



Published in final edited form as:

J Happiness Stud. 2018 March ; 19(3): 699–718. doi:10.1007/s10902-016-9843-9.

Development and Evaluation of the PROMIS® Pediatric Positive Affect Item Bank, Child-Report and Parent-Proxy Editions

Christopher B. Forrest, MD, PhD^[1], Ulrike Ravens-Sieberer, PhD, MPH^[2], Janine Devine, PhD^[2], Brandon D. Becker, MPH^[1], Rachel Teneralli, MS^[1], JeanHee Moon, PhD^[1], Adam Carle, PhD^[3], Carole A. Tucker^[4], and Katherine B. Bevans, PhD^[1]

^[1]Department of Pediatrics, School of Medicine, University of Pennsylvania, Philadelphia, PA

^[2]Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, University Medical Center Hamburg-Eppendorf, Germany

^[3]Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

^[4]Department of Physical Therapy, College of Public Health, Temple University, Philadelphia, PA

Abstract

Objective—The purpose of this study is to describe the psychometric evaluation and item response theory calibration of the PROMIS Pediatric Positive Affect item bank, child-report and parent-proxy editions.

Methods—The initial item pool comprising 53 items, previously developed using qualitative methods, was administered to 1,874 children 8–17 years old and 909 parents of children 5–17 years old. Analyses included descriptive statistics, reliability, factor analysis, differential item functioning, and construct validity. A total of 14 items were deleted, because of poor psychometric performance, and an 8-item short form constructed from the remaining 39 items was administered to a national sample of 1,004 children 8–17 years old, and 1,306 parents of children 5–17 years old. The combined sample was used in item response theory (IRT) calibration analyses.

Results—The final item bank appeared unidimensional, the items appeared locally independent, and the items were free from differential item functioning. The scales showed excellent reliability and convergent and discriminant validity. Positive affect decreased with children's age and was lower for those with a special health care need. After IRT calibration, we found that 4 and 8 item short forms had a high degree of precision (reliability) across a wide range of the latent trait (>4 SD units).

Conclusion—The PROMIS Pediatric Positive Affect item bank and its short forms provide an efficient, precise, and valid assessment of positive affect in children and youth.

Keywords

positive affect; experienced well-being; subjective well-being; PROMIS; child; Item Response Theory

Introduction

Positive affect comprises the emotional experiences resulting from pleasant engagement with the environment (Clark et al. 1989), inclusive of the activated states of happiness, joy, excitement, and enthusiasm and the low arousal states of contentment, peacefulness, satisfaction, and calm (Pressman and Cohen 2005). It results from social interactions, activity, and goal attainment (Forbes and Dahl 2005). The processes of seeking and obtaining rewards and overcoming developmentally appropriate challenges are inherently pleasurable (Campos et al. 1994). Among children, positive affect facilitates prosocial behavior, helping individuals initiate and maintain friendships and enhance their relationships with family members (Lennon and Eisenberg 1987).

Positive affect is part of the multidimensional concept of subjective well-being, which includes experiential (emotional states), evaluative (how satisfying life is), and eudaimonic (sense of meaning and purpose) components (National Research Council 2013). Experienced well-being refers to emotions that result from individuals' everyday interactions in their environments; it subsumes physical and emotional distress experienced as pain, fatigue, sadness, anxiety, and anger as well as the positive feelings of happiness, joy, and contentment (i.e., positive affect). Much research on experienced well-being has focused on the health effects of negative affect, even though positive affect is an essential element of the lived experience (Fredrickson 1998).

To address the need for efficient (short), precise (reliable across a wide range of the latent trait), and valid measures of self-reported physical, mental, and social health that can be used in clinical research and practice, the National Institutes of Health launched in 2004 a program of research called the Patient Reported Outcome Measurement Information System (PROMIS[®]) (Cella et al. 2007). PROMIS comprises a cooperative group of research sites and centers as well as a unique mixed-methods approach for developing instruments that assess the lived experiences of physical, mental, and social health (Cella et al. 2007; Forrest et al. 2012). PROMIS has produced numerous measures that are applicable across the life course and are in widespread use internationally (see www.healthmeasures.net for more information).

The PROMIS suite of pediatric measures includes Depressive Symptoms and Anxiety (Irwin et al. 2010), Anger (Irwin et al. 2012), and Psychological Stress Experiences (Bevans et al. 2013). However, it does not include assessments of positive affect, which precludes uses of these tools for comprehensive assessments of experienced well-being, nor does it include measures of evaluative or eudaimonic well-being. To fill these gaps in the PROMIS suite of pediatric measures, we embarked on a program of research to develop measures of pediatric positive psychological functioning using the PROMIS methodology for item bank development (Forrest et al. 2012).

Previously, we reported the qualitative development of pediatric subjective well-being measures, including items pools for life satisfaction, meaning and purpose, and positive affect (Ravens-Sieberer et al. 2014). In this manuscript, we describe the psychometric evaluation and item bank calibration using item response theory for the PROMIS Pediatric

Positive Affect item bank. Data were collected during two cross-sectional surveys that included 2,878 children 8–17 years-old and 2,215 parents of children 5–17 years-old who provided data for the psychometric and statistical analyses.

Methods

Two studies were done to evaluate the psychometric properties of the item bank. Study 1 was conducted April 2011 to March 2012. Its purpose was to evaluate scale dimensionality, reliability, and concurrent validity, and to identify items with poor item-level reliability. Study 2 was done June 2014 to September 2014. Its purpose was to obtain a national sample of participants for further psychometric evaluation and item bank calibration. The Institutional Review Board of the Children’s Hospital of Philadelphia approved both study protocols (IRB #10-007684 and IRB #13-010404). Informed consent was obtained from parents, and assent was obtained from children.

Study 1 Data Collection

Study 1 enrolled 1,874 children aged 8–17 years and 909 parents of children aged 5–17 years. About half (53% of children [n=992] and 51% [n = 487] of parents) were recruited from a convenience Internet panel. Others were recruited from schools (40% of children [n=760], 33% of parents [n=302]) or the Children’s Hospital of Philadelphia clinics (7% of children [n = 122], 16% of parents [n=143]). The child samples sizes from the clinics were: primary care (n=63), emergency department (n=26), gastroenterology (n=15), rheumatology (n=9), healthy weight (n=6), and the dialysis center (n=3). Study 1 participants responded to the full positive affect item pool.

A sample of children (n=114) and parents (n=52) from the convenience Internet panel completed the items twice, approximately 3 weeks apart (mean and median=23 days, SD=3 days, range=16–32 days) to evaluate the instruments’ test-retest reliability.

The Study 1 Internet sample was recruited through the Op4G.com online research panel, a private research community with approximately 250,000 members who participate in a range of research activities on their home computers. Adult participants known to have children aged 5–17 years were notified by email of their eligibility. Parents who provided informed consent were emailed a link to the online parent questionnaire. After completing their measures, parents asked their child to complete the survey. The online research panel firm did not retain the number of parents invited to participate, which precluded calculation of a participation rate. Parents with missing data were sent automated reminder emails every 3 days until they/their child completed the questionnaire(s) or a maximum of 3 email reminders were sent.

School-based data collection took place in three school districts located in New Hampshire, Vermont, and Texas. Fifteen schools were selected to diversify the sample with respect to race, ethnicity, socio-economic status, and geographic location. All students in grades 3–12 (approximate age range 8–17 years-old), for whom parental informed consent could be obtained in English, except those in self-contained special education classrooms, were invited to participate. Parental consent forms were sent home at the beginning of the school

year. Parents returned consent forms to the school. Children in 5th–12th grade (approximate age range 10–17 years-old) with parental consent completed self-administered paper-and-pencil questionnaires, and staff members from the our study team read questions to groups of children in 3rd–4th grade (approximate age range 8–9 years-old), either in their classrooms or in cafeterias. After completing the survey, children were given a parent questionnaire in an envelope and instructed to take it home. Parents returned the questionnaire by mail to the study team in postage paid envelopes.

For clinic-based survey administration, parents and children who met study inclusion criteria (child age and capacity to self-report) were approached by study staff in the reception areas and were provided information about the study. If interested, parents provided consent, and then responded to the survey, while they waited in the reception area, at the same time their children completed the survey on tablet computers.

Study 2 Data Collection

We recruited a sample of parents of children aged 5–17 years old from GFK Knowledge Panel, an existing dual-frame (random-digit dial and address-based) online probability panel (DiSogra et al. 2010; Dennis 2010). Weights were used to render the online panel assembled for this study a national sample of the US non-institutionalized population. The initial weights adjusted for oversampling of individuals living in minority communities and Spanish-language dominant Hispanic areas and other sources of error (e.g., non-response). The weights were then iteratively adjusted (raked) until the weighted sample's distributions of gender, age, race, ethnicity, education, U.S. Census region, metropolitan area, household internet access, and language (English/Spanish) matched those in the most recent Current Population Survey (Lohr 2009). The weights for our study's participants went through the raking process again to render the weighted study sample representative of the non-institutionalized US population. Data integrity can be a concern with Internet panels (Hays et al. 2015), because panelists may rush through a questionnaire so they can receive whatever incentive is offered. To address this potential data quality issue, we presented 1 item per screen, rather than a cluster of items per screen, to focus respondent's attention on each question.

Study 2 participants completed responded to an 8-item short form of the positive affect instrument. Parents who provided informed consent were emailed a link to the online parent questionnaire. After completing their measures, parents of children aged 8–17 years were instructed to ask their child to complete a questionnaire. The child survey was administered as an audio-assisted computerized questionnaire, although children could stop the audio by advancing forward after reading the question and recording their answer. Data collection continued until age and gender quotas were met for each form.

Item Pool

The previously developed item pool evaluated in this study contained 53 items measuring positive emotions within the content areas of contentment, pride, love, happiness, and energy (Ravens-Sieberer et al. 2014). Items have a 7-day recall period and include five response

categories (1: never, 2: rarely, 3: sometimes, 4: often, 5: always). Parent-proxy versions replace the pronoun “I” with “my child.”

Data obtained from children’s reports of their gender, age, race (White, African American, Asian, Other), and ethnicity (Hispanic, non-Hispanic) were used in the analysis of child-report instruments. Likewise, parent reports of child gender, age, race, and ethnicity were used in analyses of the parent-report instruments. Analyses for Study 2 only used parents’ responses for presence of a special health care need (Bethell et al. 2002a; Bethell et al. 2002b), annual household income, and parental educational attainment.

Classical Test Theory Analyses

All classical test theory analyses were done use data from Study 1. Each item’s mean, standard deviation, skewness, and percentage with scores at the ceiling (score of 5) or floor (score of 1) were computed. At the scale level, we examined the range of the IRT-based Positive Affect T-score (see below for scoring details) and the percentage of individuals at the floor and ceiling of these scores. Reliability was evaluated with an IRT-based estimate of Cronbach's alpha called marginal reliability (Green et al. 1984), item-total correlations, and test-retest correlations (intraclass correlation coefficient). Multivariable regression analyses that controlled for the covariates described above were done to explore the hypotheses that positive affect decreases with age and presence of a chronic condition (i.e., special health care need).

Convergent and discriminant validity were assessed by examining the correlation (Pearson’s r) between an IRT-based Positive Affect T-score (see below for scoring details) and scores on other measures. Convergent validity (expected positive correlations with other measures of positive psychological functioning) was examined with the NIH Toolbox Positive Affect Fixed Form v2.0 (nihtoolbox.org) NIH Toolbox General Life Satisfaction Fixed Form v2.0 (nihtoolbox.org), PROMIS Pediatric Short Form v1.0 - Life Satisfaction 8a (Ravens-Sieberer et al. 2014), PROMIS Pediatric Short Form v1.0 - Meaning and Purpose 8a (Ravens-Sieberer et al. 2014), and the KIDSCREEN Psychological Well-Being scales (Ravens-Sieberer et al. 2005; Ravens-Sieberer et al. 2008). Discriminant validity (expected negative correlations with physical and emotional distress) was examined using the PROMIS Pediatric Short Form 1.0 - Anger 6a (Irwin et al. 2012), PROMIS Pediatric Short Form 1.0 - Anxiety 8a (Study 1) and Short Form 1.1 –Anxiety 8b (Study 2) (Irwin et al. 2010), PROMIS Pediatric Short Form 1.0 - Depressive Symptoms 8a (Irwin et al. 2010), PROMIS Pediatric Short Form 1.0 - Pain Interference 8a (Varni et al. 2010) and PROMIS Pediatric Short Form 1.0 - Physical Stress Experiences 8a (Bevans et al. 2013).

Testing Assumptions of IRT Analysis

Unidimensionality, local independence, and monotonicity are prerequisites for IRT modeling (Edelen and Reeve 2007; Embretson and Reise 2000). Using the full sample from Study 1, unidimensionality was examined using Exploratory and Confirmatory Factor Analyses (EFA and CFA) with the weighted least squares means and variance adjusted estimator and an oblique rotation technique implemented using Mplus 6.1. Unidimensionality was supported

in EFA if a single factor explained a large share of the variance and the ratio of the 1st and 2nd eigenvalues was >4 . Three indices provided information about CFA model fit: the Comparative Fit Index, Tucker–Lewis fit index, and Root Mean Square Error of Approximation (Hu and Bentler 1999). A criterion of >0.60 for CFA factor loadings was used; those with lower loadings were considered for removal. Items were considered locally dependent if their residual correlations were >0.20 in the single factor CFA model (Edelen and Reeve 2007), and when present, one of the two correlated items was removed. Graphs of item mean scores conditional on the total test scale score minus the item score were examined to confirm item monotonicity (i.e., the probability of item endorsement should increase as the measured trait increases). Non-monotonic items were removed.

Differential Item Functioning

To identify item bias by age, sex, race, ethnicity, or study sample, differential item functioning analyses were done using the Lordif package in R (Choi et al. 2011). Lordif regresses an item's ordered responses on an IRT-derived scale score and the indicator variables, such as age and sex, to test for both uniform and non-uniform differential item functioning (DIF). Items that showed a 1% change in the McFadden pseudo R^2 measure were considered to demonstrate DIF (Crane et al. 2007) and were removed from the item pool.

Item Bank Calibration

The final item pool, selected based on results from descriptive analyses, reliability, monotonicity, differential item functioning, factor analyses, and local dependence assessments, was calibrated using Samejima's Graded Response Model (Samejima 1997). In IRT, calibration refers to estimating the IRT discrimination and threshold parameters for each item using a sample's item responses. The discrimination statistic (also referred to as slope and designated by a) measures the capacity of the item to differentiate respondents by their level of the latent trait (i.e., positive affect). The IRT model also produces threshold parameters (referred to as item difficulty, and designated as b), which measure the "difficulty" of endorsing the item and where along the latent trait continuum the threshold between two response options functions best. Thus, for a measure with five response categories, the IRT model produces four item threshold statistics.

Ideally, one calibrates items using a representative sample from the population of interest. Although we had a sample of the general US pediatric population (Study 2), those children answered eight short form items only rather than all items in the item pool. Thus, we could not calibrate all of the items in the bank using the Study 2 sample. Although the children in Study 1 answered all of the items in the bank, it was not a representative sample, and by design included a much larger proportion of children with health conditions than the general population.

To address these challenges, we used a multiple group IRT approach (Millsap and Yun-Tein 2004). This allowed us to use all of the available data from both samples to conduct our calibration analyses. We statistically identified the latent variable's metric by constraining

the Study 2 group IRT mean and variance to 0 and 1 respectively. We freely estimated the Study 1 group's IRT mean and variance, so this group's mean and variance did not set the latent variable metric. To statistically identify the model across the groups, we constrained each short form item's parameters to equality across groups. We treated the items not presented in Study 2 as missing for the Study 2 group and did not introduce cross-group constraints for these items. As a result, the values for these parameters were estimated in the same metric as the short form item parameters but using only Study 1 sample data. Importantly, this metric is interpretable in terms of the general US pediatric population. Setting the IRT mean and variance to 0 and 1 respectively sets the metric for the latent variable and parameters (Millsap and Yun-Tein 2004). Because the Study 2 sample represents the general US population and because we identified the latent variable's metric using the Study 2 group only, the IRT parameters and scores estimated from these parameters can be interpreted relative to the general US pediatric population. We implemented these analyses in Mplus 7.2 using maximum likelihood estimation with a logit link.

Development of Fixed Length Forms

To create 4- (SF4) and 8-item (SF8) fixed length forms, items were selected from the item bank to provide measurement precision across a wide range of children's positive affect experiences. Full bank and short form test information functions were plotted to assess each instrument version's precision across all levels of the positive affect continuum. Most PROMIS item banks provide 4- and 8-item short forms to give users a couple of fixed length form options (see www.healthmeasures.net).

Scoring

After finalizing item parameters, we used Firestar v1.2.2, an R-based simulation software (Choi 2009), and estimated full bank, SF8, and SF4 theta scores using Bayesian Expected A Posteriori (EAP) estimation (Bock and Aitkin 1981). EAP scoring uses an individual's pattern of responses and the model's parameters to estimate an individual's theta score. The theta scores were linearly transformed to T-scores by multiplying theta scores by 10 and adding 50. Thus, a score of 50 represents the average positive affect level for children in the national sample used for calibration and centering of scores.

Results

The descriptive characteristics of the child and parent samples in the two studies are shown in Table 1. All analyses were replicated for the parent-proxy banks, and are shown in the Appendix; results were consistent with those found for the child self-report edition.

Item Deletions

The item pool evaluated during Study 1 included 53 items. Fourteen of these were deleted for local dependence (residual correlation >0.20) with one or more items, non-monotonicity, or low factor loadings (<0.40) in CFA. Items with local dependence included: *very strong*, *very healthy*, *very active*, *very enthusiastic*, *very positive*, *very relaxed*, *very calm*, *full of*

energy, felt loved by parents, and friendly. Note that 7 of the 10 items with local dependence used the adverb *very*, and the item was locally dependent with a similarly phrased item that excluded “*very*.” The three items excluded because of low factor loadings were *alert, amused, and enjoyed the things I did.* Finally, one item, *loved*, was removed because of non-monotonic increases among the four item thresholds.

Item-level Classical Test Analyses

Item-level means ranged from a low of 3.52 (*wide awake*) to a high of 4.16 (*healthy*)—Table 2. Floor effects (% endorsing never) were minimal, while ceiling effects (% endorsing always) were more common, with 31% of children endorsing always on average across all items. The range of item-total correlations was 0.61 to 0.86, and the range of item-level test-retest correlations was comparable, 0.57 to 0.86.

Dimensionality

The ratio of the 1st and 2nd eigenvalues from EFA using all 39 items in the item bank was 21, and the first extracted factor accounted for 66% of the variance. The model fit statistics from a confirmatory factor analyses also supported unidimensionality (CFI 0.94, TLI 0.94, and RMSEA 0.09). The factor loadings ranged from 0.64 to 0.91 (Table 2). These findings indicated that the 39 items appeared to converge on a single unidimensional factor.

Item Response Theory Analyses

The IRT item discrimination and threshold parameters from the graded response model for the 39 Positive Affect items are shown in Table 3. The item with the largest slope, providing the greatest level of information about Positive Affect, was *joyful*. The range of threshold parameters was a low of -3.30 (*healthy*) to a high of 0.75 (*refreshed, blissful, merry*).

Short Forms

We selected items to constitute 8-item and 4-item short forms (Table 3). The SF4 is a subset of the SF8. Items with high levels of item discrimination were preferred, because this statistic reflects an item’s ability to differentiate among individuals at different ranges of Positive Affect, and we chose items that provided discrimination across as much of the latent trait as possible. The test information function plot (Figure) illustrates that the short forms provide acceptable levels of precision between T-scores of 20 to 65.

Scale-Level Analyses

The range in T-scores for the full item bank spanned 6 standard deviations (12–72), while the range for SF8 was 4.5 standard deviations (19–66) and 4.0 standard deviations (22–63) for SF4. Fewer than 1% of individuals had a floor effect, although ceiling effects (highest levels of positive affect) were seen among 12% and 20% of the sample for SF8 and SF4, respectively. Marginal reliability exceeded 0.90 and test-retest reliability exceeded 0.75 for the item bank and both short forms. The item bank and short forms showed excellent convergent validity with measures of positive affect, life satisfaction, meaning and purpose,

and psychological well-being (Table 4). Regarding discriminant validity, the largest negative correlations were observed with anger and depressive symptoms, were lowest for pain, and intermediate for anxiety and psychological stress. Children with the highest quartile of depressive symptoms versus the lowest had markedly less positive affect (T-score of 39 versus 50). Similarly, children with the highest quartile of anxiety had less positive affect than others in the lowest quartile (T-score of 42 versus 50).

Positive Affect T-scores decreased with age: an average of 52 for 8–9 year-olds, 51 for 10–11 year-olds, 49 for 12–13 year-olds, 48 for 14–15 year-olds, and 47 for 16–17 year-olds. In multivariable regression that regressed Positive Affect T-scores on the covariates in Table 1 from Study 2, we found significant effects of age (-0.4 , 95% CI -0.6 to -0.2) and special health care need (-3.4 , 95% CI -4.7 to -2.2), but no significant differences by gender, race, ethnicity, household income, parental educational attainment, parental relationship to the child, or parental age.

Differential Item Functioning

None of the 39 items in the item bank showed DIF by age, sex, race, or ethnicity. In addition, we did not detect DIF between the child samples in Studies 1 and 2 for the items in the Positive Affect SF8.

Discussion

Development of the PROMIS Pediatric Positive Affect item bank involved formative, qualitative research that included child, parent, and content expert semi-structured interviews, a systematic literature review, readability analysis, translatability review, and cognitive interviews (Ravens-Sieberer et al. 2014). The initial item pool included 53-items representing 6 content categories – low activation states of contentment, love, and pride, and high activation states of happiness, excitement, and energy. A total of 14 items were removed from the item pool because they had low factor loadings, local dependence with another item, or non-monotonicity. The resulting item pool demonstrated unidimensionality and local independence, important assumptions of IRT modeling. The item pool was calibrated using the IRT graded response model (Samejima 1997). Based on results from these analyses, the SF4 and SF8 fixed length questionnaires were developed. The final PROMIS Pediatric Positive Affect item bank and short forms assess a child's experiences of well-being measured as momentary positive or rewarding affective experiences using a 7-day reporting period. A Parent-Proxy edition with comparable psychometric properties was also developed, and it can be used for children 5–17 years old, whereas the child-report edition is for children 8–17 years-old.

The PROMIS Pediatric Positive Affect item bank measures the positive feelings of experienced well-being. Additional PROMIS Pediatric measures can be used to evaluate physical and psychological distress, such as the feelings of pain, fatigue, and negative affect. Similar to the PROMIS Pediatric pain interference, fatigue, anger, anxiety, depressive symptoms, and psychological stress experiences item banks, the positive affect item bank uses a 7-day recall period. Thus, a comprehensive conceptualization of experienced well-

being in childhood can be examined using a profile of PROMIS Pediatric measures. Very few studies have examined the independent and joint contributions of negative affect and positive affect (Pressman and Cohen 2005), and the PROMIS suite of measures will prove useful for this type of research.

The item bank and short form scales have excellent reliability, as assessed by marginal reliability and test-retest reliability. The Positive Affect scales provide adequate levels of precision (marginal reliability >0.90) from T-scores as low as 20 to as high as 65 with precision decreasing at the scales' edges. Like the SF8, the SF4 measures positive affect with a high degree of precision when scores range from about 25 to 60; however, the SF8 form provides slightly more precision at the high end of the distribution (>60). Unless an application calls for differentiating a sample at the highest levels of positive affect, the SF4 might be preferred because it has half as many items.

The Positive Affect measures showed convergent validity with other, legacy measures of positive affect and positive psychological functioning, and discriminant validity with measures of emotional distress, psychological stress, and pain. The tripartite model of anxiety and depression, proposed by Clark and Watson for adults (Clark and Watson 1991), was not supported by this study. The model suggests that positive affect, among adults, is low among patients with depression but not anxiety. In this study, we found that children with either high anxiety or depressive symptoms also had low positive affect, which is consistent with other pediatric research (Jacques and Mash 2004).

The lack of longitudinal data is an important limitation of this study. We were not able to examine the predictive validity of positive affect, nor were we able to assess the intra-individual trajectories of negative and positive affect to better understand how their associations change during childhood and adolescence and into adulthood.

The newly developed PROMIS Pediatric Positive Affect measure has similarities to the Positive and Negative Affect Schedule for Children (PANAS-C), which includes a positive affect subscale that has been reduced to 5 items, including *joyful*, *cheerful*, *happy*, *lively*, and *proud* items (Ebesutani et al. 2012). All but *lively*, which we found to be poorly understood by children (Ravens-Sieberer et al. 2014), are included in the PROMIS Pediatric Positive Affect item bank, and the SF4 includes *joyful*, *cheerful*, and *happy*. In graded response model item response theory analyses, the item *joyful* provided the highest level of item discrimination for both the PANAS-C and PROMIS Pediatric Positive Affect measure. Similar to the PROMIS measure, the PANAS-C positive affect scale provides adequate measurement precision up to about 1 standard deviation unit above the sample mean (Ebesutani et al. 2012). The results of this study and work on the PANAS-C indicate that a small batch of positive emotions can provide useful information about positive affect across a wide range of the latent trait among children.

Much of the data for this study was collected from two Internet panels. Study 1 used a non-probability convenience panel and Study 2 used a probability-based panel. Advantages of Internet panels include the efficiency with which large amounts of data can be collected, the

accessibility of diverse populations, and standardization of the data collection process (Hays et al. 2015).

In summary, the PROMIS Pediatric Positive Affect item bank provides efficient, precise, and valid short forms that can be used to assess a child's level of happiness and positive affect. The scales have excellent precision across a wide range of the latent trait, and evidence for their concurrent validity. The child-report edition can be used for children 8–17 years-old, and a Parent-Proxy edition is available for children ages 5–7 years-old. Assessment of experienced well-being can be done by combining the PROMIS Pediatric Positive Affect measure with other PROMIS measures of emotional distress, pain, or stress to create a profile of suffering and happiness. Reducing suffering while promoting happiness is a common objective of pediatric care; thus, the measure should find widespread use in pediatric clinical applications.

Acknowledgments

The Patient-Reported Outcomes Measurement Information System (PROMIS) is an NIH Roadmap initiative to develop a computerized system measuring PROs in respondents with a wide range of chronic diseases and demographic characteristics. PROMIS II was funded by cooperative agreements with a Statistical Center (Northwestern University, PI: David Cella, PhD, 1U54AR057951), a Technology Center (Northwestern University, PI: Richard C. Gershon, PhD, 1U54AR057943), a Network Center (American Institutes for Research, PI: Susan (San) D. Keller, PhD, 1U54AR057926) and thirteen Primary Research Sites which may include more than one institution (State University of New York, Stony Brook, PIs: Joan E. Broderick, PhD and Arthur A. Stone, PhD, 1U01AR057948; University of Washington, Seattle, PIs: Heidi M. Crane, MD, MPH, Paul K. Crane, MD, MPH, and Donald L. Patrick, PhD, 1U01AR057954; University of Washington, Seattle, PIs: Dagmar Amtmann, PhD and Karon Cook, PhD, 1U01AR052171; University of North Carolina, Chapel Hill, PI: Darren A. DeWalt, MD, MPH, 2U01AR052181; Children's Hospital of Philadelphia, PI: Christopher B. Forrest, MD, PhD, 1U01AR057956; Stanford University, PI: James F. Fries, MD, 2U01AR052158; Boston University, PIs: Stephen M. Haley, PhD and David Scott Tulskey, PhD (University of Michigan, Ann Arbor), 1U01AR057929; University of California, Los Angeles, PIs: Dinesh Khanna, MD and Brennan Spiegel, MD, MSHS, 1U01AR057936; University of Pittsburgh, PI: Paul A. Pilkonis, PhD, 2U01AR052155; Georgetown University, PIs: Carol M. Moinpour, PhD (Fred Hutchinson Cancer Research Center, Seattle) and Arnold L. Potosky, PhD, U01AR057971; Children's Hospital Medical Center, Cincinnati, PI: Esi M. Morgan DeWitt, MD, MSCE, 1U01AR057940; University of Maryland, Baltimore, PI: Lisa M. Shulman, MD, 1U01AR057967; and Duke University, PI: Kevin P. Weinfurt, PhD, 2U01AR052186). NIH Science Officers on this project have included Deborah Ader, PhD, Vanessa Ameen, MD, Susan Czajkowski, PhD, Basil Eldadah, MD, PhD, Lawrence Fine, MD, DrPH, Lawrence Fox, MD, PhD, Lynne Haverkos, MD, MPH, Thomas Hilton, PhD, Laura Lee Johnson, PhD, Michael Kozak, PhD, Peter Lyster, PhD, Donald Mattison, MD, Claudia Moy, PhD, Louis Quatrano, PhD, Bryce Reeve, PhD, William Riley, PhD, Ashley Wilder Smith, PhD, MPH, Susana Serrate-Sztejn, MD, Ellen Werner, PhD and James Witter, MD, PhD.

References

- Bethell CD, Read D, Neff J, Blumberg SJ, Stein REK, Sharp V, et al. Comparison of the children with special health care needs screener to the questionnaire for identifying children with chronic conditions--revised. *Academic pediatrics*. 2002a; 2(1):49–57.
- Bethell CD, Read D, Stein REK, Blumberg SJ, Wells N, Newacheck PW. Identifying children with special health care needs: development and evaluation of a short screening instrument. *Academic pediatrics*. 2002b; 2(1):38–48.
- Bevans KB, Gardner W, Pajer K, Riley AW, Forrest CB. Qualitative development of the PROMIS[®] pediatric stress response item banks. *Journal of pediatric psychology*. 2013; 38(2):173–191. [PubMed: 23124904]
- Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*. 1981; 46:443–459.
- Campos JJ, Mumme DL, Kermoian R, Campos R. A functionalist perspective on the nature of emotion. *Monographs of the Society for Research in Child Development*. 1994; 59(2–3):284–303. [PubMed: 7984165]

- Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care*. 2007; 45(5 Suppl 1):S3–S11.
- Choi SW. Firestar: Computerized Adaptive Testing Simulation Program for Polytomous Item Response Theory Models. *Applied Psychological Measurement*. 2009; 33(8):644–645.
- Choi SW, Gibbons LE, Crane PK. Iordif: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations. *Journal of statistical software*. 2011; 39(8):1–30.
- Clark LA, Watson D. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*. 1991; 100:316–336. [PubMed: 1918611]
- Clark LA, Watson D, Leeka J. Diurnal variation in the positive affects. *Motivation and Emotion*. 1989; 13:205–234.
- Craig BM, Hays RD, Pickard AS, Cella D, Revicki DA, Reeve BB. Comparison of US panel vendors for online surveys. *Journal of Medical Internet Research*. 2013; 15:e260. [PubMed: 24292159]
- Crane PK, Gibbons LE, Narasimhalu K, Lai JS, Cella D. Rapid detection of differential item functioning in assessments of health-related quality of life: The Functional Assessment of Cancer Therapy. *Qual Life Res*. 2007; 16(1):101–114. [PubMed: 17111233]
- Dennis, JM. [Accessed June 6 2016] KnowledgePanel®: Processes & Procedures Contributing to Sample Representativeness & Tests for Self-Selection Bias. 2010. <http://www.knowledgenetworks.com/ganp/docs/KnowledgePanelR-Statistical-Methods-Note.pdf>
- DiSogra C, Dennis JM, Fahimi M. On the quality of ancillary data available for address-based sampling. *Proceedings of the American Statistical Association, Section on Survey Research Methods*. 2010:4174–4183.
- Ebesutani C, Regan J, Smith A, Reise S, Higa-McMillan C, Chorpita BF. The 10-Item Positive and Negative Affect Schedule for Children, Child and Parent Shortened Versions: Application of Item Response Theory for More Efficient Assessment. *J Psychopathol Behav Assess*. 2012; 34(2):191–203.
- Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Qual Life Res*. 2007; 16(Suppl 1):5–18. [PubMed: 17375372]
- Embretson, SE., Reise, SP. *Item Response Theory for Psychologists*. London: Lawrence Erlbaum Associates; 2000.
- Forbes EE, Dahl RE. Neural systems of positive affect: relevance to understanding child and adolescent depression? *Dev Psychopathol*. 2005; 17(3):827–850. [PubMed: 16262994]
- Forrest CB, Bevans KB, Tucker C, Riley AW, Ravens-Sieberer U, Gardner, et al. Commentary: the patient-reported outcome measurement information system (PROMIS®) for children and youth: application to pediatric psychology. *J Pediatr Psychol*. 2012; 37(6):614–621. [PubMed: 22362923]
- Fredrickson BL. What Good Are Positive Emotions? *Rev Gen Psychol*. 1998; 2(3):300–319. [PubMed: 21850154]
- Green BF, Bock D, Humphres RL, Linn MD. Technical Guidelines for Assessing Computerized Adaptive Tests. *Journal of Educational Measurement*. 1984; 21(4):347–360.
- Hays RD, Liu H, Kapteyn A. Use of Internet panels to conduct surveys. *Behav Res*. 2015; 47(3):685–690.
- Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- Irwin DE, Stucky B, Langer MM, Thissen D, Dewitt EM, Lai JS, et al. An item response analysis of the pediatric PROMIS anxiety and depressive symptoms scales. *Qual Life Res*. 2010; 19(4):595–607. [PubMed: 20213516]
- Irwin DE, Stucky BD, Langer MM, Thissen D, Dewitt EM, Lai JS, et al. PROMIS Pediatric Anger Scale: an item response theory analysis. *Qual Life Res*. 2012; 21(4)
- Jacques HAK, Mash EJ. A test of the tripartite model of anxiety and depression in elementary and high school boys and girls. *J Abnorm Child Psychol*. 2004; 32(1):13–25. [PubMed: 14998108]
- Lennon R, Eisenberg N. Emotional displays associated with preschoolers' prosocial behavior. *Child Dev*. 1987; 58:992–1000. [PubMed: 3608668]

- Lohr, S. Sampling: design and analysis. Nelson Education; 2009.
- Millsap RE, Yun-Tein J. Assessing Factorial Invariance in Ordered-Categorical Measures. *Multivariate Behavioral Research*. 2004; 39(3):479–513.
- National Research Council. *Subjective Well-Being: Measuring Happiness, Suffering, and Other Dimensions of Experience*. Washington, DC: The National Academies Press; 2013.
- Pressman SD, Cohen S. Does positive affect influence health? *Psychol Bull*. 2005; 131(6)
- Ravens-Sieberer U, Devine J, Bevans K, Riley AW, Moon J, Salsman JM, et al. Subjective well-being measures for children were developed within the PROMIS project: presentation of first results. *Journal of clinical epidemiology*. 2014; 67(2):207–218. [PubMed: 24295987]
- Ravens-Sieberer U, Gosch A, Rajmil L, Erhart M, Bruil J, Duer W, et al. KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Rev Pharmacoecon Outcomes Res*. 2005; 5(3): 353–364. [PubMed: 19807604]
- Ravens-Sieberer U, Gosch A, Rajmil L, Erhart M, Bruil J, Power M, et al. The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Value Health*. 2008; 11(4):645–658. [PubMed: 18179669]
- Samejima, F. Graded response model. In: van der Linden, WJ., Hambleton, RK., editors. *Handbook of modern item response theory*. New York, NY: Springer; 1997. p. 85-100.
- Varni JW, Stucky BD, Thissen D, Dewitt EM, Irwin DE, Lai JS, et al. PROMIS Pediatric Pain Interference Scale: an item response theory analysis of the pediatric pain item bank. *J Pain*. 2010; 11(11):1109–1119. [PubMed: 20627819]

APPENDIX

Table A1

PROMIS Pediatric Positive Affect, Parent-Proxy Edition, item-level descriptive statistics, reliability, and factor loadings; data are from Study 1.

Item Stem	Mean (SD)	Floor, (% Never)	Ceiling (% Always)	Item-total correlation (r value)	Test-retest reliability (Intraclass Correlation)	CFA Factor Loadings
My child felt calm.	3.88 (0.82)	0.3	23.0	0.76	0.75	0.83
My child felt peaceful.	3.99 (0.83)	0.3	29.6	0.84	0.73	0.89
My child felt satisfied.	4.03 (0.80)	0.6	29.2	0.88	0.80	0.93
My child felt content.	4.03 (0.82)	0.6	29.2	0.86	0.69	0.91
My child felt grateful.	3.95 (0.90)	0.9	30.8	0.78	0.77	0.84
My child felt thankful.	3.90 (0.92)	1.1	28.7	0.70	0.75	0.77
My child felt positive.	4.07 (0.80)	0.6	30.9	0.86	0.80	0.91
My child felt carefree.	3.87 (0.92)	0.8	26.8	0.78	0.64	0.81
My child felt relaxed.	3.86 (0.84)	0.6	23.0	0.76	0.70	0.82
My child felt comfortable.	4.15 (0.78)	0.4	35.0	0.81	0.72	0.86

Item Stem	Mean (SD)	Floor, (% Never)	Ceiling (% Always)	Item-total correlation (r value)	Test-retest reliability (Intraclass Correlation)	CFA Factor Loadings
My child felt fulfilled.	3.91 (0.84)	0.7	25.1	0.86	0.74	0.91
My child felt respected.	4.07 (0.81)	0.2	33.2	0.77	0.79	0.82
My child felt appreciated.	4.12 (0.80)	0.3	35.6	0.76	0.76	0.82
My child felt proud.	4.07 (0.80)	0.3	31.9	0.83	0.77	0.89
My child had much to be proud about.	4.06 (0.82)	0.2	32.9	0.77	0.67	0.84
My child felt pleased.	3.94 (0.80)	0.3	24.5	0.81	0.78	0.87
My child felt great.	3.98 (0.84)	0.4	28.6	0.88	0.72	0.93
My child felt cheerful.	4.03 (0.79)	0.4	28.9	0.87	0.83	0.93
My child felt happy.	4.14 (0.75)	0.2	33.0	0.87	0.79	0.92
My child had very strong happy feelings.	3.93 (0.87)	0.8	27.2	0.86	0.76	0.90
My child felt joyful.	3.95 (0.85)	0.4	27.8	0.89	0.81	0.94
My child felt delighted.	3.93 (0.86)	0.7	27.6	0.88	0.86	0.92
My life was pleasurable.	4.02 (0.83)	0.8	29.6	0.87	0.82	0.92
My child smiled a lot.	4.13 (0.80)	0.6	35.1	0.77	0.85	0.83
My child laughed a lot.	4.07 (0.85)	0.3	34.3	0.78	0.86	0.84
My child was merry.	3.87 (0.87)	0.8	24.6	0.81	0.82	0.85
My child was in a good mood.	4.04 (0.74)	0.3	26.0	0.81	0.83	0.87
My child was in good spirits.	4.07 (0.73)	0.3	27.3	0.82	0.75	0.88
My child felt good.	4.09 (0.75)	0.4	28.9	0.81	0.65	0.87
My child felt blissful.	3.75 (0.95)	1.7	23.0	0.83	0.84	0.87

Item Stem	Mean (SD)	Floor, (% Never)	Ceiling (% Always)	Item-total correlation (r value)	Test-retest reliability (Intraclass Correlation)	CFA Factor Loadings
My child felt enthusiastic.	3.97 (0.85)	0.4	29.3	0.80	0.76	0.84
My child felt energetic.	3.97 (0.88)	0.4	31.2	0.78	0.73	0.87
My child had a lot of energy.	3.99 (0.89)	0.9	33.2	0.76	0.77	0.87
My child felt wide awake.	3.83 (0.81)	0.4	20.5	0.66	0.70	0.70
My child felt refreshed.	3.82 (0.84)	0.7	21.3	0.83	0.58	0.88
My child felt active.	4.05 (0.89)	0.6	35.8	0.71	0.72	0.77
My child felt full of pep.	3.89 (0.90)	0.9	27.1	0.80	0.63	0.86
My child felt strong.	4.07 (0.80)	0.4	31.8	0.78	0.55	0.83
My child felt healthy.	4.13 (0.84)	0.7	37.7	0.70	0.75	0.75

Table A2

Item Response Theory Item Parameters for the 39 PROMIS Pediatric Positive Affect Item Bank, Parent-Proxy Edition, and Short Form Item Assignment

Item stem	Short Forms		Item Discrimination (a)	Item Thresholds			
	8-item	4-item		b1	b2	b3	b4
My child felt calm.	X		2.10	-3.98	-2.21	-0.77	0.94
My child felt peaceful.	X		2.87	-3.39	-2.14	-0.73	0.65
My child felt satisfied.			4.43	-2.93	-2.09	-0.92	0.53
My child felt content.			3.58	-3.09	-2.06	-0.96	0.56
My child felt grateful.			2.59	-3.05	-1.97	-0.76	0.57
My child felt thankful.			1.81	-3.47	-2.13	-0.81	0.77
My child felt positive.			3.70	-3.04	-2.20	-1.00	0.49
My child felt carefree.			2.49	-3.27	-1.89	-0.69	0.72
My child felt relaxed.			2.27	-3.59	-2.17	-0.72	0.93
My child felt comfortable.			2.96	-3.40	-2.35	-1.20	0.38
My child felt fulfilled.			3.89	-2.90	-1.99	-0.70	0.72
My child felt respected.			2.49	-4.25	-2.44	-0.99	0.48
My child felt appreciated.			2.36	-3.89	-2.60	-1.10	0.40
My child felt proud.			3.02	-3.57	-2.30	-0.99	0.50

Item stem	Short Forms		Item Discrimination (a)	Item Thresholds			
	8-item	4-item		b1	b2	b3	b4
My child had much to be proud about.			2.33	-4.37	-2.39	-1.00	0.51
My child felt pleased.			2.78	-3.73	-2.29	-0.85	0.80
My child felt great.	X	X	4.23	-2.92	-1.97	-0.88	0.53
My child felt cheerful.	X	X	4.69	-3.02	-2.17	-0.92	0.53
My child felt happy.	X	X	3.81	-2.88	-2.33	-1.23	0.46
My child had very strong happy feelings.			3.61	-2.91	-1.89	-0.75	0.64
My child felt joyful.	X	X	4.95	-2.80	-1.79	-0.74	0.56
My child felt delighted.			4.11	-2.88	-1.91	-0.69	0.61
My life was pleasurable.			4.05	-2.81	-2.01	-0.90	0.52
My child smiled a lot.			2.50	-3.45	-2.42	-1.19	0.41
My child laughed a lot.			2.57	-3.80	-2.19	-1.01	0.42
My child was merry.			2.73	-3.16	-1.97	-0.69	0.81
My child was in a good mood.			2.97	-3.60	-2.42	-1.14	0.73
My child was in good spirits.			3.04	-3.60	-2.42	-1.20	0.67
My child felt good.			2.86	-3.43	-2.44	-1.22	0.61
My child felt blissful.			3.21	-2.59	-1.67	-0.49	0.85
My child felt enthusiastic.			2.54	-3.60	-2.20	-0.82	0.63
My child felt energetic.			2.43	-3.67	-2.10	-0.77	0.56
My child had a lot of energy.			2.22	-3.32	-2.16	-0.81	0.51
My child felt wide awake.			1.60	-4.46	-2.65	-0.74	1.27
My child felt refreshed.	X		3.02	-3.03	-2.02	-0.56	0.94
My child felt active.			1.89	-3.90	-2.38	-0.98	0.44
My child felt full of pep.			2.54	-3.17	-1.93	-0.69	0.73
My child felt strong.			2.45	-3.63	-2.44	-1.07	0.54
My child felt healthy.			1.87	-3.82	-2.62	-1.19	0.37

Table A3

Scale-Level descriptive statistics, reliability, and concurrent validity

	Item Bank, 39 Items	SF8, 8 Items	SF4, 4 Items
Scale Descriptive Statistics (S1, n=909)			
Range	10.2–72.8	13.2–67.3	16.8–63.9
Mean (SD)	49.8 (11.2)	49.6 (10.5)	49.3 (9.9)
Floor, %	1 (0.11%)	2 (0.22%)	2 (0.22%)
Ceiling, %	56 (6%)	109 (12%)	176 (19.4%)
Reliability			

	Item Bank, 39 Items	SF8, 8 Items	SF4, 4 Items
Marginal Reliability (S1)	0.97	0.93	0.91
Test-Retest ICC (S1, n=52)	0.70	0.71	0.70
Convergent Validity, Pearson's r			
NIH Toolbox Proxy Positive Affect (S1, n=107)	0.69	0.70	0.70
NIH Toolbox Proxy Life Satisfaction (S1, n=102)	0.83	0.81	0.80
PROMIS Parent Proxy Life Satisfaction (S1, n=502)	0.66	0.65	0.64
PROMIS Parent Proxy Meaning and Purpose (S1, n=464)			
Discriminant Validity, Pearson's r			
PROMIS Parent Proxy Anger (S2, n=1,306)	-	-0.56	-0.50
PROMIS Parent Proxy Anxiety (S2, n=1,306)	-	-0.40	-0.36
PROMIS Parent Proxy Pain Interference (S2, n=1,305)	-	-0.26	-0.26
PROMIS Parent Proxy Physical Stress Experiences (S2, n=1,285)	-	-0.38	-0.36

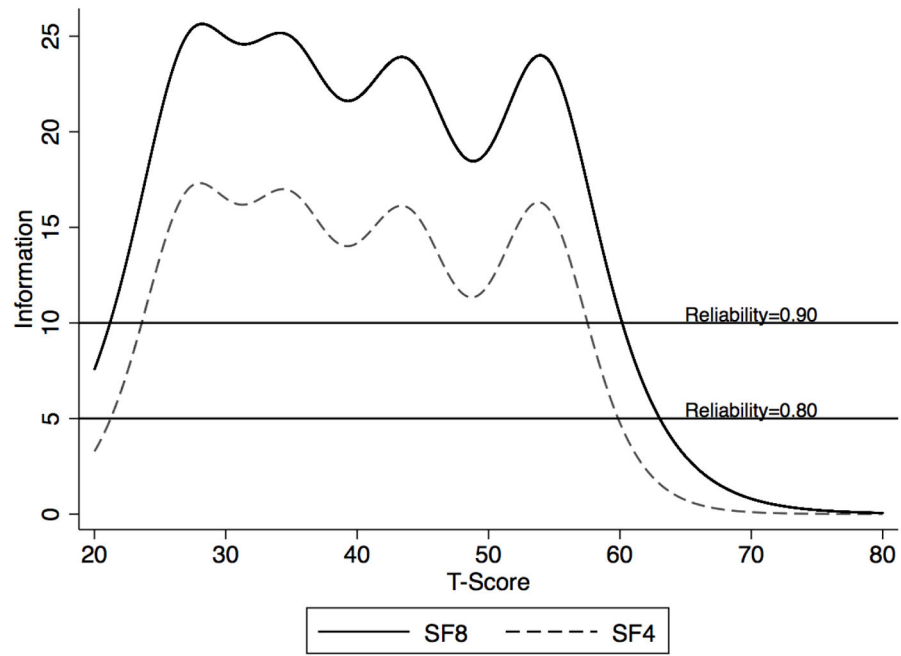


Figure. Test Information Functions for the Positive Affect, Child-Report Edition, Item Bank, 8-item Short Form (SF8), and 4-item Short Form (SF4).

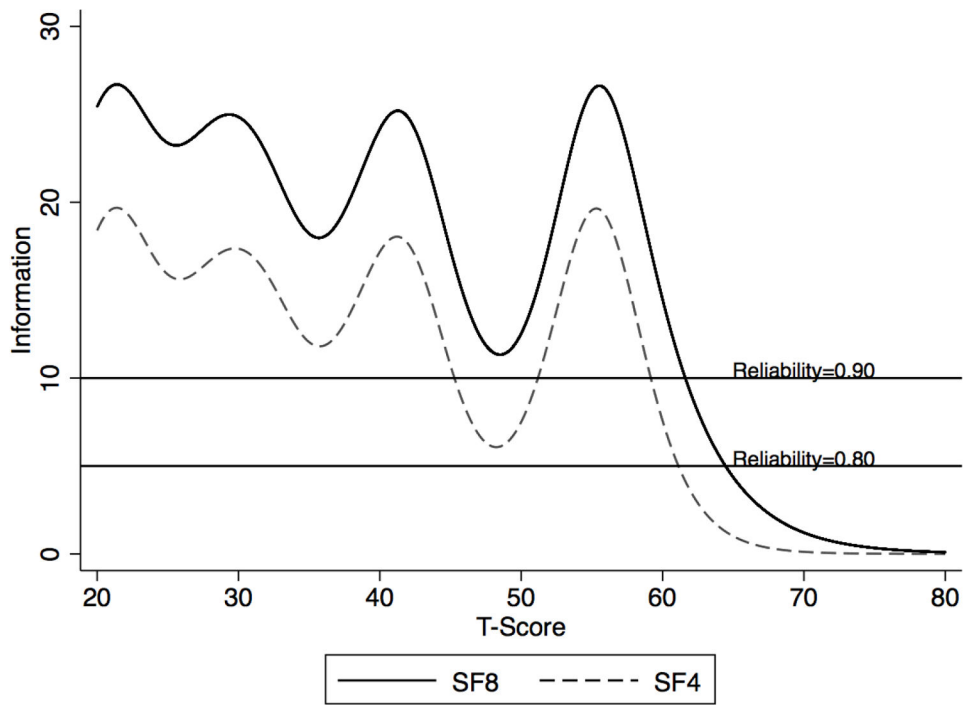


Figure A. Test Information Functions for the Positive Affect, Proxy-Report Edition, 8-item and 4-item Short Forms.

Table 1

Child and parent participants.

Characteristic	Study 1		Study 2	
	Children	Parents	Children	Parents
Sample size	1,874	909	1,004	1,306
Survey location				
Home	992 (53%)	768 (84%)	1,004 (100%)	1,306 (100%)
Clinic	122 (7%)	143 (16%)	0	0
School	760 (40%)	0	0	0
Child Characteristics				
Age, y				
5	0	10 (1%)	0	99 (8%)
6	0	9 (1%)	0	100 (8%)
7	0	8 (1%)	0	100 (8%)
8	253 (14%)	76 (8%)	95 (9%)	96 (7%)
9	190 (10%)	74 (8%)	108 (11%)	108 (8%)
10	155 (8%)	93 (10%)	90 (9%)	91 (7%)
11	181 (10%)	94 (10%)	97 (10%)	97 (7%)
12	156 (8%)	88 (10%)	104 (10%)	105 (8%)
13	251 (13%)	119 (13%)	109 (11%)	109 (8%)
14	206 (11%)	94 (10%)	100 (10%)	101 (8%)
15	166 (9%)	70 (8%)	103 (10%)	103 (8%)
16	164 (9%)	93 (10%)	102 (10%)	101 (8%)
17	152 (8%)	81 (9%)	96 (10%)	96 (7%)
Gender				
Male	992 (53%)	453 (50%)	509 (51%)	661 (51%)
Female	880 (47%)	455 (50%)	495 (49%)	645 (49%)
Race				
White	1439 (77%)	677 (76%)	791 (79%)	1,011 (77%)
African-American or Black	182 (10%)	116 (13%)	82 (8%)	114 (9%)
Asian or Pacific Islander	86 (4%)	36 (4%)	35 (4%)	43 (3%)
Other	167 (9%)	60 (7%)	96 (9%)	138 (10%)
Ethnicity				
Hispanic	293 (16%)	100 (11%)	149 (15%)	212 (16%)
Non-Hispanic	1551 (84%)	784 (89%)	855 (85%)	1,094 (84%)
Special Healthcare Need	---	---	276 (28%)	333 (26%)
Parental and Family Characteristics				

Characteristic	Study 1		Study 2	
	Children	Parents	Children	Parents
Annual household income				
Less than \$40,000	---	---	241 (24%)	341 (26%)
\$40,000 or more			763 (76%)	965 (74%)
Relationship to child				
Mother	---	704 (79%)	---	934 (72%)
Father		141 (16%)		314 (24%)
Other		50 (6%)		55 (4%)
Parental Age (years)				
18–34	---	122 (26%)	---	294 (23%)
35–44		221 (48%)		512 (39%)
45+		120 (26%)		500 (38%)
Parental Educational attainment		136 (29%)		206 (16%)
High school or less	---	299 (65%)	---	474 (36%)
Some college		28 (6%)		626 (48%)
College degree or higher				

Table 2

PROMIS Pediatric Positive Affect, Child Self-Report Edition, item-level descriptive statistics, reliability, and factor loadings; data are from Study 1.

Item Stem	Mean (SD)	Floor (% Never)	Ceiling (% Always)	Item-total correlation (r value)	Test-retest reliability (Intraclass Correlation)	CFA Factor Loadings
I felt calm.	3.72 (0.99)	2.1	23.6	0.73	0.74	0.79
I felt peaceful.	3.85 (1.01)	2.6	30.7	0.79	0.78	0.85
I felt satisfied.	3.91 (0.99)	2.1	31.7	0.82	0.67	0.87
I felt content.	3.87 (0.96)	1.8	28.2	0.78	0.81	0.83
I felt grateful.	3.98 (0.92)	1.2	33.6	0.74	0.81	0.81
I felt thankful.	3.97 (0.95)	1.3	34.4	0.68	0.65	0.75
I felt positive.	3.98 (0.94)	1.3	33.4	0.82	0.75	0.87
I felt carefree.	3.54 (1.16)	6.3	23.7	0.62	0.75	0.66
I felt relaxed.	3.74 (1.03)	2.8	26.1	0.71	0.73	0.77
I felt comfortable.	4.01 (0.93)	1.5	34.3	0.78	0.75	0.84
I felt fulfilled.	3.72 (0.99)	2.3	24.2	0.83	0.79	0.87
I felt respected.	3.73 (1.05)	2.9	27.1	0.74	0.79	0.79
I felt appreciated.	3.83 (1.07)	2.9	32.8	0.74	0.75	0.79
I felt proud.	3.90 (1.02)	2.1	33.4	0.79	0.83	0.84
I had much to be proud about.	3.89 (1.04)	2.3	34.5	0.72	0.79	0.78
I felt pleased.	3.89 (0.91)	1.2	27.2	0.76	0.70	0.81
I felt great.	3.92 (0.97)	1.8	32.1	0.86	0.76	0.90
I felt cheerful.	3.92 (0.97)	2.1	30.9	0.83	0.69	0.89
I felt happy.	4.05 (0.93)	1.8	36.3	0.83	0.84	0.89
I had very strong happy feelings.	3.82 (1.04)	2.7	30.5	0.84	0.74	0.88
I felt joyful.	3.88 (0.97)	1.7	29.7	0.86	0.81	0.91
I felt delighted.	3.80 (1.01)	2.4	28.1	0.85	0.86	0.90
My life was pleasurable.	3.92 (0.98)	1.9	32.7	0.83	0.73	0.88
I smiled a lot.	4.03 (0.98)	1.6	38.5	0.69	0.85	0.74
I laughed a lot.	4.12 (0.95)	1.3	42.7	0.69	0.76	0.75
I was merry.	3.66 (1.07)	4.2	24.0	0.75	0.77	0.80

Item Stem	Mean (SD)	Floor (% Never)	Ceiling (% Always)	Item-total correlation (r value)	Test-retest reliability (Intraclass Correlation)	CFA Factor Loadings
I was in a good mood.	3.95 (0.88)	1.4	27.8	0.77	0.70	0.82
I was in good spirits.	3.90 (0.92)	1.8	27.2	0.79	0.67	0.84
I felt good.	4.04 (0.91)	1.5	34.0	0.77	0.57	0.82
I felt blissful.	3.63 (1.05)	3.5	23.4	0.78	0.86	0.83
I felt enthusiastic.	3.87 (1.01)	2.2	31.6	0.75	0.73	0.79
I felt energetic.	3.90 (1.00)	2.0	32.8	0.74	0.73	0.82
I had a lot of energy.	3.92 (1.03)	2.1	35.8	0.73	0.79	0.81
I felt wide awake.	3.52 (1.07)	3.9	20.8	0.60	0.79	0.64
I felt refreshed.	3.67 (1.04)	3.4	24.2	0.76	0.72	0.81
I felt active.	4.03 (1.00)	2.1	39.5	0.67	0.71	0.73
I felt full of pep.	3.73 (1.11)	4.0	29.9	0.74	0.75	0.78
I felt strong.	3.99 (1.00)	2.1	37.5	0.71	0.74	0.76
I felt healthy.	4.16 (0.94)	1.5	44.2	0.61	0.66	0.68

Table 3

Item Response Theory Item Parameters for the 39 PROMIS Pediatric Positive Affect Item Bank, Child Self-Report Edition, and Short Form Item Assignment.

Item stem	Short Forms		Item Discrimination	Item Thresholds			
	8-item	4-item		b1	b2	b3	b4
I felt calm.	X		2.40	-2.56	-1.57	-0.48	0.71
I felt peaceful.	X		3.03	-2.24	-1.52	-0.52	0.50
I felt satisfied.			3.49	-2.32	-1.54	-0.64	0.40
I felt content.			2.81	-2.55	-1.65	-0.65	0.55
I felt grateful.			2.64	-2.83	-1.91	-0.69	0.40
I felt thankful.			1.99	-3.10	-2.01	-0.74	0.44
I felt positive.			3.51	-2.58	-1.64	-0.73	0.35
I felt carefree.			1.64	-2.30	-1.40	-0.30	0.93
I felt relaxed.			2.11	-2.56	-1.56	-0.47	0.73
I felt comfortable.			2.98	-2.63	-1.75	-0.80	0.35
I felt fulfilled.			3.43	-2.29	-1.46	-0.39	0.67
I felt respected.			2.44	-2.37	-1.46	-0.43	0.64
I felt appreciated.			2.42	-2.38	-1.46	-0.52	0.44
I felt proud.			2.92	-2.44	-1.56	-0.59	0.38
I had much to be proud about.			2.21	-2.62	-1.65	-0.60	0.40
I felt pleased.			2.58	-2.87	-1.85	-0.67	0.62
I felt great.	X	X	4.06	-2.32	-1.57	-0.66	0.36
I felt cheerful.	X	X	3.84	-2.22	-1.49	-0.62	0.45
I felt happy.	X	X	3.80	-2.36	-1.78	-0.91	0.29
I had very strong happy feelings.			3.75	-2.14	-1.34	-0.49	0.44
I felt joyful.	X	X	4.44	-2.24	-1.46	-0.55	0.45
I felt delighted.			3.98	-2.19	-1.40	-0.48	0.50
My life was pleasurable.			3.60	-2.39	-1.58	-0.61	0.37
I smiled a lot.			2.02	-3.01	-1.89	-0.88	0.29
I laughed a lot.			2.09	-3.09	-2.01	-0.94	0.15

Item stem	Short Forms		Item Discrimination	Item Thresholds			
	8-item	4-item		b1	b2	b3	b4
I was merry.			2.58	-2.14	-1.38	-0.40	0.75
I was in a good mood.	X		2.81	-2.72	-1.99	-0.85	0.55
I was in good spirits.			2.88	-2.59	-1.72	-0.73	0.60
I felt good.			2.69	-2.71	-1.89	-0.90	0.37
I felt blissful.			2.77	-2.18	-1.39	-0.27	0.75
I felt enthusiastic.			2.43	-2.56	-1.66	-0.58	0.48
I felt energetic.			2.32	-2.69	-1.70	-0.64	0.44
I had a lot of energy.			2.20	-2.72	-1.70	-0.64	0.36
I felt wide awake.			1.51	-2.79	-1.51	-0.15	1.18
I felt refreshed.	X		2.67	-2.20	-1.44	-0.32	0.75
I felt active.			1.88	-2.91	-1.91	-0.89	0.28
I felt full of pep.			2.29	-2.25	-1.42	-0.39	0.56
I felt strong.			2.14	-2.75	-1.86	-0.75	0.31
I felt healthy.			1.67	-3.30	-2.23	-1.17	0.12

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Scale-Level descriptive statistics, reliability, and concurrent validity

	Item Bank, 39 Items	SF8, 8 Items	SF4, 4 Items
Descriptive Statistics (Study 1, n=1,874)			
Range	12.3–72.3	19.1–66.2	22.0–63.0
Mean (SD)	49.6 (10.4)	49.4 (9.6)	49.4 (9.3)
Floor, n (%)	3 (0.2%)	8 (0.4%)	16 (0.9%)
Ceiling, n (%)	101 (5.4%)	221 (11.8%)	367 (19.5%)
Reliability			
Marginal Reliability (Study 1, n=1,874)	0.98	0.94	0.92
Test-Retest ICC (Study 1 Retest Sample, n=114)	0.75	0.75	0.77
Convergent Validity, Pearson's r			
NIH Toolbox Positive Affect (Study 1, n=177)	0.88	0.87	0.85
NIH Toolbox Life Satisfaction (Study 1, n=181)	0.61	0.58	0.57
PROMIS Pediatric Life Satisfaction (Study 1, n=1,033)	0.83	0.81	0.81
PROMIS Pediatric Meaning and Purpose (Study 1, n=969)	0.71	0.69	0.68
KIDSCREEN Psychological Well Being Scale (Study 2, n=498)	-	0.68	0.67
Discriminant Validity, Pearson's r			
PROMIS Pediatric Anger (Study 1, n=224)	-0.55	-0.52	-0.51
PROMIS Pediatric Anger (Study 2, n=1,004)	-	-0.54	-0.49
PROMIS Pediatric Anxiety (Study 1, n=226)	-0.33	-0.31	-0.31
PROMIS Pediatric Anxiety (Study 2, n=1,004)	-	-0.39	-0.34
PROMIS Pediatric Depressive Symptoms (Study 1, n=226)	-0.54	-0.50	-0.52
PROMIS Pediatric Pain Interference (Study 2, n=1,004)	-	-0.33	-0.29
PROMIS Pediatric Physical Stress Experiences (Study 2, n=999)	-	-0.43	-0.39