	.83 (.05)	0.60	0.65	0.35
SMU 4	.85 (.05)	0.61	0.63	0.38
SMU 9	.90 (.05)	0.65	0.58	0.42
SMU 10	.85 (.05)	0.61	0.62	0.38
Social Avoidance				
SMU 11	.84 (.08)	0.62	0.62	0.38
SMU 17	.67 (.10)	0.49	0.77	0.24
SMU 18	.85 (.06)	0.63	0.80	0.39
SMU 19	.51 (.13)	0.37	0.65	0.14
SMU 22	.73 (.12)	0.53	0.69	0.29

Positive Social				
Attributes				
SMU 3	.35 (.10)	0.25	0.94	0.06
SMU 7	.86 (.08)	0.61	0.63	0.37
SMU 8	.04 (.11)	0.03	0.99	0.00
SMU 13	.69 (.08)	0.49	0.76	0.24
SMU 14	.38 (.11)	0.27	0.93	0.07
SMU 16	.27 (.14)	0.19	0.96	0.04
Teasing and				
Marginalization				
SMU 1	.50 (.14)	0.36	0.87	0.13
SMU 5	.75 (.18)	0.55	0.70	0.30

Note. Standardized estimates are taken from the completely standardized solution.

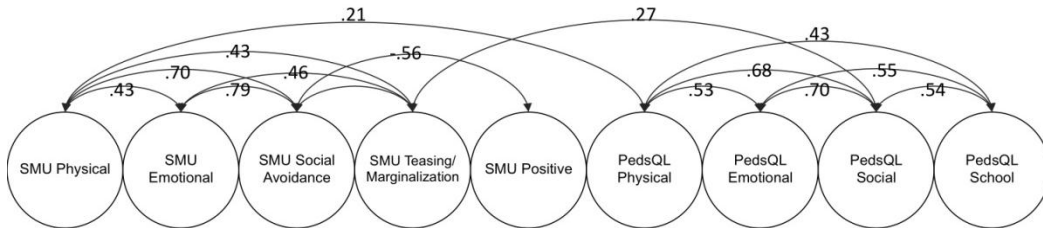
Nine-factor null model. In order to assess the associations between *Sizing Me Up* and the *PedsQL*, a null model was specified with 44 items from *Sizing Me Up* and the *PedsQL*. As stated previously, one *PedsQL* item was excluded due to restriction of range. The null model was specified as all 44 manifest variables with error terms freely estimated and factor loadings fixed to 0.0. This is the mathematical equivalent of the statement, “*Sizing Me Up* and *PedsQL* items do not measure latent constructs and capture only error.” The null model evidenced poor fit to the data, $\chi^2(1016, n = 134) = 6489.77, p < 0.001$. All nine-factor alternative models were assessed using this null model.

Nine-factor alternative model. The nine-factor model including both *Sizing Me Up* and the *PedsQL* was estimated using the RML estimator with a ridge constraint of 1.0. Again, the polychoric correlation matrix was used to account for the violation of the assumption of continuous data and the asymptotic covariance matrix was used to account for non-normality in the observed data. As in the previous model, BMI%ile and sex were allowed to enter the model freely. Only the three paths identified as significant in the evaluation of *Sizing Me Up* alone remained significant. Nonsignificant regression paths were pruned from the final nine-factor model. Results of the final nine factor solution revealed acceptable fit to the data $\chi^2(974, n = 134) = 1453.50, p < 0.001, RMSEA = .061, CFI = .91, NNFI = .91$.

Convergent validity between QOL scales. To assess convergent validity between the two QOL scales, latent intercorrelations between *Sizing Me Up* factors and *PedsQL* scales were estimated as well as latent correlations between *Sizing Me Up* factors. The hypothesis of convergent validity was partially supported (see Figure 2). The significant latent correlations among *Sizing Me Up* factors identified by Zeller and Modi (2009) were replicated in the current sample. The correlation between the physical ($\psi = .22$) scales and the *PedsQL* social scale and *Sizing Me Up* teasing and marginalization scale ($\psi = .27$) were significant. Hypothesized intercorrelations stated between the *Sizing Me Up* emotional and social avoidance scales and *PedsQL* social scale were not significant.

Figure 2

Latent Intercorrelations of the Aim 2 Nine-Factor Model



$\chi^2(974, n = 134) = 1453.50, p < 0.001, RMSEA = .061, CFI = .91, NNFI = .91$

Note. For clarity factor loadings and standard errors are presented in Table 4.

Table 4

Estimated and Standardized Factor Loadings, Residuals, and R² Values for Each Sizing Me Up and PedsQL Indicator

Indicator	Estimated	Standardized	Theta	R ²
	Loading (SE)	Loading		
Sizing Me Up				
Physical				
SMU 6	.65 (.15)	0.48	0.77	0.23
SMU 12	.53 (.17)	0.38	0.85	0.15
SMU 15	.70 (.11)	0.52	0.73	0.27
SMU 20	.70 (.09)	0.50	0.75	0.25
SMU 21	.58 (.14)	0.43	0.82	0.18
Emotion				
SMU 2	.78 (.05)	0.56	0.68	0.32
SMU 4	.81 (.05)	0.58	0.66	0.34
SMU 9	.88 (.06)	0.63	0.60	0.40
SMU 10	.84 (.05)	0.60	0.64	0.36
Social Avoidance				
SMU 11	.83 (.08)	0.61	0.63	0.37
SMU 17	.67 (.10)	0.49	0.76	0.24
SMU 18	.82 (.06)	0.59	0.65	0.35
SMU 19	.47 (.14)	0.34	0.89	0.11

SMU 22	.80 (.12)	0.58	0.67	0.33
Positive Social Attributes				
SMU 3	.35 (.10)	0.25	0.94	0.06
SMU 7	.88 (.08)	0.62	0.61	0.39
SMU 8	.05 (.11)	0.04	0.99	0.00
SMU 13	.68 (.08)	0.48	0.77	0.23
SMU 14	.37 (.11)	0.26	0.93	0.07
SMU 16	.25 (.14)	0.18	0.97	0.03
Teasing and Marginalization				
SMU 1	.73 (.18)	0.53	0.72	0.28
SMU 5	.50 (.18)	0.36	0.87	0.13
<i>PedsQL</i>				
Physical				
<i>PedsQL</i> 1	.63 (.11)	0.44	0.80	0.20
<i>PedsQL</i> 2	.86 (.06)	0.62	0.62	0.38
<i>PedsQL</i> 3	.75 (.09)	0.54	0.71	0.29
<i>PedsQL</i> 4	.44 (.12)	0.31	0.90	0.10
<i>PedsQL</i> 6	.52 (.12)	0.37	0.86	0.14
<i>PedsQL</i> 7	.62 (.09)	0.44	0.81	0.19
<i>PedsQL</i> 8	.70 (.08)	0.50	0.76	0.25
Emotional				

<i>PedsQL</i> 9	.73 (.07)	0.51	0.74	0.26
<i>PedsQL</i> 10	.80 (.05)	0.56	0.68	0.32
<i>PedsQL</i> 11	.58 (.08)	0.41	0.83	0.17
<i>PedsQL</i> 12	.64 (.07)	0.45	0.80	0.21
<i>PedsQL</i> 13	.80 (.06)	0.57	0.68	0.32
Social				
<i>PedsQL</i> 14	.63 (.08)	0.45	0.80	0.21
<i>PedsQL</i> 15	.69 (.07)	0.50	0.75	0.25
<i>PedsQL</i> 16	.66 (.08)	0.48	0.77	0.23
<i>PedsQL</i> 17	.71 (.07)	0.51	0.74	0.26
<i>PedsQL</i> 18	.75 (.07)	0.55	0.70	0.30
School				
<i>PedsQL</i> 19	.66 (.08)	0.47	0.78	0.22
<i>PedsQL</i> 20	.76 (.07)	0.54	0.71	0.29
<i>PedsQL</i> 21	.57 (.11)	0.40	0.84	0.16
<i>PedsQL</i> 22	.60 (.08)	0.43	0.82	0.18
<i>PedsQL</i> 23	.52 (.10)	0.37	0.87	0.13

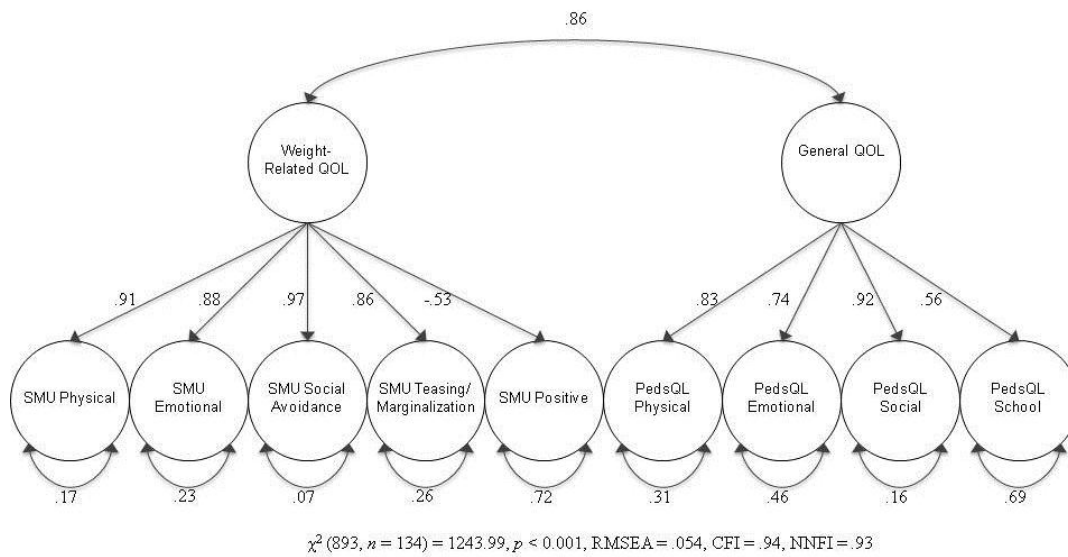
Note. Standardized estimates are taken from the completely standardized solution.

Second-order factor structure. To test the overall construct validity of weight-related QOL and determine the utility of a total score for *Sizing Me Up*, second order weight-related QOL and general QOL constructs were specified. Weight-related QOL was made up of the five latent constructs derived from *Sizing Me Up*. General

QOL was made up of the four latent constructs derived from the *PedsQL*. The two-factor second-order model demonstrated close fit to the data, χ^2 (892, $n = 134$) = 1171.66, $p < 0.001$, RMSEA = .049, CFI = .95, NNFI = .94. Similar to the results from the first-order structure, sex was not associated with either second-order factor and BMI%ile was significantly associated with the weight-related QOL factor (i.e., *Sizing Me Up* total score; $\gamma = .11$) and the general QOL total score (i.e., *PedsQL* total score; $\gamma = .09$). Due to a relatively low loading of the positive social attributes scale ($\gamma = -.53$) a three-factor higher order model with positive social attributes as a unique second-order factor was tested to ensure that the five *Sizing Me Up* scales represent a unitary construct. This model demonstrated significantly worse fit to the data ($\Delta\chi^2 = 3.86$, $p < .05$) and was less parsimonious than the two-factor model, χ^2 (891, $n = 134$) = 1175.52, $p < 0.001$, RMSEA = .049, CFI = .95, NNFI = .94, and was rejected.

Figure 3

Second-Order Factor Structure



Aim 2

To address the implicit theoretical assumption that overweight and obesity confers a unique impairment on QOL, the nine-factor model with significant BMI%ile regressions was used. To allow for testing of nested models, the measurement model for the current test included nonsignificant estimates for the latent correlations between the two different QOL social and emotional scales. The measurement model demonstrated acceptable fit to the data, $\chi^2 (972, n = 134) = 1465.61, p < 0.001, RMSEA = .062, CFI = .91, NNFI = .91$. The hypothesis that *Sizing Me Up* measures unique weight-related QOL constructs was supported. Specifically, when the latent correlations of each corresponding pair of scales for the *Sizing Me Up* and *PedsQL* were constrained to 1.0 the resulting nested model chi-

square comparison indicated that the constraint was untenable due to significant change in the chi-square statistic (*Sizing Me Up* physical and *PedsQL* physical, $\Delta\chi^2 = 129.35, p < .05$; *Sizing Me Up* emotion and *PedsQL* emotion, $\Delta\chi^2 = 226.38, p < .05$; *Sizing Me Up* social avoidance *PedsQL* social, $\Delta\chi^2 = 141.43, p < .05$; *Sizing Me Up* teasing and marginalization and *PedsQL* social, $\Delta\chi^2 = 9.5, p < .05$). Finally, in order to test the hypothesis that weight-related QOL and general QOL are distinct constructs, the latent the second-order measurement model was used and the latent correlation between the two higher order QOL constructs was fixed to 1.0. The nested model chi-square comparison indicated that the constraint was untenable due to a significant change in the chi-square statistic ($\Delta\chi^2 = 30.05, p < .05$).

Transformed Means and Standard Deviations

In order to aid with interpretation, the *PedsQL* and *Sizing Me Up* were each transformed to a 0-100 scale as recommended by their respective authors. Scaled means and standard deviations are available in Table 5. One method of establishing cut-off scores for population level QOL measures is to subtract one standard deviation from the total mean score (Varni et al., 2003). Following a similar procedure, the current *Sizing Me Up* mean score minus one standard deviation ($78.30 - 10.93 = 67.37$) was almost identical to the mean reported in Zeller and Modi's (2009) initial validation study (~68) of treatment-seeking children with overweight or obesity.

Table 5

Scaled Means and Standard Deviations

Scale	<i>M</i>	<i>SD</i>
SMU Physical	90.44	12.79
SMU Emotional	85.25	21.30
SMU Social Avoid.	95.00	11.87
SMU Positive Social	47.80	18.53
SMU Teas/Marg.	93.07	13.12
SMU Total	78.30	10.93
<i>PedsQL</i> Physical	89.20	12.34
<i>PedsQL</i> Emotional	85.25	21.30
<i>PedsQL</i> Social	86.83	16.44
<i>PedsQL</i> School	78.25	18.76
<i>PedsQL</i> Total	83.94	12.64

Table 6

Scaled Means and Standard Deviations by Weight Category

Scale	Overweight (<i>n</i> =56)		Obese (<i>n</i> = 62)		Very Obese (<i>n</i> = 16)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SMU Physical	92.78	12.00	89.30	11.88	86.69	17.55
SMU Emotional	86.94	22.85	84.60	20.24	81.86	20.40
SMU Social Avoid.	95.38	14.02	95.19	10.16	92.95	10.16
SMU Positive Social	49.03	17.59	49.01	18.32	38.47	20.94
SMU Teas/Marg.	94.66	12.32	93.60	11.00	85.47	20.10
SMU Total	79.82	11.25	78.34	10.26	72.83	11.21
<i>PedsQL</i> Physical	90.94	9.92	88.88	13.02	84.38	16.28
<i>PedsQL</i> Emotional	79.29	20.75	78.95	20.15	81.25	18.93
<i>PedsQL</i> Social	87.59	16.35	86.45	17.40	85.63	13.40
<i>PedsQL</i> School	80.71	17.25	76.37	21.22	76.88	12.63
<i>PedsQL</i> Total	85.20	11.28	83.23	14.08	82.24	11.57

Note. The 0-100 point scales for each measure are derived by linearly transforming the item level data such that 0 indicates poorer QOL and 100 indicates higher quality QOL.

Reliability of First and Second Order Factors and Minimal Clinically Important Difference

Reliability statistics (calculated as ρ to provide factor reliability statistics) for the current study were marginally acceptable for all of the first-order factors except for the *Sizing Me Up* positive social attributes and teasing and marginalization scales. However, the total score reliability estimate was much higher for both the *Sizing Me Up* and *PedsQL* total scores than for their subscales. MCIDs are provided for all of the factors examined in the nine-factor and second-order models; however, relatively lower reliability estimates for the first-order factors make the second-order MCIDs the most meaningful estimate of clinically significant changes. Previous data is not available for MCID for *Sizing Me Up*; however, the total score MCID identified for the *PedsQL* is consistent with previous reports (e.g., Varni et al., 2003).

Table 7

Sizing Me Up Reliability and MCID Statistics

Scale	Reliability	MCID
SMU Physical	$\rho = .58$	8.29
SMU Emotion	$\rho = .69$	11.86
SMU Social Avoidance	$\rho = .65$	7.02
SMU Positive Social Attributes	$\rho = .39$	14.47
SMU Teasing/ Marginalization	$\rho = .33$	10.74
SMU Total	$\rho = .93$	3.09

Table 8

PedsQL Reliability and MCID Statistics

Scale	Reliability	MCID
<i>PedsQL</i> Physical	$\rho = .65$	7.30
<i>PedsQL</i> Emotional	$\rho = .64$	12.08
<i>PedsQL</i> Social	$\rho = .62$	10.13
<i>PedsQL</i> School	$\rho = .54$	12.72
<i>PedsQL</i> Total	$\rho = .86$	4.38

Discussion

Construct Validity and Invariance

The current study was an evaluation of a weight-related QOL measure in a nontreatment-seeking sample of 4th and 5th grade children with overweight and obesity. This study fills a gap in the QOL assessment literature by evaluating the construct validity of *Sizing Me Up* in this population and by testing the implicit theoretical assumption that overweight and obesity confers a unique experience to children's QOL not captured by broad and general assessment tools. Findings from the CFA confirmed the *Sizing Me Up* five-factor first-order structure with one second-order factor previously proposed by Zeller and Modi (2009) among treatment-seeking children. Moreover, construct validity was partially established for the *Sizing Me Up* scales and total score. Criterion validity was established for the physical, social avoidance, teasing and marginalization scales, and the total score.

This lends evidence that these scales assess latent variables that are significantly associated with weight and should fluctuate with changes in weight.

Generally, results support the use of the *Sizing Me Up* in community samples of children with overweight and obesity. This was an important finding because it speaks to the generalizability of the measure in the context of research studies. Greater generalizability in measurement allows for easier transportation of research findings in treatment-seeking populations to community-based interventions. That is, when mechanisms of change in weight-related QOL are identified in a treatment-seeking sample, it is possible to confidently test similar mechanisms in a community sample with well-established measures. Clinically, this means that measures of weight-related QOL may be appropriate for use in nontreatment-seeking populations such as school- or community-based intervention programs or one-on-one with children targeted for motivational changes as a starting point for clinical intervention.

Adding to the evidence for construct validity of the *Sizing Me Up*, convergent validity was established with the *PedsQL* for the *Sizing Me Up* physical scale, teasing and marginalization scale, and total score. Departing from the associations discovered by Zeller and Modi (2009), the current investigation did not observe significant associations between the *Sizing Me Up* and *PedsQL* emotion scales or the social avoidance and social scales. Several explanations are available for this finding. First, Zeller and Modi's (2009) initial validation sample was comprised of only treatment-seeking children with obesity while the current sample included children

with overweight and obesity who were not seeking treatment. It is possible that a difference in sample characteristics can account for the failure to replicate the correlations in the latent factor analysis. Specifically, the amount of impairment in the current sample may not be sufficient to elicit the associations found in a treatment-seeking sample. Second, the magnitude of the correlations observed in the initial validation study was small ($r = .35-.36$) and it is possible that shared measurement error accounted for some of this association. By removing measurement error, the current analysis produced a more accurate true score estimate (Brown, 2006) and it may be that these scales do not share an association in latent space. This is important at a practical level because it may indicate that clinicians should use different scoring conventions than researchers (see discussion of this issue in clinical implications section). Additional studies are needed to definitively determine whether these correlations are limited to obese samples or if they reliably disappear when measurement error is removed from the analysis.

The results of the invariance test comparing children with overweight and obesity to children with a healthy weight revealed that the assumption of factorial invariance does not hold across these two distinct groups. The results of the current study indicate that *Sizing Me Up* should not be administered to children with a healthy weight due to a different underlying measurement structure. Therefore, weight-related QOL measures are not appropriate for population-level assessments and comparisons across disease conditions. For these purposes, broad and general QOL measures are still the most appropriate choice (Palermo et al., 2008).

Recently, it has been suggested that not every chronic illness should be assessed using a disease-specific instrument; and that the decision to do so should be made on the particular characteristics of a given illness population (Connelly, Fulmer, Smith, Anson, & Poull, 2011). This is a reasonable assertion given that assessing disease-specific QOL assumes that there is something unique about the disease experience that affects QOL in a way that those without the condition will not experience; thus, leading to greater measurement sensitivity of disease-specific instruments than broad and general measures (Palermo et al., 2008). This is an assumption that underlies all disease-specific QOL instruments, but commonly goes untested in the empirical literature. Results from the current study provide support for the theoretical assumption that weight-related QOL measures add additional information to a QOL assessment battery over and above what is available from broad and general measures (Palermo et al., 2008); and that the *Sizing Me Up* assesses the weight-related QOL construct in nontreatment-seeking children with overweight and obesity. Each first-order scale and the one second-order *Sizing Me Up* scale appear provide additional information about a nontreatment-seeking child's QOL experience beyond what is available from an assessment using only the *PedsQL*.

Statistically demonstrating that *Sizing Me Up* adds incrementally to the understanding of a theoretical construct is an important strength of this study, and has several implications. First, the current study provides support for the widely espoused belief that disease-specific instruments add additional information to a QOL assessment over and above broad and general measures. Given the relatively low cost

of administration, children with overweight and obesity submitting to a QOL assessment should be administered the *Sizing Me Up* measure as well as a broad and general tool. Second, when considering markers of clinical progress, children should be compared to themselves on the *Sizing Me Up* instrument because of the greater sensitivity to change. Third, information gained from *Sizing Me Up* should be considered as different from information gained from the *PedsQL*. That is, a low QOL scores on each instrument may have different causes and should be explored carefully before proceeding with an intervention.

In addition to the evidence for criterion validity noted above, the scaled means and standard deviations provide information that speaks to *Sizing Me Up*'s association with weight status. While not part of a primary aim of the current study, visual inspection of the means in Table 6 may indicate that *Sizing Me Up* differentiates between overweight, obese, and very obese (i.e., BMI_{99th} \geq 99) groups on each of the component scales. Table 6 also indicates that the primary cause of low weight-related QOL as measured by *Sizing Me Up* is low scores on positive social attributes. This presents different targets for intervention than other psychosocial measures, which may lend themselves to cognitive change. For example, if a child “felt worried because of your size” or they are not “happy because of your size” the clinician may not be interested in modifying beliefs about the child’s size. That is, it is not useful to help the very obese child reframe their thinking to become happy about their size. Instead, problems identified by these items may be more useful as

vehicles for discussing motivation to change a set of health behaviors such as diet and exercise habits.

Clinical Implications

It is important to note, the current results indicate that the *Sizing Me Up* should not be used in exactly the same way in all evaluations of weight-related QOL. In school-based or community studies where large samples and sophisticated statistical techniques are available, investigators may choose to use either the five-factor first-order structure or the second-order one-factor structure within a SEM framework. However, the low latent reliability estimates for the first-order structure limit the utility of these scaled scores in clinical practice or as summary scores in traditional regression models. That is, the amount of error variance captured by these scales is unacceptably high and would result in summary scores that have little or no association with other variables in an inferential statistical analysis, and perhaps even less clinical utility. In either of these two cases, however, the total *Sizing Me Up* score can be computed and used to produce meaningful statistical inferences or provide useful clinical information.

For example, a school nurse participating in a tiered healthy lifestyle intervention may identify a child as having overweight or obesity at an annual health check. The nurse may refer the identified child to a psychologist or medical professional working in the school to determine if the child should receive the universal, selected, or targeted arm of the program. The relatively smaller MCID

means that the *Sizing Me Up* total score should provide a more sensitive assessment of QOL than the *PedsQL* in children with overweight and obesity, and should be included in the QOL assessment. The descriptive statistics reported in the current study combined with Zeller and Modi's (2009) initial evaluation suggests that 68 may be a reasonable total score that could prompt the psychologist or pediatrician to have a conversation with the child and their family about the way weight is impairing the child's QOL. In an initial discussion with the child and his/her family the practitioner may encounter resistance to their suggestions for lifestyle change. In this scenario, the health care provider may be able to use the *Sizing Me Up* items to point out areas where the child experiences impairment. The professional may then be able to use these items to call attention to the impairment and effectively increase motivation for behavior change.

As the psychologist continues to track the child during the school year, the baseline total score yielded by the *Sizing Me Up* instrument can be assessed for changes of approximately three or more points (i.e., MCID) to determine the impact of the lifestyle changes for that particular child. The evidence from the current investigation suggests that such an assessment will yield substantively different and more sensitive information about a child with overweight and obesity than a broad and general QOL assessment.

Limitations

The findings of the current study are limited by several factors. First, the *Sizing Me Up* is designed for use among 5-13 year-old children (Zeller & Modi, 2009). However, the current study sample was limited to children aged 8-12. Therefore, the findings may not generalize to all participants who could respond to the measure. While one of the strengths of the current study was the use of a nontreatment-seeking sample, the current results cannot address treatment seeking samples. It is possible that subtle, as of yet untested differences exist between how the instrument behaves in community-based and treatment-seeking samples. Therefore, additional confirmatory work with the *Sizing Me Up* is necessary in treatment seeking samples including CFA, tests of equivalence of *Sizing Me Up* scales and *PedsQL* scales, and factorial invariance tests.

Future Directions

As noted above, additional confirmatory work is needed to enhance the understanding of the factor structure of *Sizing Me Up*. First, future validation studies of *Sizing Me Up* should attempt to determine whether a wider range of children can participate in group administrations of the instrument. This would provide information regarding whether or not the instrument can be used in studies of entire grade schools, potentially informing community or school-based interventions. Second, a CFA of *Sizing Me Up* is still needed in a treatment-seeking sample that closely resembles Zeller and Modi's (2009) original sample. As noted above, exploratory factor analysis can be a useful tool early in questionnaire development.

However, CFA provides a much more theoretically sound and stringent test of a measure. Specifically, a CFA in a treatment seeking sample would provide information about scoring conventions for clinicians working with treatment-seeking children. Additionally, there is still cause for concern that the positive social attributes scale does not constitute a QOL factor, and may measure a unique construct. A CFA in a treatment-seeking sample that tested a two-factor versus a one-factor second-order model would add confidence to the assertion that the measure assesses a unitary weight-related QOL construct. Similarly, longitudinal invariance studies are needed to demonstrate that the instrument is stable across measurement occasions in the same sample. This information will assure researchers interested in using the measure in the context of longitudinal work or an intervention study that the instrument will not be subject to measurement fluctuations as a function of repeated administrations or the passage of time. Finally, the current study collapsed three potentially meaningful groups for the final analysis (i.e., overweight, obese, and very obese). A larger sample is necessary to allow for a test of factorial invariance across these weight categories. Such a test would provide confidence that *Sizing Me Up* has a similar factor structure that underlies each weight category and can be used to compare members of all three groups.

The results of the current investigation indicate that *Sizing Me Up* should be sensitive to changes in BMI%ile. However, it is unknown how physical fitness and dietary changes (i.e., the behavioral variables that underlie Δ BMI%ile) might affect scores on *Sizing Me Up*. Future studies, should attempt to examine these associations

given that even successful behavioral interventions produce slow changes in BMI%ile. It is possible that changes in QOL occur as a result of healthy lifestyle rather than or in addition to changes in BMI%ile. This would be a positive finding because it would mean that changes in QOL would be available to children more rapidly than changes in weight, and would have implications for treatment planning.

Conclusion

In conclusion, the *Sizing Me Up* is a reliable and valid instrument for use in nontreatment-seeking samples of children with overweight and obesity. However, it is not appropriate for children with a healthy weight. Consistent with the larger theoretical literature on QOL assessment (Palermo et al., 2008), the *Sizing Me Up* total score appears to be a more precise measure of QOL in children with overweight and obesity. The current study offers evidence of the importance of testing theoretical assumptions about newly developed assessment instruments. In particular, a large number of disease-specific QOL tools are currently available; although, tests of their incremental merit relative to broad and general measures are lacking from the empirical literature. *Sizing Me Up* appears to offer researchers and clinicians more specific information about a child's experience of overweight or obesity and a total score that should demonstrate meaningful change at smaller intervals than the *PedsQL*. It is recommended that clinicians interested in QOL among children with overweight and obesity use the *Sizing Me Up* total score as part of their assessment battery. Researchers interested in assessing QOL in school-aged children should continue to follow the evidence-based assessment recommendations and include both

a broad and general QOL instrument as well as the *Sizing Me Up* scale (Palermo et al., 2008).

References

- Barlow, S. E., & the Expert Committee. (2007). Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: Summary report. *Pediatrics*, *120*, S164-S192.
doi:10.1542/peds.2007-2329C
- Braet, C., Tanghe, A., Decaluwe, V., Moens, E., & Rosseel, Y. (2004). Inpatient treatment for children with obesity: Weight loss, psychological well-being, and eating behavior. *Journal of Pediatric Psychology*, *29*, 519-529.
doi:10.1093/jpepsy/jsh054
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255.
doi: 10.1207/S15328007SEM0902_5
- Connelly, M., Fulmer, D., Smith, A., & Anson, L. (2011). Predictors of pre-operative and post-operative quality of life in children and adolescents undergoing spinal fusion surgery for idiopathic scoliosis. Paper presented at the National Conference in Pediatric Psychology, San Antonio, TX.
- Connolly, M. A., & Johnson, J. A. (1999). Measuring quality of life in paediatric patients. *Pharmacoeconomics*, *16*, 605-625.
<http://adisonline.com/pharmacoeconomics/Pages/default.aspx>
- Dietz, W. H., & Bellizzi, M. C. (1999). Introduction: The use of body mass index to

- assess obesity in children. *American Journal of Clinical Nutrition*, 70, 123S-125S. <http://www.ajcn.org/>
- 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. doi:10.1006/jado.2002.0470
- Eiser, C., & Morse, R. (2001). Can parents rate their child's health-related quality of life? Results of a systematic review. *Quality of Life Research*, 10, 347-357. doi: 10.1023/A:101225372372
- Freedman, D. S. (2002). Clustering of coronary heart disease risk factors among obese children. *Journal of Pediatric Endocrinology and Metabolism*, 15, 1099-1108. <http://www.degruyter.de/journals/jpem/detailEn.cfm>
- Friedlander, S. L., Larkin, E. K., Rosen, C. L., Palermo, T. M., & Redline, S. (2003). Decreased quality of life associated with obesity in school-aged children. *Archives of Pediatrics and Adolescent Medicine*, 157, 1206-1211. doi:10.1001/archpedi.157.12.1206
- Gilliland, F. D., Berhane, K., Islam, T., McConnell, R., Gauderman, W. J., Gilliland, S. S., ... Peters, J. M. (2003). Obesity and the risk of newly diagnosed asthma in school-age children. *American Journal of Epidemiology*, 158, 406-415. doi:10.1093/aje/kwg175
- Goodwin, D. A. J., Boggs, S. R., & Graham-Pole, J. (1994). Development and validation of the Pediatric Oncology Quality of Life Scale. *Psychological Assessment*, 6, 321–328. doi:10.1037//1040-3590.6.4.321

- Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research, 31*, 197-218. doi:10.1207/s15327906mbr3102_3
- Guyatt, G. H., Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality of life. *Annals of Internal Medicine, 118*, 622-629. doi:10.1016/0735-1097(93)90488-M
- Ingersoll, G. M., & Marrero, D. G. (1991). A modified quality of life measure for youths: Psychometric properties. *Diabetes Care, 9*, 114-118. doi:10.1177/014572179101700219
- Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials, 10*, 407-415. doi:10.1016/0197-2456(89)90005-6
- Jöreskog, K.G. & Sörbom, D. (2006). LISREL 8.8 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.
- Juniper, E. F., Guyatt, G. H., Feeny, D. H., Ferrie, P. J., Griffith, L. E., & Townsend, M. (1996). Measuring quality of life in children with asthma. *Quality of Life Research, 5*, 35-46. doi:10.1007/BF00435967
- Kitzmann, K. M., Dalton, W. T., Stanley, C. M., Beech, B. M., Reeves, T. P., Buscemi, J., ... Midgett, D. L. (2010). Lifestyle interventions for youth who are overweight: A meta-analytic review. *Health Psychology, 29*, 91-101. doi:10.1037/a0017437

- Kolotkin R. L., Zeller, M., Modi, A. C., Samsa, G. P., Quinlan, N. P., Yanovski, J. A.,... Roehrig, H. R.. (2006). Assessing weight-related quality of life in adolescents. *Obesity, 14*, 448-457. doi:10.1038/oby.2006.59
- Landgraf, J. M., Abets, L., & Ware, J. E. (1996). *The CHQ User's Manual* (1st ed.). Boston, MA: The Health Institute, New England Medical Center.
- Lawson, M. L., Kirk, S., Mitchell, T., Chen, M. K., Loux, T. J., Daniels, S. R., ... Inge, T. H. (2006). One-year outcomes of Roux-en-Y gastric bypass for morbidly obese adolescents: A multicenter study from the Pediatric Bariatric Study Group. *Journal of Pediatric Surgery, 41*, 137-143. doi:10.1016/j.jpedsurg.2005.10.017
- Limbers, C. A., Newman, D. A., & Varni, J. W. (2008a). Factorial invariance of child self-report across age subgroups: A confirmatory factor analysis of ages 5 to 16 years utilizing the PedsQL 4.0 Generic Core Scales. *Value in Health, 11*, 659-668. doi:10.1111/j.1524-4733.2007.00289.x
- Limbers, C. A., Newman, D. A., & Varni, J. W. (2008b). Factorial invariance of child self-report across healthy and chronic health condition groups: A confirmatory factor analysis utilizing the PedsQL™ 4.0 Generic Core Scales. *Journal of Pediatric Psychology, 33*, 630-639. doi:10.1093/jpepsy/jsm131
- Little, T. D. (1997). Mean and covariance structures (MACS) of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53-76. doi:10.1207/s15327906mbr3201_3

- MacCallum, R. C., Brown, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149. doi:10.1037/1082-989X.1.2.130
- Mallory, G. B., Fiser, D. H., & Jackson, R. (1989). Sleep-associated breathing disorders in morbidly obese children and adolescents. *Journal of Pediatrics, 115*, 892-897. doi:10.1016/S0022-3476(89)80738-3
- Matza, L. S., Swensen, A. R., Flood, E. M., Sexnik, K., & Leidy, N. K. (2004). Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value in Health, 7*, 79-92. doi:10.1111/j.1524-4733.2004.71273.x
- McGovern, L., Johnson, J. N., Paulo, R., Hettinger, A., Singhal, V., Kamath, C., ... Montori, V. M. (2008). Clinical review: Treatment of pediatric obesity: a systematic review and meta-analysis of randomized trials. *Journal of Clinical Endocrinology and Metabolism, 93*, 4600-4605. <http://jcem.endojournals.org/>
- Must, A., & Strauss, R. S. (1999). Risks and consequences of childhood and adolescent obesity. *International Journal of Obesity Related Metabolic Disorders, 23*, S2-S11. doi:10.1038/sj/ijjo/0800852
- Muthén, L. K., & Muthén B. O. (2002). Teacher's corner: How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599-620. <http://www.tandf.co.uk/journals/titles/10705511.asp>
- Ogden, C. L., Carroll, M. D., Curtin, Lamb, M. M., & Flegal, K. M. (2010).

- Prevalence of high body mass index in US children and adolescents, 2007-2008. *Journal of the American Medical Association*, 303, 242-249.
doi:10.1001/jama.2009.2012
- Palermo, T. M., Long, A. C., Lewandowski, A. S., Drotar, D., Quittner, A. L., & Walker, L. S. (2008). Evidence-based assessment of health-related quality of life and functional impairment in pediatric psychology. *Journal of Pediatric Psychology*, 33, 983-996. doi:10.1093/jpepsy/jsn038
- Palermo, T. M., Witherspoon, D., Valenzuela, D., & Drotar, D. (2004). Development and validation of the Child Activity Limitations Interview: A measure of pain-related functional impairment in school-age children and adolescents. *Pain*, 109, 461-470. doi:10.1016/S0304-3959(04)00103-4
- Ping, R. A. (2004). On assuring valid measures for theoretical models using survey data. *Journal of Business Research*, 57, 125-141. doi: 10.1016/S0148-2963(01)00297-1
- Preacher, K. J., & Coffman, D. L. (2006, May). Computing power and minimum sample size for RMSEA [Computer software]. Available from <http://quantpsy.org/>.
- Quinlan, N. P., Kolotkin, R. L., Fuemmeler, B. F., & Costanzo, P. R. (2009). Psychosocial outcomes in a weight loss camp for overweight youth. *International Journal of Pediatric Obesity*, 4, 134-142.
doi:10.1080/17477160802613372
- Quittner, A. L., Buu, A., Messer, M. A., Modi, A. C., & Watrous, M. (2005).

- Development and validation of the Cystic Fibrosis Questionnaire in the United States: A health-related quality-of-life measure for cystic fibrosis. *Chest*, *128*, 2347–2354. doi:10.1378/chest.128.4.2347
- Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy*, *35*, 299-331. doi: 10.1016/S0005-7894(04)80041-8
- Remsberg, K. E., Demerath, E. W., Schubert, C. M., Chumela, W. C., Sun, S. S., & Siervogel, R. M. (2005). Early menarche and the development of cardiovascular disease risk factors in adolescent girls: The Fels longitudinal study. *Journal of Clinical Endocrinology and Metabolism*, *90*, 2718-2724. doi:10.1210/jc.2004-1991
- Retsch-Bogart, G. Z., Quittner, A. L., Gibson, R. L., Oermann, C. M., McCoy, K. S., Montgomery, A. B., & Cooper, P. J. (2009). Efficacy and safety of inhaled aztreonam lysine for airway pseudomonas in cystic fibrosis. *Chest*, *135*, 1223-1232. doi: doi:10.1378/chest.08-1421
- Revicki, D. A., Osoba, D., Fairclough, D., Barofsky, I., Berzon, R., Leidy, N. K., & Rothman, M. (2000). Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. *Quality of Life Research*, *9*, 887-900. doi: 10.1023/A:1008996223999
- Rhodes, S. K., Shimoda, K. C., Waid, L. R., Mahlen, P., Oexmann, M. J., Collop, N.

- A., & Willi, S. M. (1993). Neurocognitive deficits in morbidly obese children with obstructive sleep apnea. *Journal of Pediatrics*, *127*, 741-744.
doi:10.1016/S0022-3476(95)70164-8
- Rubin, D.B. (1976) Inference and missing data. *Biometrika*, *63*, 581-592.
doi:10.2307/2335739
- Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, *8*, 3-15. doi:10.1191/096228099671525676
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147-177. doi: 10.1037/1082-989X.7.2.147
- Schlomer, G. L., Bauman, S., & Card., N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, *57*, 1-10. doi:10.1037/a0018082
- Schwimmer, J. B., Burwinkle, T. M., & Varni, J. W. (2003). Health-related quality of life of severely obese children and adolescents. *Journal of the American Medical Association*, *289*, 1813-1819. doi:10.1001/jama.289.14.1813
- Singh, G., Arthreya, B., Fries, J. F., & Goldsmith, D. P. (1994). Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis & Rheumatism*, *37*, 1761-1769. doi:10.1002/art.1780371209
- 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. doi: 10.1093/jpepsy/21.2.175
- Spilker, B., & Revicki, D. A. (1999). Taxonomy of quality of life. In B. Spilker (Ed.),

Quality of life and Pharmacoeconomics in Clinical Trials (pp. 25-32).

Philadelphia: Lippincott-Raven.

Starfield, B., Riley, A. W., & Green, B. F. (1999). *Manual for the child health and illness profile: Adolescent edition (CHIP-AE)*. Baltimore: The Johns Hopkins University.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics, Fourth Edition*. Needham Heights, MA: Allyn & Bacon.

Turner, R., Quittner, A. L., Parasuraman, B. M., & Cleeland, C. S. (2007). Patient-reported outcomes: Instrument selection issues. *Value in Health, 2*(Suppl2), S86-S93. doi:10.1111/j.1524-4733.2007.00271.x

Varni, J. W., Burwinkle, T. M., & Lane, M. M. (2005). Health-related quality of life measurement in pediatric clinical practice: An appraisal and precept for future research and application. *Health Quality of Life Outcomes, 3*, 1-9. doi: 10.1186/1477-7525-3-34

Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL™ 4.0 as a pediatric population health measure: Feasibility, Reliability, and Validity. *Ambulatory Pediatrics, 3*, 329-341. doi:10.1367/1539-4409(2003)003<0329:TPAAPP>2.0.CO;2

Varni J. W., Limbers, C. A., & Burwinkle, T. A. (2007). How young can children reliably and validly self-report their health-related quality of life?: An analysis of 8,591 children across age subgroups with the PedsQL™ 4.0 Generic Core

Scales. *Health and Quality of Life Outcomes*, 5, 1-13. doi:10.1186/1477-7525-5-1

Varni, J. W., Limbers, C. A., Newman, D. A., & Seid, M. (2008). Longitudinal factorial invariance of the PedsQL™ 4.0 Generic Core Scales child self-report version: One year prospective evidence from the California State Children's Health Insurance Program (SCHIP). *Quality of Life Research*, 17, 1153-1162. doi: 10.1007/s11136-008-9389-3

Varni J. W., Seid M., & Kurtin P. S. (2001). PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 Generic Core Scales in healthy and patient populations. *Medical Care*, 39, 800–812.
<http://journals.lww.com/lww-medicalcare/pages/default.aspx>

Williams, J., Wake, M., Hesketh, K., Maher, E., & Waters, E. (2005). Health-related quality of life of overweight and obese children. *Journal of the American Medical Association*, 293, 70-76. doi: 10.1016/j.accreview.2005.04.008

World Health Organization. (1947). The constitution of the World Health Organization. *WHO Chronicles*. I. 29.

Wyrwich, K., Tierney, W., & Wolinsky, F. (1999). Further evidence supporting an SEM-based criterion for identifying meaningful intraindividual changes in health-related quality of life. *Journal of Clinical Epidemiology*, 52, 861-873. doi: 10.1016/S0895-4356(99)00071-2

Young, T. K., Dean, H. J., Flett, B., & Wood-Steiman, P. (2000). Childhood obesity

in a population at high risk for type 2 diabetes. *Journal of Pediatrics*, *136*, 365-369. doi: 10.1067/mpd.2000.103504

Zeller, M. H., & Modi, A. C. (2006). Predictors of health-related quality of life in obese youth. *Obesity*, *14*, 122-130. doi: 10.1038/oby.2006.15

Zeller, M. H., & Modi, A. C. (2008). Psychosocial factors related to obesity in children and adolescents. In E. Jelalian & R G. Steele (Eds.), *Handbook of childhood and adolescent obesity* (pp. 25-42). New York: Springer. doi: 10.1007/978-0-387-76924-0_3

Zeller, M. H., & Modi, A. C. (2009). Development and initial validation of an obesity-specific quality-of-life measure for children: Sizing Me Up. *Obesity*, *17*, 1171-1177. doi:10.1038/oby.2009.47

.

Appendix A

ID #: _____

SMU

Please check the box that corresponds with your answers.
During the past month, tell us how much you:

	None of the time	A little	A lot	All the time
1. Were teased by other kids because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Felt sad because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Were told you are healthy or growing well	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Felt mad because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Felt left out because of your size (no one talks or sits with you).....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Found it hard to swing, climb, skip, bounce a ball, or jump rope because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Like yourself because of your size	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Stood up for or helped other kids because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Felt frustrated because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Felt worried because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Chose not to go to school because of your size	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Had problems fitting into your desk at school because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Felt happy because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Were picked first for recess or gym because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Were teased by other kids when physically active (e.g. move your body) because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Felt you have a good sense of humor.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Did not want to go to the swimming pool or park because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Felt uncomfortable sleeping at a friend's house because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Got upset at mealtimes (e.g. cried, fussed, argued).....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Found it hard to keep up with other kids because of your size	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. Got out of breath and had to slow down because of your size.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. Chose not to participate in gym or recess at school because of your size... ..	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>