AT WHAT AGE CAN CHILDREN RELIABLY AND VALIDLY SELF-REPORT THEIR HEALTH-RELATED QUALITY OF LIFE?: AN INVESTIGATION USING THE PEDSQL™ 4.0 GENERIC CORE SCALES DATABASE

A Thesis

by

CHRISTINE A. LIMBERS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2006

Major Subject: Psychology

AT WHAT AGE CAN CHILDREN RELIABLY AND VALIDLY SELF-REPORT THEIR HEALTH-RELATED QUALITY OF LIFE?: AN INVESTIGATION USING THE PEDSQL™ 4.0 GENERIC CORE SCALES DATABASE

A Thesis

by

CHRISTINE A. LIMBERS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Co-Chairs of Committee,	James W. Varni
	Robert W. Heffer
Committee Member,	Daniel Newman
Head of Department,	Leslie C. Morey

December 2006

Major Subject: Psychology

ABSTRACT

At What Age Can Children Reliably and Validly

Self-Report Their Health-Related Quality of Life?:

An Investigation Using the PedsQL™ 4.0 Generic

Core Scales Database. (December 2006)

Christine A. Limbers, B.A., Santa Clara University

Co-Chairs of Advisory Committee: Dr. James W. Varni
Dr. Robert W. Heffer

Health-related quality of life (HRQOL) assessment has emerged as a vital health outcome measure in clinical trials, healthcare services and evaluation, and population health outcomes research. Reliability, validity, and parent-child agreement of the PedsQL™ 4.0 Generic Core Scales were examined using child self-report and parent proxy-report age subgroup data on over 8,000 children ages 5-16 years from the PedsQL 4.0 Generic Core Scales Database^SM^. The PedsQL™ 4.0 Generic Core Scales demonstrated good internal consistency reliability for children as young as 5 years; healthy children across the age subgroups demonstrated a statistically significant difference in HRQOL (better HRQOL) than children with a known chronic health condition. Confirmatory factor analysis demonstrated that a 5-factor model fit almost identically across the age subgroups, providing further evidence that children as young as 5 years are reliable and valid self-reporters of their HRQOL. Parent-child agreement was in the moderate-to-good range, with parents reporting significantly higher

PedsQL™ 4.0 scores across the age subgroups. In conclusion, the analyses support the reliability and validity of child self-report in children as young as 5 years old.

TABLE OF CONTENTS

INTRODUCTION: HEALTH-RELATED QUALITY OF LIFE ASSESSMENT

Health-related quality of life (HRQOL) assessment has emerged as a vital health outcome measure in clinical trials, healthcare services research and evaluation, and population health outcomes research (Fekkes, Theunissen, Brugman, & Veen, 2000; Varni, Burwinkle, & Lane, 2005; Varni, Burwinkle, & Seid, 2006). It has been increasingly recognized that traditional outcome measures such as mortality and morbidity may be too limited in focus to provide an accurate measurement of health in the general population (Forrest, SA, Dougherty, & MR, 2003; Newacheck & Taylor, 1992; Varni, Burwinkle, Seid, & Skarr, 2003). Consequently, the last decade has witnessed a dramatic increase in the development and utilization of multidimensional HRQOL measures that are generic and disease-specific (Eiser & Morse, 2001a; Varni, Burwinkle et al., 2003). Generic HRQOL measures encompass broader health domains such as physical, emotional, social, and role functioning (World Health Organization, 1948;Varni, Burwinkle et al., 2003) and enable comparisons across patient populations (Varni, Burwinkle, & Seid, 2005). Disease-specific measures include condition-specific symptoms and may improve measurement sensitivity for health domains germane to a disease or condition (Varni, Burwinkle, & Seid, 2005). There is a growing recognition that for chronic health conditions, administering both generic and disease-specific HRQOL measures may result in a more thorough assessment of the patient's HRQOL

____

This thesis follows the style of the *Journal of Clinical Child Psychology*.

(Patrick & Deyo, 1989; Sprangers, Cull, Bjordal, Groenvold, & Aaronson, 1993; Varni et al., 1999).

There are distinct advantages to measuring HRQOL in pediatric populations. The utilization of HRQOL measures can help determine the broad physical and psychosocial impact of a specific disease or disability, assist in evaluating changes in treatment delivery, and aid in the identification of interventions which impose fewer limitations on the quality of life of children (Eiser, 2004; Eiser, Mohay, & Morse, 2000; Varni, Burwinkle, Seid et al., 2003). Furthermore, HRQOL measures can help evaluate the health care needs of children within the community and schools, assist in the identification of subgroups of children who are at risk for health problems, and help shape public policy decisions related to community health and the allocation of health care resources for children (Varni, Burwinkle, & Seid, 2006; Varni, Burwinkle, Seid et al., 2003).

Central to the measurement of HRQOL is the notion that a person's HRQOL is highly individualized and dependent on his or her subjective experience of health (De Civita et al., 2005; Eiser & Morse, 2001a). Assessing HRQOL from the individual's perspective is essential given that for certain HRQOL concepts (i.e. pain intensity, cognitive fatigue), there are often no externally observable or physical measures (FDA, 2006). Thus, for these domains of functioning, individual self-report may be the only source of data (FDA, 2006). Furthermore, reduction of symptoms on clinical measures may not correspond with improvements in how a patient feels or is able to perform his or her daily activities (FDA, 2006). For example, a patient may have clinically significant improvements in lung functioning as assessed by a physiological measure such as spirometry (FDA, 2006). However, these improvements may not correspond with improvements in asthma-related symptoms and the impact these symptoms have on the patient's ability to perform activities of daily living (FDA, 2006). HRQOL assessment has the potential to enhance what is known about an individual's health from physiological measures and the physician's perspective (FDA, 2006).

The Proxy Problem

While self-report measures are considered the standard method for assessing HRQOL in adult populations (De Civita et al., 2005; Eiser & Morse, 2001a), HRQOL assessment for children has historically not taken into account young children's perceptions of their health and well-being (Eiser & Morse, 2001a). Few self-report HRQOL measures have been developed for use in pediatric populations under the age of 8 years (De Civita et al., 2005; Eiser et al., 2000; Varni, Seid, & Kurtin, 2001a). As a result, measures of HRQOL for young children are often completed by parent proxy-respondents on behalf of the child (De Civita et al., 2005; Eiser et al., 2000; Eiser & Morse, 2001c; Varni et al., 2001a). The importance of assessing children's self-reported HRQOL has been increasingly recognized given a growing body of literature demonstrating the imperfect concordance between child self-report and parent proxy-report (Clancy, McGrath, & Oddson, 2005; Koot & Wallander, 2001; Levi & Drotar, 1999; Varni, Katz, Colegrove, & Dolgin, 1995; Varni, Katz, Seid, Quiggins, Friedman-Bender et al., 1998; Varni, Seid, & Rode, 1999; Yasuda et al., 2004). Parent-child concordance has been shown to be influenced by a number of variables, including the observability of the HRQOL domain (Baxt, Kassam-Adams, Nance, Vivarelli-O'Neill, & Winston, 2004; Eiser & Morse, 2001a; Levi & Drotar, 1999; Varni et al., 1999). In a meta-analysis of studies evaluating parent-child agreement on different measures of HRQOL, Eiser and Morse (2001) found generally good agreement ($r > 0.50$) between child and parent report for domains reflecting physical activity, functioning and some symptoms, but weaker agreement ($r < 0.30$) between child and parent report for emotional and social HRQOL domains (Eiser & Morse, 2001a). Furthermore, parent-child concordance may be affected by the child's age and illness status (Cremeens, Eiser, & Blades, 2006; Varni, Katz, Seid, Quiggins, & Friedman-Bender, 1998). In a sample of children with asthma, Annett et al. (2003) found parent-child agreement to be greater as children increased in age (Annett, Bender, DuHamel, & Lapidus, 2003). Conversely, in a sample of healthy children ages 5.5 to 8.5 years, Creemeens et al. (2006) found no median differences between child self-report and parent proxy-report scores for children in the youngest age group (5.5 to 6.5 years) and larger parent-child median differences for children in the older age groups (Cremeens et al., 2006).

Further complicating the use of proxy reporters, who typically are the child's parents, are concerns regarding the impact of parental variables on parents' perceptions of their children's health (De Los Reyes & Kazdin, 2004; Richters, 1992). Variables such as distress, gender, and health have been shown to influence parents' reports of their children's health (Duhig, Renk, Epstein, & Phares, 2000; Landgraf & Abetz, 1997; Luoma, Koivisto, & Tamminen, 2004; N. Theunissen et al., 1998; Truetler & Epkins, 2003; Waters, Doyle, Wolfe, Wright, & Wake, 2000). Waters et al. (2000) found that mothers with poor global health reported a higher incidence of child illness than mothers with good global health, an effect which was not found in fathers (Waters et al., 2000). Taken together, the evidence is quite compelling that parent's reports of their children's HRQOL are influenced by numerous variables and are not synonymous with children's own ratings of their HRQOL. Thus, we contend that parent proxy-report should not be used as a substitute for child self-report when children are capable of self-reporting their own HRQOL (Eiser et al., 2000; N. C. M. Theunissen et al., 1998). At the very least, parent proxy-reports should be used to complement children' self-reported information. The demonstration of imperfect concordance between parent proxy-report and child self-report and the recognition that HRQOL derives from an individual's perceptions indicate a critical need in pediatric HRQOL assessment for reliable and valid child self-report measures for children of the youngest possible age (Varni, Burwinkle, & Seid, 2005).

Cognitive Development

The construction of reliable and valid self-report measures for young children is complicated by young children's developmental immaturity (Eiser et al., 2000). The development of self-report measures for young children must therefore be based on a thorough understanding of the child's cognitive developmental level (Eiser et al., 2000). According to Piaget's theoretical framework, children between the ages of 2 and 7 years are in the preoperational stage of cognitive development (Piaget, 1930). Children in the preoperational stage rely predominantly on concrete experiences for acquiring and reflecting information (Perrin & Gerrity, 1981; Santrock, 2000). Research examining young children's self descriptions indicate that preoperational children make differentiations between themselves and others primarily on the basis of observable behaviors and characteristics rather than internal experiences (De Civita et al., 2005; La Greca, 1990). As a result, young children are more likely to describe themselves in terms of physical attributes, possessions, and overt behaviors (Keller, Ford, & Meacham, 1978; La Greca, 1990). While children acquire a greater understanding of mental operations when they reach the concrete operational stage at approximately age 7, they continue to have difficulty understanding hypothetical and abstract concepts until they reach the formal operations stage at approximately age 11 (Perrin & Gerrity, 1981; Santrock, 2000). Taken as a whole, the research indicates a need to link questions to children's experiences and to use concrete language and examples on self-report HRQOL measures for young children (Eiser et al., 2000).

Children's Understanding of Emotions

Also critical to young children's ability to self-report their HRQOL is their capacity to recognize and interpret emotional states. During childhood, the comprehension of emotions improves as children learn to make more complicated causal inferences, accept the possibility of feeling two conflicting emotions simultaneously, and recognize mental aspects of emotions (La Greca, 1990). It has been well documented in the developmental literature that by the time children reach three years old, they recognize themselves and others as emotional beings (La Greca, 1990; Santrock, 2000). Furthermore, by this age children have the capacity to communicate about basic emotions in everyday scenarios, label simple feeling states of others and themselves, make basic causal interpretations of feelings, and participate in pretend play that includes emotional expression (Bretherton, Fritz, Zahn, & Ridgeway, 1986; La Greca, 1990). Research examining the effect of color on children's emotions has demonstrated that young children not only show specific color preferences, but are also able to attribute emotional characteristics to colors in a reliable manner (Terwogt & Hoeksma, 1995; Zentner, 2001). Zentner (2001) found that when asked to match colors with emotional expressions, 3 and 4 year old children consistently attributed bright colors with happy emotional expressions and dark colors with sad emotional expressions, a finding that is consistent with older children and adults (Terwogt & Hoeksma, 1995; Zentner, 2001).

A distinct shift in children's conceptions of emotions takes place between the ages of 6 and 9 years (La Greca, 1990). While children ages 6 and younger are likely to use idiosyncratic body cues (e.g. a smile) or situational cues (e.g. birthday party) to identify their emotions, children 8 to 9 years tend to use their own inner experiences and mental cues (Carroll & Steward, 1984; Harris, Olthof, & Terwogt, 1981). Between the age of 6 and 8 years, children also begin to accept the possibility of possessing two different feelings if the feelings are separated temporally, such as feeling scared in a dark room and then feeling happy when leaving the dark room (Harter, 1986; La Greca, 1990). It is not until between the ages of 8 and 12 years that children are able to accept the simultaneous co-occurrence of two emotions, such as feeling sad and happy at the same times (Carroll & Steward, 1984; Harter, 1986; La Greca, 1990).

Children's Language Skills

Children's language skills also influence their ability to self-report their HRQOL. Research on the development of language suggests that children acquire basic linguistic skills between the ages of 1 and 5 years (Bee & Mitchell, 1980; La Greca, 1990). Shifts in vocabulary skills, syntactic skills, and semantics occurs in a regular and fixed order in all children (La Greca, 1990). Between the ages of 8 and 18 months, children begin to speak their first meaningful words (La Greca, 1990). As children progress beyond speaking 2-word utterances, they begin to learn and apply rules of language such as the use of plural and possessive nouns, appropriate verb endings, and prepositions (Santrock, 2000). The child's vocabulary grows rapidly during the preschool years increasing from 50 words at age 1 ½ to approximately 900 words at age 3, to over 2000 words by age 5 (La Greca, 1990). During the elementary school years, changes in phonology, syntax, and semantics continue (Hetherington & Parke, 1979; La Greca, 1990). Included in these changes is the growth in vocabulary (over 2600 words by age 7), the learning of irregular verbs, the use of varied sentence structures, putting together a greater number of passive sentences, and comprehension of unique versus general attributes of concepts by age 8 to 9 years (La Greca, 1990).

While most children begin reading at approximately 5 to 6 years, young children's limited vocabulary has been shown to limit their reading comprehension (Biemiller, 2003). Research has also demonstrated great variability in decoding and reading comprehension abilities among young children (Biemiller, 2003; Cunningham & Stanovich, 1997). One study found that children in the second grade whose vocabulary size placed them in the highest quartile had an average estimated vocabulary of 7,100 root words, while children from the lowest quartile had an average vocabulary size of 3,000 words (cited in Biemiller 2003). Given such findings, it has been suggested that assessment techniques that are appropriate for older children, such as self-report questionnaires, should not be used for children under the age of 8 (De Civita et al., 2005; Eiser et al., 2000) Instead, it has been recommended that self-report questionnaires be administered as a structured interview, using simple and concrete language (Eiser, Mohay et al. 2000). Based on evidence indicating that young children have difficulty deciphering between multiple response choices and are more likely to endorse responses at the extreme ends of the continuum when using 5-point Likert-type rating scales (Chambers & Craig, 1998; Chambers & Johnston, 2002), it has also been recommended that self-report questionnaires for young children use a limited number of response choices, such as a 3-point Likert- type rating scale (Eiser et al., 2000; Ollendick, 1983). The tendency of young children to relate to images with which they can identify suggests that the use of clear and unambiguous pictures that relate to the topic may also enhance their comprehension on self-report measures (Eiser et al., 2000). Furthermore, it has been suggested that measures of HRQOL for young children be brief given the short attention span of young children (Eiser et al., 2000; Santrock, 2000).

Summary of Hypotheses

The primary objective of the current study is to investigate the age at which children can reliably and validly self-report their HRQOL using child self-report age subgroup data on over 8,000 children ages 5-16 years from the PedsQL™ 4.0 Generic

Core Scales Database[SM]. The secondary objective is to examine parent-child concordance on the PedsQL™ 4.0 Generic Core Scales using child self-report and parent proxy-report age subgroup data. We hypothesized that (a) items on the PedsQL™ 4.0 Generic Core Scales would have minimal missing responses across each age subgroup (b) the PedsQL™ 4.0 Generic Core Scales would demonstrate acceptable internal consistency reliability values (.70 or greater) for children as young as 5 years old and, (c) alpha coefficients would increase slightly with age, (d) healthy children would report significantly higher PedsQL™ 4.0 Generic Core Scales (better HRQOL) than pediatric patients with a known chronic health condition, showing reliable differences within each age subgroup, (e) factor loadings would be equivalent across age subgroups, (f) Intraclass Correlations (ICC) between child self-report and parent proxy-report will be greater for Physical Health Scores compared to Psychosocial Health Scores (g) Intraclass Correlations (ICC) between child self-report and parent proxy-report would be greater as the child increases in age, (h) mean differences between child self-report and parent proxy-report would be smaller for Physical Health Scores compared to Psychosocial Health Scores, and (i) mean differences between child self-report and parent proxy-report would decrease as the child's age increases.

METHOD

Participants

The sample contains composite published (n=8,086, 94.1%) and unpublished (n=505, 5.9%) child self-report and parent proxy-report age subgroup data on 8,591 children ages 5 to 16 years from the PedsQL™ 4.0 Generic Core Scales Database[SM] (Bhat et al., 2005; Goldstein et al., 2006; Hays et al., 2006; Palmer, Meeske, Katz, Burwinkle, & Varni, in press; Pohl, Anding, Wong, Grimes, & Varni, in press; Schwimmer, Burwinkle, & Varni, 2003; Seid, Sobo, Gelhard, & Varni, 2004; Uzark, Jones, Burwinkle, & Varni, 2003; Varni, Burwinkle, Berrin et al., 2006; Varni, Burwinkle, Jacobs et al., 2003; Varni, Burwinkle, Katz, Meeske, & Dickinson, 2002; Varni, Burwinkle, Rapoff, Kamps, & Olson, 2004; Varni, Burwinkle, Seid et al., 2003; Varni et al., in press; Varni, Seid et al., 2002; Varni et al., 2001a). Participants were recruited from general pediatric clinics, subspecialty clinics, and hospitals in which children were being seen for well-child checks, mild acute illness, or chronic illness care (n=2,603, 30.3%), and from a State Children's Health Insurance Program (SCHIP) in California (n=5,988, 69.7%). Participants recruited from general pediatric clinics, subspecialty clinics, and hospitals were assessed in-person or by telephone. For in-person mode of administration, research assistants obtained written parental informed consent and child assent. Paper-and-pencil questionnaires were self-administered for parents and for children ages 8 to 16 and interview administered for children ages 5 to 7 and in situations in which the child was unable to read or write as a consequence of either physical or cognitive impairment. For telephone administration, parents of children ages 5 to 16 were called by a research assistant who explained the study, and obtained verbal parental informed consent and child assent. The research assistant verbally administered the PedsQL™ 4.0 individually to the parent and their child. If the child was not home at the time of the initial call, the research assistant arranged for a call at another time.

Participants recruited from the State Children's Health Insurance Program (SCHIP) were assessed via statewide mailing. PedsQL™ 4.0 paper-and-pencil surveys were mailed separately for each of the months of February and March 2001 to families with children ages 5-16 years throughout the State of California who were all new enrollees in SCHIP. Parents and children ages 8-16 were instructed to complete the survey separately, while parents of children ages 5-7 were instructed to assist their child in completing the questionnaire after completing the proxy-report. A reminder postcard followed the initial mailing, with a second survey mailed to nonrespondents. At year 1 only, nonrespondents to the second survey received a telephone reminder. Since this survey was conceived of as an operational requirement and was conducted to comply with a California Insurance Code, parents and children did not complete informed consent forms.

As depicted in Table 1, for all forms combined (n=8,591), the number of children within each age subgroup is as follows: 757 five-year-olds (8.8%), 932 six-year-olds (10.8%), 891 seven-year-olds (10.4%), 882 eight-year-olds (10.3%), 841 nine-year-olds (9.8%), 841 ten-year-olds (9.8%), 683 eleven-year-olds (7.9%), 683 twelve-year-olds (7.9%), 614 thirteen-year-olds (7.1%), 572 fourteen-year-olds (6.7%), 563 fifteen-year-olds (6.6%), and 332 sixteen-year-olds (3.9%). The sample contains 4,391 boys (51.1%), 4,185 girls (48.7%), and 15 missing (0.2%). The sample is heterogeneous with respect to race/ethnicity with 4,403 Hispanics (51.3%), 1,995 White non-Hispanics (23.2%), 759 Asian or Pacific Islanders (8.8%), 405 Black non-Hispanics (4.7%), 41 American Indians or Alaskan Natives (0.5%), 115 other (1.3%), and 873 missing (10.2%). Child surveys were completed in English (n=4,859, 56.6%), Spanish (n= 3,377, 39.3%), Chinese (n= 184, 2.1%), Korean (n= 93, 1.1%), and Vietnamese (n= 46, 0.5%; missing=32, 0.4%). Response equivalence has been previously demonstrated across language for the PedsQL™ by examining the percent missing data, floor and ceiling effects, and scale internal consistency across language, as well as across mode of administration (Varni, Seid, & Kurtin, 2001).

The sample included healthy children, who were assessed either in physicians' offices during well-child checks and/or whose parents did not report the presence of a chronic health condition (n=5,491, 63.9%), acutely ill children, whose parents did not report the presence of a chronic health condition, but who were assessed at one of the pediatric clinics or hospitals (n= 142, 1.7%), and chronically ill children, whose parents reported the presence of a chronic health condition (n=2,627, 30.6%; missing=331, 3.9%). The chronically ill sample (n=2,627) is heterogeneous in terms of diagnoses, as shown in Table 1, with 374 children diagnosed with asthma (4.4%), 358 with cancer (4.2%), 291 with diabetes (3.4%), 269 with a gastrointestinal condition (3.1%), 268 with a rheumatic condition (3.1%), 199 with a cardiac condition (2.3%), 103 diagnosed as obese (1.2%), 103 with sickle cell anemia (1.2%), 78 with ADHD (0.9%), 72 with renal disease (0.8%), 71 with cerebral palsy (0.8%), 45 with mental health conditions (0.5%), and 396 with other conditions (15.1%).

Measures

Health-Related Quality of Life. Children and their parents completed the PedsQL™ 4.0 (Pediatric Quality of Life Inventory™ Version 4.0). The 23-item PedsQL™ 4.0 Generic Core Scales encompass: 1) Physical Functioning (8 items), 2) Emotional Functioning (5 items), 3) Social Functioning (5 items), and 4) School Functioning (5 items), and were developed through focus groups, cognitive interviews, pre-testing, and field testing measurement development protocols (Varni et al., 2001). The instrument takes approximately 5 minutes to complete (Varni et al., 2001).

The PedsQL™ 4.0 Generic Core Scales comprise parallel child self-report and parent proxy-report formats. Child self-report includes ages 5-7, 8-12, and 13-18 years. Parent proxy-report includes ages 2-4 (toddler), 5-7 (young child), 8-12 (child), and 13-18 (adolescent), and assesses parent's perceptions of their child's HRQOL. The items for each of the forms are essentially identical, differing in developmentally appropriate language (i.e. the use of simple and concrete language for the 5-7 child self-report), or first or third person tense. The instructions ask how much of a problem each item has been during the past one month. A 5-point response scale is utilized across child self-report for ages 8-18 and parent proxy-report (0 = never a problem; 1 = almost never a problem; 2 = sometimes a problem; 3 = often a problem; 4 = almost always a problem). To further increase the ease of use for the young child self-report (ages 5-7), the

response scale is reworded and simplified to a 3-point scale (0 = not at all a problem; 2 = sometimes a problem; 4 = a lot of a problem), with each response choice anchored to a happy to sad faces scale (Varni, Thompson, & Hanson, 1987; Varni et al., 1996).
Items are reverse-scored and linearly transformed to a 0-100 scale (0=100, 1=75, 2=50, 3=25, 4=0), so that higher scores indicate better HRQOL. Scale Scores are computed as the sum of the items divided by the number of items answered (this accounts for missing data). If more than 50% of the items in the scale are missing, the Scale Score is not computed. This accounts for the differences in sample sizes for scales reported in the Tables. Although there are other strategies for imputing missing values, this computation is consistent with the previous PedsQL™ peer-reviewed publications, as well as other well-established HRQOL measures (Fairclough, 1996; McHorney, Kosinski, & Ware, 1994; Varni et al., 2001a). The Physical Health Summary Score (8 items) is the same as the Physical Functioning Subscale. To create the Psychosocial Health Summary Score (15 items), the mean is computed as the sum of the items divided by the number of items answered in the Emotional, Social, and School Functioning Subscales

Demographic Information. The PedsQL™ Family Information Form (Varni et al., 2001a) or survey items adapted from the PedsQL™ Family Information Form were completed by parents. The PedsQL™ Family Information Form contains demographic information including the child's date of birth, sex, race/ethnicity, and information required to calculate the Hollingshead socioeconomic status (SES) index (Hollingshead, 1975). One survey question asks the parent to report on the presence of a chronic health condition ("In the past 6 months, has your child had a chronic health condition?")

defined as a physical or mental health condition that has lasted or is expected to last at least 6 months and interferes with the child's activities. If the parents check "Yes" to this question, they are asked to write in the name of the chronic health condition.

Analyses

The feasibility of child self-report was determined from the percentage of missing values across each age subgroup from 5 to 16 years (Fairclough, 1996; McHorney et al., 1994; Varni et al., 2001a). Scale internal consistency reliability was determined by calculating Cronbach's coefficient alpha across individual age subgroups (Cronbach, 1951). Scales with reliabilities of 0.70 or greater are recommended for comparing patient groups, while a reliability criterion of 0.90 is recommended for analyzing individual patient scale scores (Nunnally & Bernstein, 1994; Pedhazur & Schmelkin, 1991).

Construct validity was determined utilizing the known-groups method and multiple- group confirmatory factor analysis (CFA). The known-groups method compares scale scores across groups known to differ in the health construct being investigated. In this study, PedsQL™ 4.0 Generic Core Scales Scores in groups differing in known health condition (healthy children and children known to have a chronic illness) were computed across each age subgroup (McHorney, Ware, & Raczek, 1993; McHorney, Ware, Rogers, Raczek, & Lu, 1992), using independent sample *t*-tests. In order to determine the magnitude of the anticipated differences, effect sizes were calculated (Cohen, 1988). Effect size as used in these analyses was calculated by taking the difference between the healthy sample mean and the chronic sample mean, divided by the healthy sample standard deviation. Effect sizes for differences in means are designated as small (0.20), medium (0.50), and large (0.80) in magnitude (Cohen, 1988).

Multiple-group confirmatory factor analysis (CFA) was used to examine factor invariance across the age subgroups. Multiple-group CFA has been widely utilized in research examining factor invariance across groups such as cross cultural comparisons, gender comparisons, and ethnic comparisons (Vandenberg & Lance, 2000). Assumptions of factor invariant properties can be tested using multiple-group CFA, such as factor structure form, factor loadings, and factor uniqueness variances across a number of groups under a hypothetical model (Yao & Wu, 2005). According to the PedsQL™ 4.0 initial field test (Varni, Seid, & Kurtin, 2001b), a 5-factor model was appropriate for the PedsQL™ 4.0. Thus, a 5-factor model for CFA was specified. Multi-sample analyses can investigate multiple models at the same time with some or all parameters constrained to be equal over groups (Yao & Wu, 2005). For this study, the "All Free Loadings" and "Loadings Constrained Equal Across Groups" models were used. The "All Free Loadings" model constrained the factor loadings to be equal across the age subgroups, while the "Loadings Constrained Equal Across Groups" model had no constraints. We had an a priori expectation that the chi squared discrepancy test between the two models would be significant based on previous findings that indicate when using large samples, a statistically significant chi squared value can occur even when there are only minor differences in factor patterns among the groups (Cheung & Rensvold, 1999; Vandenberg & Lance, 2000). Thus, the change in CFI (comparative fit index) between the two models was calculated, in addition to three other indices of practical fit (RMSEA; PNFI; NNFI).

Concordance between child self-report and parent proxy-report was determined through two-way, mixed effect, absolute agreement, single measure Intraclass Correlations (ICC). The ICC offers an index that reflects the ratio between subject variability and total variability (Cremeens et al., 2006). Intraclass Correlations (ICC) are designated as $\leq$ 0.40 poor to fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 good agreement, and 0.81-1.00 excellent agreement (Bartko, 1966; Wilson, Dowling, Abdolell, & Tannock, 2001).

Agreement between child self-report and parent proxy-report was assessed across the age subgroups by calculating the absolute mean difference for the PedsQL™ Total Scale Scores, Physical Health Scores, and Psychosocial Health Scores (Cremeens et al., 2006). Paired sample *t*-tests were calculated to determine whether there was a significant difference between child self-report and parent proxy-report.

Statistical analyses were conducted using SPSS Version 13.0 for Windows and LISREL 8.8 for Windows.

RESULTS

Feasibility

Table 2 contains the percentage of missing item responses for child self-report across each age subgroup. Items on the PedsQL™ 4.0 Generic Core Scales had minimal missing responses for children as young as 5 years old, with the percentage of missing item responses decreasing slightly with age. The largest percentage of missing item responses was found in the 5 year old subgroup (2.1%), while the smallest percentage of missing item responses was found in the 16 year old subgroup (0.79%).

Internal Consistency Reliability

Internal consistency reliability alpha coefficients across individual age subgroups are presented for the PedsQL™ 4.0 Generic Core Scales Total Score in Table 3, Physical Health Score in Table 4, Psychosocial Health Score in Table 5, Emotional Functioning Score in Table 6, Social Functioning Score in Table 7, and School Functioning Score in Table 8. The majority of the child self-report scales across the age subgroups, including for children as young as 5 years, exceed the minimum reliability standard of 0.70 required for group comparisons, while the Total Scale Scores across the age subgroups approaches or exceeds the reliability criterion of 0.90 recommended for analyzing individual patient scale scores. Alpha values are lower for the School Functioning Scale across the age subgroups. Specifically, the alpha value is lowest on the School Functioning Scale for the 6 year old subgroup (α =0.59). Across the PedsQL™ scales and summary scales, internal consistency reliability alpha coefficients increase slightly with age.

Construct Validity

Tables 3 through 8 demonstrate comparisons between children's self-reported PedsQL™ 4.0 Generic Core Scales Total Scores, Physical Health Scores, Psychosocial Health Scores, Emotional Functioning Scores, Social Functioning Scores, and School Functioning Scores, for healthy children and children with a known chronic health condition by individual age subgroups. For each PedsQL™ scale and summary scale, across age subgroups, including children as young as 5 years, healthy children demonstrated a statistically significant difference in HRQOL (better HRQOL) than children with a known chronic health condition, with most effect sizes in the medium to large range (Cohen, 1988).

Table 9 demonstrates the χ^2 values with degrees of freedom (*df*) for model comparisons for the "All Free Loadings" and "Loadings Constrained Equal Across Groups" models. According to the χ^2 discrepancy test between the "All Free Loadings" model and "Loadings Constrained Equal Across Groups" model, contrasts were significant (χ^2 critical with 253 *df* = 291.10; χ^2 observed with 253 *df*=575.70; χ^2 observed > χ^2 critical therefore $p<.05$) suggesting that factorial invariance does not exist and "Loadings Constrained Equal Across Groups" model fits worse. However, as discussed in the analyses section, a statistically significant χ^2 value can occur even when there are only minor differences between the factor patterns in the age groups (Mulaik et al., 1989; Vandenberg & Lance, 2000), given that the χ^2 test is extremely sensitive to sample size (i.e., larger sample sizes produce larger chi-squares).

Thus, it has been recommended in the literature that a variety of indices of practical fit be utilized to supplement the χ^2 test as the goodness of model fit (Bentler, 1990; Mulaik et al., 1989; Vandenberg & Lance, 2000). In particular, Cheung & Rensvold (1999) suggest that examining the change in CFI (Comparative Fit Index) is superior to the χ^2 test (Cheung & Rensvold, 1999; Vandenberg & Lance, 2000). For multi-group confirmatory factor analysis, they propose that changes in CFI of -.01 or less or between -.01 and -.02 suggest factor invariance across groups (Cheung & Rensvold, 1999; Vandenberg & Lance, 2000). Table 9 demonstrates the CFI values for the "All Free Loadings" and "Loadings Constrained Equal Across Groups" models. The change in CFI is .001 (.944-.943). Thus, following these guidelines set forth by Cheung & Rensvold (1999), this suggests that the 5-factor model fits almost identically across the age subgroups.

Furthermore, the other indices of practical fit presented in Table 9 (RMSEA; PNFI; NNFI) indicate that the 5-factor model was acceptable across the age subgroups. However, it is worth noting that the practical fit indices (RMSEA; PNFI; NNFI) for the 6-year-old subgroup indicate that the 5-factor model does not fit as well in this age subgroup.

The PNFI (Parsimony Normed Fit Index) penalizes models that are less parsimonious so that more complex models have lower fit indexes (Mulaik et al., 1989). In Table 9, the PNFI value for the "All Free Loadings" model is .810 and the PNFI value for the "Loadings Constrained Equal Across Groups" model is .885. These values suggest that the "Loadings Constrained Equal Across Groups" model is a more parsimonious model (closer to .90 equals a better fit). Thus, based on this criterion, the "Loadings Constrained Equal Across Groups" model is chosen as the best fitting model.

Also illustrated in Table 9 are the factor loadings across the age subgroups and the "Loadings Constrained Equal Across Groups" model for each item on the PedsQL™. Given that the factor loadings for the 23 items are very similar across the age subgroups and the "Loadings Constrained Equal Across Groups" model, this is further evidence that the items are measuring similar HRQOL constructs for children across the age subgroups. In summary, children between the ages of 5 years and 16 years have nearly the same 5-factor HRQOL structure when self-reporting on the PedsQL™.

Parent/Child Concordance

Table 10 shows the two-way, mixed effect, absolute agreement, single measure Intraclass Correlations (ICC) between PedsQL™ 4.0 child self-report and parent proxy-report across the PedsQL™ scales and summary scales by individual age subgroups. Most intercorrelations are in the range of moderate to good agreement. Across the scales and summary scales, Intraclass Correlations (ICC) increase with the child's age. Intraclass Correlations (ICC) are greater for the Psychosocial Health Score as compared to the Physical Health Score across the age subgroups.

Tables 11 through 13 show mean differences between child self-report and parent proxy-report on the PedsQL™ 4.0 Generic Core Scales Total Scores, Physical Health Scores, and Psychosocial Health Scores across age subgroups. Children's ratings of their HRQOL were better than parents' proxy ratings of their children's HRQOL across age subgroups. Child self-report and parent proxy-report were significantly different for PedsQL™ 4.0 Total Scores, Physical Health Scores and Psychosocial Health Scores ($p < .001$) with the exception of the 6 and 7 year old subgroups on the Psychosocial Health Score. Across the age subgroups, mean differences were greatest for the Physical Health Scores (differences ranging from 3.58 to 6.57) and smallest for the Psychosocial Health Scores (differences ranging from 0.73 to 4.45). Mean differences did not become smaller as the child's age increased on the Total Scores, Physical Health Scores, and Psychosocial Health Scores.

DISCUSSION AND CONCLUSIONS

This study presents child self-report age subgroup data on over 8,000 children ages 5-16 years from the PedsQL™ 4.0 Generic Core Scales Database^{SM} to investigate the age at which children can reliably and validly self-report their HRQOL. The analyses support the reliability and validity of child self-report in children as young as 5 years old. Based on the literature that suggests the percentage of missing item responses have serious implications on parameter estimates as values approaches 15-20% of the total items on an instrument (Newman, 2003) , items on the PedsQL™ had minimal missing responses across the age subgroups (ranging from 0.79% to 2.1%), suggesting that children as young as 5 years are willing and able to provide quality data regarding their HRQOL. While the percentage of missing items was greatest for children in the 5 year old subgroup (2.1%), this can be attributed to a greater number of items left blank on the School Functioning Scale, which can be expected given that some 5-year-olds may not yet be enrolled in school. In fact, the percentage of missing item responses decreases to 1.1% for the 5 year old subgroup when the school items are excluded.

As hypothesized, internal consistency reliabilities generally exceeded the recommended minimum alpha coefficient standard of 0.70 for group comparisons across the age subgroups (Nunnally & Bernstein, 1994; Pedhazur & Schmelkin, 1991). PedsQL™ 4.0 Generic Core Scales Total Score for child self-report across the age subgroups approached or exceeded an alpha of 0.90, recommended for individual patient analysis (Nunnally & Bernstein, 1994), making the Total Scale Score suitable as a summary score for the primary analysis of HRQOL outcome in clinical trials and other group comparisons. The demonstration of alpha coefficients increasing slightly with age across scales and summary scales suggests that as children develop more advanced mental operations, they become more reliable reporters of their HRQOL. There are two additional explanations for the lower alpha values for young children 5-7 years, especially on the School Functioning Scale. There is evidence suggesting that 2-and-3 point item Likert scales have lower reliability than 5-or-7 point item Likert scales (McHorney & Cohen, 2000; Preston & Colman, 2000). Thus, the use of a 3-point Likert scale on the young child self-report form may also be contributing to the lower alpha values for young children across the across scales and summary scales. In addition, inter-item correlations among items affect Cronbach's alpha values (Schmitt, 1996). Thus, if inter-item correlations are high, there is evidence that the items are measuring a single unidimensional latent construct, and consequently, the alpha value for the scale will also be high (Schmitt, 1996). However, if the average inter-item correlations are low, there is evidence that the items are measuring more than one unidimensional latent construct, and alpha values will consequently be lower (Schmitt, 1996). Given evidence from the PedsQL™ 4.0 initial field test (Varni et al., 2001a) that demonstrates the School Functioning items split into two different factors, the low inter-item correlation among items on this scale may be contributing to the lower alphas values.

The PedsQL™ 4.0 Generic Core Scales performed as hypothesized utilizing the known-groups method. Across each individual age group, including children as young as 5 years, the PedsQL™ 4.0 differentiated HRQOL in healthy children as a group in comparison to children with a known chronic health condition, suggesting that children as young as 5 years are valid reporters of their HRQOL.

The demonstration of the 5-factor model fitting almost identically across the age subgroups is further evidence that children as young as 5 years old are reliable and valid self-reporters of their HRQOL. Given that items on the PedsQL™ represent similar constructs for children between the ages of 5 years and 16 years, meaningful comparisons can be made between mean differences in PedsQL™ scores across the age subgroups. While the indices of practical fit for the 6-year-old subgroup demonstrated that the 5-factor model did not fit as well for this age group, it is interesting to note that the factor loadings across the 23 items for the 6-year-old subgroup were similar to the factor loadings across the other age subgroups. Furthermore, there is a general trend on some items (Phys 2, 3, 4, 8; Psyc 1, 3, 4; Soc 1, 2, 3; Sch 2; MsSch 1) to have higher loading as children become older. This suggests that despite adequate fit for the "All Loadings Constrained Equal Across Groups" model, there is also a trend on some items for construct validity to become slightly better as the respondents get older.

The cross-informant variance observed in the ICC's and mean differences support the need to measure the perspectives of child and parent informants in evaluating HRQOL in pediatric populations. Contrary to our hypotheses that there would be greater parent-child concordance for more observable HRQOL, we found greater parent-child concordance on the Psychosocial Health Scores compared to the Physical Health

Scores across age subgroups using both ICC's and parent-child mean score differences. While there is evidence that parent-child agreement is generally greater for more observable domains such as physical functioning (Baxt et al., 2004; Eiser & Morse, 2001a; Levi & Drotar, 1999; Varni et al., 1999), in a review of the literature, Eiser & Morse (2001) found limited support for this widely held belief (Eiser & Morse, 2001b). In particular, as cited in Eiser & Morse (2001), Bruil (1999) found good parent-child agreement for physical complaints in a sample of chronically ill children, but not in a comparable group of healthy children (Bruil, 1999; Eiser & Morse, 2001b). Thus, given that our sample was predominantly healthy children (63.9%), consistent with the findings of Bruil (1999), lower parent-child agreement on the Physical Health Scores may have been influenced by the children's illness status. Our hypothesis was supported that ICC's would be greater as children's age increased, however, mean child-parent differences did not become smaller as the child's age increased. This can in part be explained by the fact that the ICC takes into account the ratio between subject variability and total variability (Cremeens et al., 2006), while mean child-parent differences does not take this variability into account.

Limitations of the Present Study

The present findings have several potential limitations. The method of contacting respondents and the medium of delivering the questionnaire varied across the sample. These differences could have had varying effects on the data collected; however, response equivalence has been previously demonstrated across mode of administration for the PedsQL™(Varni, Seid, & Kurtin, 2001), and differences in data collection methods may increase the generalizability of the findings. Another potential limitation is the use of a 3 point Likert-type scale for the young child self-report form (ages 5-7 years), compared to the 5 point Likert-type scale used for the child self-report forms for children ages 8 to 16 years. However, given the literature that indicates young children have difficulty deciphering between multiple response choices and are more likely to endorse responses at the extreme ends of the continuum when using 5 point Likert-type rating scales (Chambers & Craig, 1998; Chambers & Johnston, 2002), the use of a 3 point Likert-type scale for the young child self-report form may have increased the likelihood of young children providing reliable and valid responses (Eiser et al., 2000; Ollendick, 1983). An additional limitation of this study is the differences in sample size across the age groups. While most of the age groups had a sample size above or near 600 participants, the 16 year old subgroup had 332 participants. Given that comparisons were made between each age subgroup, it would have been ideal to have more evenly distributed sample sizes across the age subgroups. On the other hand, the large sample sizes make these findings robust.

Implications for Future Research

Future research will use multiple-group confirmatory factor analysis to investigate the factor structure among parent versus child groups and chronic versus healthy groups. Multiple-group confirmatory factor analysis will also be used to compare the factor structures across various chronic health conditions. In addition, future research will also examine the estimates of alpha's standard error in order to compare the significance of alpha coefficients among the age subgroups. Finally, future research will include conducting a meta-analysis of the HRQOL literature to examine various issues surrounding parent-child agreement. In particular, the specific types of statistical analyses used to determine parent-child agreement will be analyzed and recommendations about the most sound statistical analyses will be made based on the literature.

REFERENCES

Annett, R., Bender, B., DuHamel, T., & Lapidus, J. (2003). Factors influencing parent reports on quality of life for children with asthma. *J Asthma, 40*, 577-587.

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports, 19*, 3-11.

Baxt, C., Kassam-Adams, N., Nance, M., Vivarelli-O'Neill, C., & Winston, F. (2004). Assessment of pain after injury in the pediatric patient: Child and parent perceptions *Journal of Pediatric Surgery 39*(6), 979-983.

Bee, H., & Mitchell, S. (1980). *The developing person: A life-span approach*. New York: Harper and Row.

Bentler, P. (1990). Comparative fit indixes in structural models. *Psychological Bulletin, 107*, 238-246.

Bhat, S. R., Goodwin, T. L., Burwinkle, T. M., Lansdale, M. F., Dahl, G. V., Huhn, S. L., Gibbs, I. C., Donaldson, S. S., Rosenblum, R. K., Varni, J. W., & Fisher, P. G. (2005). Profile of daily life in children with brain tumors: An assessment of health-related quality of life. *Journal of Clinical Oncology, 23*, 5493-5500.

Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology 24*, 323-335.

Bretherton, I., Fritz, J., Zahn, C., & Ridgeway, D. (1986). Learning to talk about emotions: A functional perspective. *Child Development, 57*, 529-548.

Bruil, J. (1999). *Development of a quality of life instrument for children with chronic illness*. Leiden: Leiden University Press.

Carroll, J., & Steward, M. (1984). The role of cognitive development in children's understanding of their own feelings. *Child Development, 55*, 1486-1492.

Chambers, C., & Craig, K. (1998). An intrusive impact of anchors in children's faces pain scales *Pain 78*, 27-37.

Chambers, C., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology 27*, 27-36.

Cheung, G., & Rensvold. (1999). *What constitutes significant differences in evaluating measurement invariance?*, Conference of the Academy of Management. Chicago

Clancy, C., McGrath, P., & Oddson, B. (2005). Pain in children and adolescents with spina bifida. *Developmental Medicine and Child Neurology 47*(1), 27-34.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed ed.). Hillsdale, NJ: Erlbaum.

Cremeens, J., Eiser, C., & Blades, M. (2006). Factors influencing agreement between child self-report and parent proxy-reports on the Pediatric Quality of Life Inventory™ 4.0 (PedsQL™) generic core scales. *Health and Quality of Life Outcomes 4*(58).

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

Cunningham, A., & Stanovich, K. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934–945.

De Civita, M., Regier, D., Alamgir, A., Anis, A., FitzGerald, M., & Marra, C. (2005). Evaluating health-related quality of life studies in paediatric populations: Some conceptual, methodological, and developmental considerations and recent applications. *Pharmacoeconomics, 23*(7), 659-685.

De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment, 16*, 330-334.

Duhig, A., Renk, K., Epstein, M., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice, 7*, 435-453.

Eiser, C. (2004). Use of quality of life measures in clinical trials. *Ambulatory Pediatrics, 4*, 395-399.

Eiser, C., Mohay, H., & Morse, R. (2000). The measurement of quality of life in young children. *Child:Care, Health and Development, 26*(5), 401-414.

Eiser, C., & Morse, R. (2001a). Can parents rate their child's health-related quality of life?: Results from a systematic review. *Quality of Life Research, 10*, 347-357.

Eiser, C., & Morse, R. (2001b). Quality of life measures in chronic diseases of childhood. *Health Technology Assessment, 5*, 1-158.

Eiser, C., & Morse, R. (2001c). A review of measures of quality of life for children with chronic illness. *Archives of Diseases in Children, 84*, 205-211.

Fairclough, D. L. (1996). Quality of life in cancer clinical trials: Now that we have the data, what do we do? *JASS, 4(4)*, 253-269.

Fekkes, M., Theunissen, N., Brugman, E., & Veen, S. (2000). Development and psychometric evaluation of the TAPQOL: A health-related quality of life instrument for 1-5 year old children. *Quality of Life Research 9*, 961-972.

Forrest, C. B., SA, S., Dougherty, D., & MR, M. (2003). Outcomes research in pediatric settings: Recent trends and future directions. *Pediatrics, 111*, 171-178.

Goldstein, S., Graham, N., Burwinkle, T., Warady, B., Farrah, R., & Varni, J. W. (2006). Health-related quality of life in pediatric patients with ESRD. *Pediatric Nephrology 21*, 846-850.

Harris, P., Olthof, T., & Terwogt, M. (1981). Children's knowledge of emotion. *Journal of Child Psychology and Psychiatry, 22*, 247-261.

Harter, S. (1986). Cognitive-developmental processes in the integration of concepts about emotions and self. *Social Cognition 4*, 119-151.

Hays, R., Valentine, J., Haynes, G., Geyer, J., Villareale, N., McKinstry, B., Varni, J. W., & Churchill, S. (2006). The Seattle Pediatric Palliative Care Project: Effects on family satisfaction and health-related quality of life. *Journal of Palliative Medicine, 9*, 716-728.

Hetherington, E., & Parke, R. (1979). *Child psychology: A contemporary viewpoint.* New York: McGraw-Hill.

Hollingshead, A. B. (1975). *Four factor index of social status*. New Haven, CT: Yale University.

Keller, A., Ford, L., & Meacham, J. (1978). Dimensions of self-concept in preschool children *Developmental Psychology, 14*, 483-489.

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

La Greca, A. (1990). *Through the eyes of the child*. Boston: Allyn and Bacon.

Landgraf, J., & Abetz, L. (1997). Measuring health outcomes in pediatric populations: Issues in psychometrics and application. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 793-802). Philadelphia: Lippincott-Raven

Levi, R., & Drotar, D. (1999). Health-related QoL in childhood cancer: Discrepancy in parent–child reports. *International Journal of Cancer, 12*, 58-64.

Luoma, I., Koivisto, A., & Tamminen, T. (2004). Fathers' and mothers' perceptions of their child and maternal depressive symptoms. *Nordic Journal of Psychiatry, 58*(3), 205-211.

McHorney, C., & Cohen, A. (2000). Equating health status measures with item response theory: Illustrations with functional status items. *Medical Care 38*(9), II-43 - II-59.

McHorney, C. A., Kosinski, M., & Ware, J. E. (1994). Comparisons of the costs and quality of norms for the SF-36 Health Survey collected by mail versus telephone interview: Results from a national survey. *Medical Care, 32*, 551-567.

McHorney, C. A., Ware, J. E., & Raczek, A. E. (1993). The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care, 31*, 247-263.

McHorney, C. A., Ware, J. E., Rogers, W., Raczek, A. E., & Lu, J. F. R. (1992). The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts: Results from the Medical Outcomes Study. *Medical Care, 30*, MS253-MS265.

Mulaik, S., James, L., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychol Bull, 105*, 430-445.

Newacheck, P. W., & Taylor, W. R. (1992). Childhood chronic illness: Prevalence, severity, and impact. *American Journal of Public Health, 82*, 364-371.

Newman, D. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6*(3), 328-362.

Nunnally, J. C., & Bernstein, I. R. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Ollendick, T. H. (1983). Reliability and validity of the revised Fear Survey Schedule for Children (FSSC-R). *Behaviour Research and Therapy, 21*, 685-692.

Palmer, S., Meeske, K., Katz, E., Burwinkle, T., & Varni, J. W. (2006). The PedsQL™ Brain Tumor Module: Initial reliability and validity. *Pediatric Blood and Cancer* (in press).

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Perrin, E., & Gerrity, S. (1981). There's a demon in your belly: Children's understanding of illness. *Pediatrics, 67*(6), 139-148.

Piaget, J. (1930). *The child's conception of physical causality* London: Kegan Paul.

Pohl, J., Anding, J., Wong, W., Grimes, R., & Varni, J. W. (2006). Pediatric obesity in Texas: Does the Texas Public School Nutrition Policy affect child nutrition *Texas Medicine* (in press).

Preston, C., & Colman, A. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica 104*(1), 1-15.

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*, 485-499.

Santrock, J. (2000). *Children*. Dallas: McGraw- Hill.

Schmitt, N. (1996). Uses and abuses of coefficient alpha *Psychological Assessment 8*(4), 350-353.

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.

Seid, M., Sobo, E. J., Gelhard, L. R., & Varni, J. W. (2004). Parents' reports of barriers to care for children with special health care needs: Development and validation of the Barriers to Care Questionnaire. *Ambulatory Pediatrics, 4*, 323-331.

Terwogt, M., & Hoeksma, J. (1995). Colours and preferences: Preferences and combinations *Journal of General Psychology 122*, 5-17.

Theunissen, N., Vogels, T., Koopman, H., Verrips, K., Zwinderman, S., Verloove, S., & Wit, J. (1998). The proxy problem: Child report versus parent report in health-related quality of life research. *Quality of Life Research, 7*, 387-397.

Theunissen, N. C. M., Vogels, T. G. C., Koopman, H. M., Verrips, G. H. W., Zwinderman, K. A. H., Verloove-Vanhorick, S. P., & Wit, J. M. (1998). The proxy problem: Child report versus parent report in health-related quality of life research. *Quality of Life Research, 7*, 387-397.

Truetler, C., & Epkins, C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent–child relationships? *Journal of Abnormal Child Psychology, 31*, 13-27.

Uzark, K., Jones, K., Burwinkle, T. M., & Varni, J. W. (2003). The Pediatric Quality of Life Inventory in children with heart disease. *Progress in Pediatric Cardiology, 18*, 141-148.

Vandenberg, R., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research *Organizational Research Methods 3*, 4-70.

Varni, J. W., Burwinkle, T. M., Berrin, S. J., Sherman, S. A., Artavia, K., Malcarne, V. L., & Chambers, H. G. (2006). The PedsQL in pediatric cerebral palsy: Reliability, validity, and sensitivity of the Generic Core Scales and Cerebral Palsy Module. *Developmental Medicine and Child Neurology, 48,* 442-449.

Varni, J. W., Burwinkle, T. M., Jacobs, J. R., Gottschalk, M., Kaufman, F., & Jones, K. L. (2003). The PedsQL in Type 1 and Type 2 diabetes: Reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales and Type 1 Diabetes Module. *Diabetes Care, 26*, 631-637.

Varni, J. W., Burwinkle, T. M., Katz, E. R., Meeske, K., & Dickinson, P. (2002). The PedsQL in pediatric cancer: Reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer, 94*, 2090-2106.

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 and Quality of Life Outcomes, 3:34*, 1-9.

Varni, J. W., Burwinkle, T. M., Rapoff, M. A., Kamps, J. L., & Olson, N. (2004). The PedsQL in pediatric asthma: Reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales and Asthma Module. *Journal of Behavioral Medicine, 27*, 297-318.

Varni, J. W., Burwinkle, T. M., & Seid, M. (2005). The PedsQL as a pediatric patient-reported outcome: Reliability and validity of the PedsQL Measurement Model in 25,000 children *Expert Review of Pharmacoeconomics and Outcomes Research, 5*, 705-719.

Varni, J. W., Burwinkle, T. M., & Seid, M. (2006). The PedsQL 4.0 as a school population health measure: Feasibility, reliability, and validity. *Quality of Life Research, 15*, 203-215.

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.

Varni, J. W., Katz, E. R., Colegrove, R., & Dolgin, M. (1995). Adjustment of children with newly diagnosed cancer: Cross-informant variance. *Journal of Psychosocial Oncologyl, 13*, 23-38.

Varni, J. W., Katz, E. R., Seid, M., Quiggins, D. J. L., & Friedman-Bender, A. (1998). The Pediatric Cancer Quality of Life Inventory-32 (PCQL-32): I. Reliability and validity. *Cancer, 82*, 1184-1196.

Varni, J. W., Katz, E. R., Seid, M., Quiggins, D. J. L., Friedman-Bender, A., & Castro, C. M. (1998). The Pediatric Cancer Quality of Life Inventory (PCQL): I. Instrument development, descriptive statistics, and cross-informant variance. *Journal of Behavioral Medicine, 21*, 179-204.

Varni, J. W., Lane, M., Burwinkle, T., Fontaine, E., Youssef, N., Schwimmer, J., Pardee, P., Pohl, J., & Easley, D. (2006). Health-related quality of life in pediatric patients with irritable bowel syndrome: A comparative analysis. *Journal of Developmental and Behavioral Pediatrics* (in press).

Varni, J. W., Seid, M., Knight, T. S., Burwinkle, T. M., Brown, J., & Szer, I. S. (2002). The PedsQL in pediatric rheumatology: Reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. *Arthritis and Rheumatism, 46*, 714-725.

Varni, J. W., Seid, M., & Kurtin, P. S. (2001). The 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.

Varni, J. W., Seid, M., & Rode, C. A. (1999). The PedsQL: Measurement model for the Pediatric Quality of Life Inventory. *Medical Care, 37*, 126-139.

Varni, J. W., Thompson, K. L., & Hanson, V. (1987). The Varni/Thompson Pediatric Pain Questionnaire: I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. *Pain, 28*, 27-38.

Varni, J. W., Waldron, S. A., Gragg, R. A., Rapoff, M. A., Bernstein, B. H., Lindsley, C. B., & Newcomb, M. D. (1996). Development of the Waldron/Varni Pediatric Pain Coping Inventory. *Pain, 67*, 141-150.

Waters, E., Doyle, J., Wolfe, R., Wright, M., & Wake, M. (2000). Influence of parental gender and self-reported health and illness on parent-reported child health. *Pediatrics, 106*(6), 1422-1428.

Wilson, K. A., Dowling, A. J., Abdolell, M., & Tannock, I. F. (2001). Perception of quality of life by patients, partners and treating physicians. *Quality of Life Research, 9*, 1041-1052.

Yao, G., & Wu, C.-h. (2005). Factorial invariance of the WHOQOL-BREF among disease groups. *Quality of Life Research, 14*, 1881-1888.

Yasuda, N., Zimmerman, S., Hawkes, W., Gruber-Baldini, A., Hebel, J., & Magaziner, J. (2004). Concordance of proxy-perceived change and measured change in multiple domains of function in older persons. *J Am Geriatr Soc, 52*(7), 1157-1162.

Zentner, M. (2001). Preferences for colors and color- emotion combinations in early childhood *Developmental Science, 4*(4), 389-398.

APPENDIX

TABLES

Table 1. Total sample by individual age subgroups and chronically ill sample by diagnosis

Total Sample (N=8,591)			Chronically Ill Sample (N=2,627)	
Age	N		Diagnosis	N
5 years	757 (8.8%)		Asthma	374 (4.4%)
6 years	932 (10.8%)		Cancer	358 (4.2%)
7 years	891 (10.4%)		Diabetes	291 (3.4%)
8 years	882 (10.3%)		Gastrointestinal Condition	269 (3.1%)
9 years	841 (9.8%)		Rheumatic Condition	268 (3.1%)
10 years	841 (9.8%)		Cardiac Condition	199 (2.3%)
11 years	683 (7.9%)		Obesity	103 (1.2%)
12 years	683 (7.9%)		Sickle Cell Anemia	103 (1.2%)
13 years	614 (7.1%)		ADHD	78 (0.9%)
14 years	572 (6.7%)		Renal Disease	72 (0.8%)
15 years	563 (6.6%)		Cerebral Palsy	71 (0.8%)
16 years	332 (3.9%)		Mental Health Conditions	45 (0.5%)
			Other Conditions	396 (15.1%)

Table 2. Percentage of missing item responses for child self-report across individual age subgroups

Total Sample (N=8,591)	
Age	% of missing item responses
5 years	2.1%
6 years	1.3%
7 years	1.5%
8 years	1.4%
9 years	1.0%
10 years	0.96%
11 years	0.85%
12 years	0.95%
13 years	0.62%
14 years	0.95%
15 years	0.95%
16 years	0.79%

Note: When the School Functioning items are excluded, the percentage of missing item responses is 1.1% for the 5 year olds, 0.97% for the 6 year olds, 0.97% for the 7 year olds, and 1.1% for the 8 year olds.

Table 3. PedsQL™ 4.0 Generic Core Scales Total Score: Child Self-Report Reliability and Validity

Age Size	Reliability n	α	Validity: Chronic Health Condition Mean	SD	Healthy Sample Mean	SD	Difference	Effect
5 yrs	693	0.86	74.83	15.41	83.22	12.18	8.39*	0.69
6 yrs	913	0.86	76.24	14.82	82.12	12.73	5.88*	0.46
7 yrs	869	0.87	71.21	15.42	80.98	12.98	9.77*	0.75
8 yrs	864	0.90	75.43	14.73	83.54	12.95	8.11*	0.63
9 yrs	827	0.91	74.16	15.75	83.71	13.76	9.55*	0.69
10 yrs	825	0.90	76.13	15.84	84.16	12.72	8.03*	0.63
11 yrs	675	0.91	76.63	15.56	85.61	12.47	8.98*	0.72
12 yrs	669	0.91	76.57	15.48	84.01	12.97	7.44*	0.57
13 yrs	609	0.92	75.08	16.06	84.23	13.15	9.15*	0.70
14 yrs	560	0.92	74.10	16.34	85.71	11.97	11.61*	0.97
15 yrs	554	0.91	75.03	14.97	84.70	12.73	9.68*	0.76
16 yrs	327	0.91	74.67	16.52	85.76	11.41	11.09*	0.97

Note: Total N= 8,385 for reliability, Total N = 8,098 for validity.
*p<.001, (independent samples t-test). Higher values equal better health-related quality of life.
Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 4. PedsQL™ 4.0 Generic Core Scales Physical Health Score: Child Self-Report Reliability and Validity

Age	Reliability		Validity					
			Chronic Health Condition		Healthy Sample			
Size	n	α	Mean	SD	Mean	SD	Difference	Effect
5 yrs	753	0.72	77.65	19.42	86.34	13.38	8.69*	0.65
6 yrs	928	0.70	80.31	17.20	86.18	13.23	5.87*	0.44
7 yrs	885	0.73	75.28	19.18	85.52	13.73	10.24*	0.75
8 yrs	875	0.79	77.47	17.73	87.36	13.77	9.88*	0.72
9 yrs	838	0.83	76.37	18.45	88.43	14.80	12.06*	0.81
10 yrs	837	0.82	77.47	19.34	88.03	13.37	10.56*	0.79
11 yrs	681	0.84	78.34	19.29	88.03	13.30	9.68*	0.73
12 yrs	679	0.83	77.92	18.84	87.13	13.95	9.21*	0.66
13 yrs	614	0.84	76.16	19.19	87.83	13.57	11.67*	0.86
14 yrs	570	0.87	74.37	22.05	89.00	13.20	14.62*	1.11
15 yrs	560	0.85	77.21	18.58	88.59	13.55	11.38*	0.84
16 yrs	331	0.87	75.55	21.75	89.13	12.74	13.58*	1.07

Note: Total N= 8,551 for reliability, Total N = 8,325 for validity.
*p<.001, (independent samples t-test). Higher values equal better health-related quality of life.
Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 5. PedsQL™ 4.0 Generic Core Scales Psychosocial Health Score: Child Self-Report Reliability and Validity

Age	Reliability		Validity					
			Chronic Health Condition		Healthy Sample			
	n	α	Mean	SD	Mean	SD	Difference	Effect Size
5 yrs	696	0.82	73.16	16.30	81.56	13.76	8.40*	0.61
6 yrs	914	0.82	74.06	16.13	79.91	14.40	5.85*	0.41
7 yrs	870	0.83	69.14	16.68	78.55	14.73	9.41*	0.64
8 yrs	867	0.86	74.27	15.53	81.53	14.25	7.26*	0.51
9 yrs	829	0.87	73.01	16.90	81.20	14.99	8.19*	0.55
10 yrs	829	0.87	75.44	16.82	82.09	14.06	6.65*	0.47
11 yrs	675	0.89	75.77	16.41	84.32	13.45	8.55*	0.64
12 yrs	672	0.88	75.86	16.14	82.20	14.45	6.33*	0.44
13 yrs	610	0.89	74.46	16.69	82.34	14.69	7.87*	0.54
14 yrs	562	0.88	73.99	16.20	83.98	13.28	9.99*	0.75
15 yrs	556	0.88	73.89	15.31	82.72	14.21	8.83*	0.62
16 yrs	328	0.88	74.21	16.36	83.98	13.03	9.78*	0.75

Note: Total N= 8,408 for reliability, Total N = 8,324 for validity.
*p<.001, (independent samples t-test). Higher values equal better health-related quality of life.
Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 6. PedsQL™ 4.0 Generic Core Scales Emotional Functioning Score: Child Self-Report Reliability and Validity

Age	Reliability		Validity					
			Chronic Health Condition		Healthy Sample			
	n	α	Mean	SD	Mean	SD	Difference	Effect Size
5 yrs	750	0.70	72.83	23.36	79.11	17.88	6.28*	0.35
6 yrs	926	0.70	75.45	21.73	78.73	18.46	3.29**	0.18
7 yrs	886	0.73	69.44	22.10	77.38	19.07	7.95*	0.42
8 yrs	874	0.75	70.96	19.37	77.53	18.97	6.57*	0.35
9 yrs	840	0.77	72.01	21.17	78.17	18.28	6.16*	0.34
10 yrs	840	0.77	73.93	19.94	79.88	17.72	5.96*	0.34
11 yrs	680	0.79	74.60	20.41	81.04	17.51	6.44*	0.37
12 yrs	678	0.79	73.68	19.72	79.22	18.24	5.54*	0.30
13 yrs	615	0.81	72.55	21.22	80.13	18.15	7.57*	0.42
14 yrs	572	0.79	72.92	20.22	81.02	17.78	8.10*	0.46
15 yrs	562	0.81	70.90	21.20	79.72	18.57	8.82*	0.47
16 yrs	332	0.81	70.74	21.97	80.15	17.99	9.41*	0.52

Note: Total N= 8,555 for reliability, Total N = 8,319 for validity.
*p<.001, ** p<.05 (independent samples t-test). Higher values equal better health-related quality of life.
Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 7. PedsQL™ 4.0 Generic Core Scales Social Functioning Score: Child Self-Report Reliability and Validity

Age Size	Reliability n	α	Validity: Chronic Health Condition Mean	SD	Healthy Sample Mean	SD	Difference	Effect
5 yrs	749	0.68	74.13	19.31	83.39	16.82	9.26*	0.55
6 yrs	925	0.68	75.87	20.29	81.06	17.73	5.19*	0.29
7 yrs	885	0.71	69.58	22.10	80.25	18.02	10.67*	0.59
8 yrs	876	0.74	78.72	18.65	85.00	16.58	6.28*	0.38
9 yrs	837	0.79	77.21	20.73	84.16	18.54	6.96*	0.38
10 yrs	837	0.77	80.08	20.36	85.47	16.84	5.39*	0.32
11 yrs	680	0.80	80.47	19.65	88.53	14.99	8.06*	0.54
12 yrs	679	0.78	81.78	19.75	87.27	15.47	5.48*	0.35
13 yrs	612	0.81	81.32	18.25	86.84	16.47	5.52*	0.34
14 yrs	571	0.83	78.88	21.53	89.82	14.65	10.94*	0.75
15 yrs	562	0.81	82.99	17.59	89.06	14.49	6.07*	0.42
16 yrs	331	0.77	83.37	17.69	90.29	12.63	6.92*	0.55

Note: Total N= 8,544 for reliability, Total N = 8,308 for validity.
*p<.001, (independent samples t-test). Higher values equal better health-related quality of life.
Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 8. PedsQL™ 4.0 Generic Core Scales School Functioning Score: Child Self-Report Reliability and Validity

Age	Reliability		Validity					
			Chronic Health Condition		Healthy Sample			
Size	n	α	Mean	SD	Mean	SD	Difference	Effect
5 yrs	702	0.63	72.24	20.78	82.02	15.92	9.77*	0.61
6 yrs	920	0.59	71.81	18.11	79.93	16.56	8.12*	0.49
7 yrs	873	0.62	68.15	18.15	77.90	17.16	9.76*	0.57
8 yrs	872	0.68	73.27	17.39	82.11	15.56	8.84*	0.57
9 yrs	831	0.74	69.75	19.59	81.05	17.15	11.30*	0.66
10 yrs	832	0.72	72.76	19.45	80.95	16.32	8.19*	0.50
11 yrs	676	0.76	72.37	19.37	83.43	15.56	11.05*	0.71
12 yrs	677	0.75	71.94	19.30	80.25	17.26	8.30*	0.48
13 yrs	612	0.79	69.56	21.34	80.17	17.13	10.61*	0.62
14 yrs	563	0.76	69.94	18.14	81.17	17.07	11.23*	0.66
15 yrs	556	0.78	67.78	19.02	79.33	17.85	11.55*	0.65
16 yrs	328	0.77	68.36	19.93	81.49	17.41	13.12*	0.75

Note: Total N= 8,442 for reliability, Total N = 8,211 for validity.
*p<.001, (independent samples t-test). Higher values equal better health-related quality of life.
Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 9. Multi-Group Confirmatory Factor Analysis: Factor Loadings, Fit Indices, and χ^2 Values (with *df*)

													A	**B**
Item	**5 yrs.**	**6 yrs.**	**7 yrs.**	**8 yrs.**	**9 yrs.**	**10 yrs.**	**11 yrs.**	**12 yrs.**	**13 yrs.**	**14 yrs.**	**15 yrs.**	**16 yrs.**	**All free loadings**	**Factor loadings constrained equal**
Phys1	.59	.57	.56	.57	.69	.66	.69	.71	.70	.73	.67	.78	x	.66
Phys2	.68	.59	.60	.65	.73	.78	.80	.80	.80	.86	.85	.84	x	.75
Phys3	.63	.53	.64	.69	.78	.78	.81	.83	.79	.86	.85	.85	x	.76
Phys4	.38	.40	.42	.55	.56	.57	.60	.61	.61	.66	.64	.70	x	.56
Phys5	.51	.38	.41	.47	.47	.37	.39	.31	.33	.49	.31	.43	x	.41
Phys6	.44	.47	.42	.46	.52	.47	.44	.42	.48	.52	.47	.59	x	.47
Phys7	.42	.52	.50	.53	.63	.59	.54	.58	.57	.60	.61	.55	x	.55
Phys8	.43	.50	.58	.63	.61	.61	.70	.62	.69	.73	.71	.66	x	.62
Psyc1	.59	.58	.63	.69	.70	.67	.73	.67	.75	.73	.75	.73	x	.69
Psyc2	.67	.64	.65	.66	.67	.72	.74	.73	.75	.73	.78	.77	x	.71
Psyc3	.53	.57	.54	.62	.59	.67	.67	.64	.69	.65	.64	.64	x	.62
Psyc4	.48	.49	.52	.50	.59	.53	.54	.58	.59	.56	.59	.65	x	.55
Psyc5	.57	.56	.62	.63	.64	.62	.66	.70	.66	.67	.70	.66	x	.64
Soc1	.58	.60	.59	.56	.66	.63	.65	.70	.68	.72	.72	.64	x	.64
Soc2	.52	.54	.58	.59	.63	.59	.68	.64	.68	.69	.78	.68	x	.63
Soc3	.48	.55	.58	.55	.63	.59	.67	.62	.69	.63	.74	.62	x	.61
Soc4	.50	.51	.53	.64	.67	.68	.66	.62	.68	.74	.55	.58	x	.63
Soc5	.64	.56	.60	.66	.66	.67	.67	.65	.69	.75	.62	.70	x	.66
Sch1	.65	.63	.63	.60	.70	.74	.76	.79	.78	.75	.73	.72	x	.71
Sch2	.52	.54	.61	.65	.68	.70	.72	.72	.70	.72	.74	.69	x	.67
Sch3	.71	.66	.68	.68	.77	.74	.78	.76	.79	.72	.76	.76	x	.74
MsSch1	.65	.68	.73	.84	.80	.75	.79	.77	.84	.79	.84	.82	x	.77
MsSch2	.69	.66	.61	.63	.71	.75	.78	.69	.69	.66	.69	.71	x	.68
Fit Indices														
RMSEA	.06	.08	.06	.06	.07	.07	.07	.08	.08	.09	.09	.09	.074	.072
CFI	.95	.79	.95	.96	.96	.96	.96	.95	.96	.95	.94	.94	.944	.943
NNFI	.94	.76	.94	.96	.96	.95	.96	.95	.95	.94	.94	.93	.936	.940
PNFI													810	.885
χ^2 value														
	745.44	1660.09	932.08	960.79	1034.52	1211.88	1018.72	1184.30	1035.00	1207.36	1084.23	820.758	12990.22	13565.92
df														
	220	220	220	220	220	220	220	220	220	220	220	220	2640	2893

Note: χ^2 difference test was significant (p<.05) between the "All Free Loadings" model and the "Loadings Constrained Equal Across Groups" model. (χ^2 critical with 253 *df* = 291.10; χ^2 observed with 253 *df*=575.70; χ^2 observed > χ^2); RMSEA= root mean squared error of approximation; CFI= comparative fit index; NNFI= non-normed fit index ; PNFI= parsimony normed fit index; x suggests that there is no single set of loadings corresponding to the "All Free Loadings" model.

Table 10. Two-way, mixed effect, absolute agreement, single measure Intraclass Correlations (ICC) between Child Self-Report and Parent Proxy-Report for PedsQL™ 4.0 Generic Core Scales and Total Scale Scores by Age

Age	N	Total Score	Physical Health	Psychosocial Health	Emotional Functioning	Social Functioning	School Functioning
5 yrs	748	0.51*	0.36*	0.56*	0.62*	0.43*	0.45*
6 yrs	912	0.44*	0.28*	0.50*	0.60*	0.38*	0.44*
7 yrs	873	0.46*	0.31*	0.53*	0.59*	0.41*	0.42*
8 yrs	863	0.57*	0.46*	0.60*	0.63*	0.50*	0.48*
9 yrs	830	0.60*	0.48*	0.63*	0.63*	0.53*	0.57*
10 yrs	826	0.70*	0.60*	0.71*	0.69*	0.63*	0.63*
11 yrs	671	0.62*	0.52*	0.64*	0.63*	0.55*	0.59*
12 yrs	660	0.67*	0.57*	0.69*	0.67*	0.59*	0.64*
13 yrs	605	0.67*	0.57*	0.69*	0.68*	0.57*	0.60*
14 yrs	554	0.70*	0.63*	0.70*	0.67*	0.62*	0.64*
15 yrs	543	0.70*	0.60*	0.71*	0.69*	0.60*	0.66*
16 yrs	321	0.69*	0.62*	0.67*	0.64*	0.53*	0.66*

N= 8,406. Note: *p<.001, **p<.05. Intraclass Correlation Coefficients (ICC) are designated as $\leq$ 0.40 poor to fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 good agreement, and 0.81-1.00 excellent agreement.

Table 11. Mean Differences Between Child Self-Report and Parent Proxy-Report on the PedsQL™ 4.0 Generic Core Scales Total Score by Age

Age	N	Child Mean Score	Parent Mean Score	Mean Difference	Effect Size
5 yrs	748	81.34	77.81	3.53***	0.29
6 yrs	912	80.85	77.98	2.87***	0.23
7 yrs	873	79.05	76.37	2.68***	0.21
8 yrs	863	80.91	78.13	2.78***	0.21
9 yrs	830	80.92	76.32	4.60***	0.33
10 yrs	825	81.54	77.97	3.57***	0.28
11 yrs	671	82.26	77.81	4.45***	0.36
12 yrs	660	81.07	76.42	4.65***	0.36
13 yrs	604	80.49	76.55	3.94***	0.30
14 yrs	554	80.68	77.03	3.65***	0.30
15 yrs	543	80.76	77.03	3.73***	0.29
16 yrs	321	80.04	75.45	4.59***	0.40

*** $p < .001$; Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 12. Mean Differences Between Child Self-Report and Parent Proxy-Report on the PedsQL™ 4.0 Generic Core Scales Physical Health Score by Age

Age	N	Child Mean Score	Parent Mean Score	Mean Difference	Effect Size
5 yrs	748	84.27	77.70	6.57***	0.49
6 yrs	912	85.01	78.53	6.48***	0.49
7 yrs	871	83.56	77.25	6.31***	0.46
8 yrs	861	84.01	80.43	3.58***	0.26
9 yrs	828	84.87	78.60	6.27***	0.42
10 yrs	820	84.42	80.78	3.64***	0.27
11 yrs	671	84.18	79.67	4.51***	0.34
12 yrs	657	83.39	77.62	5.77***	0.41
13 yrs	604	82.75	78.03	4.72***	0.35
14 yrs	554	82.71	78.36	4.35***	0.33
15 yrs	542	83.92	79.15	4.77***	0.35
16 yrs	320	82.15	77.16	4.99***	0.39

*** $p < .001$; Effect sizes are designated as small (.20), medium (.50), and large (.80).

Table 13. Mean Differences Between Child Self-Report and Parent Proxy-Report on the PedsQL™ 4.0 Generic Core Scales Psychosocial Health Score by Age

Age	N	Child Mean Score	Parent Mean Score	Mean Difference	Effect Size
5 yrs	745	79.72	77.83	1.89***	0.14
6 yrs	913	78.57	77.66	0.91	0.06
7 yrs	871	76.67	75.94	0.73	0.05
8 yrs	863	79.28	76.95	2.33***	0.16
9 yrs	827	78.91	75.12	3.79***	0.25
10 yrs	823	80.03	76.52	3.51***	0.25
11 yrs	670	81.25	76.80	4.45***	0.33
12 yrs	661	79.69	75.67	4.02***	0.28
13 yrs	604	79.27	75.76	3.51***	0.24
14 yrs	556	79.63	76.25	3.38***	0.25
15 yrs	543	79.09	75.88	3.21***	0.23
16 yrs	319	78.85	74.55	4.30***	0.33

*** $p < .001$; Effect sizes are designated as small (.20), medium (.50), and large (.80).

VITA

Name: Christine A. Limbers

Address: 5000 Maidstone Ct., College Station, TX 77845

Email Address: Climbers@tamu.edu

Education: B.A., Communication, Santa Clara University, 2003