Commentary: The Patient-Reported Outcome Measurement Information System (PROMIS®) for Children and Youth: Application to Pediatric Psychology

Christopher B. Forrest,[1,2] MD, PhD, Katherine B. Bevans,[1,2] PhD, Carole Tucker,[3] PhD, Anne W. Riley,[4] PhD, Ulrike Ravens-Sieberer,[5] PhD, William Gardner,[6,7,8,9,10] PhD, and Kathleen Pajer,[1,2,11] MD, MPH

[1]The Children's Hospital of Philadelphia, [2]Department of Pediatrics, University of Pennsylvania School of Medicine, [3]College of Health Professions & Social Work, Temple University, [4]Department of Population and Family Health Sciences, Johns Hopkins Bloomberg School of Public Health, [5]Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Research Unit Child Public Health, University Medical Center Hamburg-Eppendorf, [6]Department of Pediatrics, [7]Department of Psychology, [8]Department of Psychiatry, The Ohio State University, Columbus, [9]Department of Obstretics/Gynaecology, [10]Department of Pediatrics, & Epidemiology, Dalhousie University, [11]IWK Health Centre, Halifax, and [12]Department of Psychiatry, Dalhousie University Faculty of Medicine

All correspondence concerning this article should be addressed to Christopher B. Forrest, MD, PhD, Professor, Children's Hospital of Philadelphia, 34th Street and Civic Center Blvd, Philadelphia, PA, 19104, USA. E-mail: forrestc@email.chop.edu

Introduction

Assessing the outcomes of healthcare interventions on children and families from their point of view has long been a central goal of pediatric psychology. This approach to outcome assessment is now being embraced in many areas of healthcare under the aegis of patient reported outcomes (PROs)—that is, evaluating health from the perspectives of patients themselves. Their growing importance in clinical research is highlighted by the 2009 guidance issued by the Food and Drug Administration (U.S. DHHS, 2010) on necessary criteria for using PROs to support claims for medical product labeling, the federal government's establishment of the Patient-Centered Outcomes Research Institute (Clancy & Collins, 2010) and their markedly increased use in clinical trials (Rahimi, Malhotra, Banning, & Jenkinson, 2010). Ample evidence has accrued in support of the validity and practicality of administering PROs to children (Bevans, Riley, Moon, & Forrest, 2010).

As the demand for pediatric participation in clinical trials has grown, the interest in trying to measure pediatric PROs has also increased. The validity and reliability of children as informants about their own states have been supported in large studies using instruments from Healthy Pathways (Bevans, Riley, Forrest, 2010), the Peds QL,(Varni et al., 2007), the Child Health and Illness Profile (Rebok et al., 2001; Riley et al., 2004), and the KIDSCREEN (Ravens-Sieberer et al., 2005, 2008).

The NIH launched in 2004 a program of research called the Patient Reported Outcome Measurement Information System (PROMIS®) (Cella, Yount et al., 2007). The goal of PROMIS is to provide clinicians and researchers access to efficient, precise, valid, and responsive adult- and child-reported measures of health (see www.nihpromis.org for more information). These measures are rapidly proliferating throughout clinical and behavioral research, epidemiology and population surveillance, and clinical practice.

PROMIS comprises a cooperative group of research sites and centers, a unique mixed-methods instrument development process, many measures of health and well-being, and an informatics platform that enables web-based static and dynamic administration (Cella, Yount et al., 2007; Cella et al., 2010; Gershon, Rothrock, Hanrahan, Bass, & Cella, 2010; Riley et al., 2010). This article

presents the science of PROMIS and discusses how these methods and tools may solve some of the issues facing pediatric psychology in measuring health outcomes in children.

The Science of PROMIS

PROMIS uses a domain-specific measurement approach. Domains are defined as clinically coherent and empirically unidimensional health attributes that cut across diseases, although disorders may have characteristic profiles of these attributes. Table I shows the PROMIS Pediatric domains and their definitions that have been developed (can be

used now) and are under development (ready for use in 2013).

The PROMIS mixed-methods approach to creating an item bank is summarized in the Figure 1. Item bank development begins with defining the breadth and depth of the content of the target domain. Input is obtained from content experts, the scientific literature, previously developed measures, analysis of existing scales, and perspectives of children and parents (Magasi et al., 2011). Either semi-structured interviews or focus groups with children and parents are done to ensure that the domain covers all facets of the health attribute.

Building an item bank that comprehensively measures the full range of the health domain's manifestations, from

Table I. *Pediatric PROMIS Item Banks Developed and Under Development*

Domain	Brief definition[b]
Self-reported physical health	
Physical functioning: upper extremity function	Activities that require use of the upper extremity including shoulder, arm, and hand activities. Examples include writing, using buttons, or opening containers.
Physical functioning: mobility	Activities of physical mobility such as getting out of bed or a chair to activities such as running.
Physical activity[a]	Level of bodily movement as assessed by performance of activities that require physical actions, ranging from habitual (activities of daily living) to more complex activities that require a combination of skills, often within a social context.
Pain interference	Impact of pain on physical, mental and social activities.
Pain quality[a]	The nature, characteristics, intensity frequency, and duration of pain.
Fatigue	Overwhelming, debilitating and sustained sense of exhaustion that decreases one's ability to carry out daily activities and to function at one's usual level in family or social roles.
Asthma impact	Symptoms specific to asthma that are not adequately captured by other general item banks.
Self-reported mental health	
Depression	Feelings of hopelessness, helplessness, and worthlessness, negative mood (e.g., sadness), decrease in positive affect (e.g., loss of interest), negative views of the self (e.g., worthlessness, low self-esteem), and negative social cognition (e.g., loneliness, interpersonal alienation).
Anxiety	Feelings of fear, anxious misery (e.g., worry), and hyperarousal (e.g., nervousness).
Anger	Angry mood (e.g., irritability, reactivity) and aggression (verbal and physical).
Stress experiences: somatic[a]	Physically experienced sensations associated with responses to internal or external challenges.
Stress experiences: psychological[a]	Thoughts or feelings about self and the world in the context of environmental or internal challenges.
Positive affect[a]	Feelings of contentment, calmness, love, pride, happiness, excitement, and energy.
Life satisfaction[a]	Global and context-specific evaluations of life overall, life conditions, and comparisons of one's life with others' lives.
Meaning and purpose[a]	One's sense that life has purpose and there are good reasons for living, including a sense of hopefulness, optimism, and goal-directedness.
Self-reported social health	
Peer relationships	Quality of relationships with friends and other acquaintances.
Family belonging[a]	Beliefs and knowledge about the extent to which one has a sense that they are a part of the family, feel loved and cared about, have sense of being part of the family, and the sense of being valued and accepted.
Family involvement[a]	Consistency and quality with which an individual experiences or engages in family behaviors and regular activities.

[a]The item bank is under development and will be ready for public use by 2013.
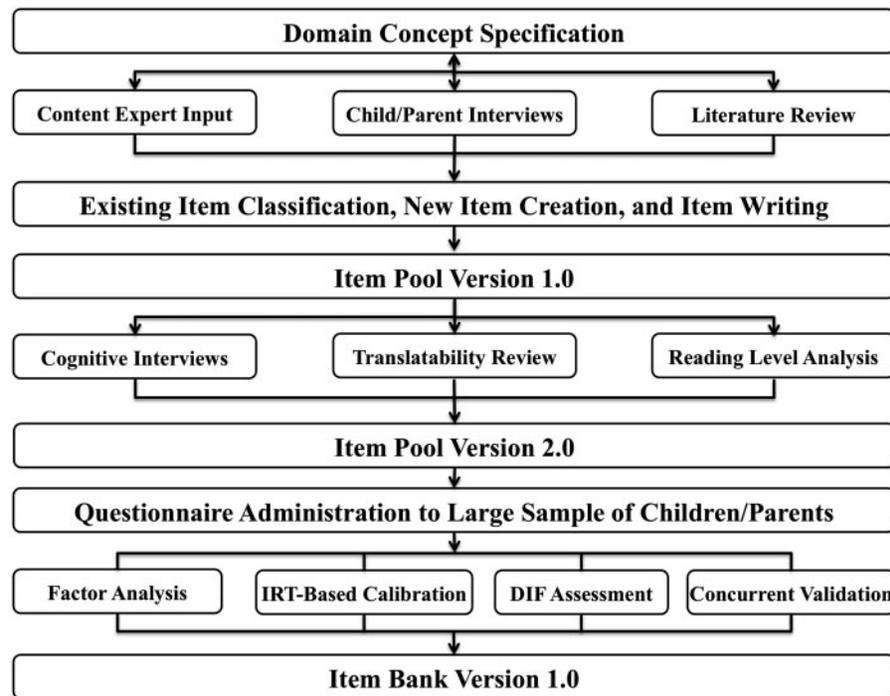[b]See http://www.nihpromis.org for more detailed descriptions of item banks.

Figure 1. Pediatric PROMIS item bank creation process. Development of a PROMIS Pediatric item bank starts with a set of qualitative methods that produce a theoretically informed item pool. Version 1.0 of the item pool is refined as a result of cognitive interviews, translatability review, and reading level analysis. The item pool version 2.0 questionnaire is then administered to a large, representative sample of children and their parents. Data from this survey administration undergo psychometric analyses, which lead to further modification of the item pool. At the end of this process, an item bank is formed, which is ready for population norming, clinical validation, and responsiveness testing.

the lowest to the highest levels, starts with a review of the measurement literature that is comprehensive in scope, systematic, and reproducible (Klem et al., 2009). Relevant articles are retrieved and abstracted to identify in-scope instruments; then instrument developers are contacted to request inclusion of their instrument's items in an item library.

Once the item library is formed, an item classification process is done to assign items to domain facets and prune redundancies (DeWalt et al., 2007). At this point, it is critical to array the items within each facet so that they are arranged from the best to worst or strongest indication of the facet to the weakest, whatever dimension is appropriate. In this way, the need for additional items to cover the entire domain becomes clear. Items identified in the literature are then rewritten to conform to PROMIS item writing standards. New items are created to cover conceptual gaps across the full range of the domain.

Every item undergoes cognitive interviews to assure children's comprehension of item content that the recall period is sensible, and other cognitive processes do not undermine usability. PROMIS Pediatric interview methods are consistent with international standards (Willis, 2004)

and have been published elsewhere (Fortune-Greeley, et al., 2009; Irwin, Varni, Yeatts, & DeWalt, 2009).

To ensure that PROMIS item banks measure the same domain concepts across languages, an expert in translation reviews items to identify idiomatic expressions, complex sentences, and concepts that are not easily translated into other languages. This translatability review leads to removal or revision of problematic items. All PROMIS item banks have been translated into Spanish and many other translations are underway. Item translation follows a universal language approach (one translation per language), which is consistent with recommendations of the ISPOR PRO Outcomes Translation and Linguistic Validation Task Force (Wild et al., 2005, 2009), international guidelines published by the IQOLA (IQOLA, 2011), MAPI (MAPI, 2011), and MOS institutes (MOS, 2011). Cognitive interviews are done after the translation process to pretest the comprehensibility of the translated version with native speaking children. Once item pools are developed and refined using qualitative methods, they are administered to large populations of individuals. Survey data are then subjected to psychometric testing using a combination of traditional and modern methods (Reeve et al., 2007). Analyses are conducted to confirm assumptions

about dimensionality of the items hypothesized to be within a single item bank, to test for differential item functioning (DIF) across sociodemographic groups, and to calibrate the items to support development of fixed-length, short forms, and computerized adaptive test (CAT) versions of the instruments.

Calibration is done using IRT-graded response models (Samejima, 1997). IRT models describe in probabilistic terms the relationship between a person's response to an item and her level of the health domain that the instrument measures (Hambleton, Swaminathan, & Rogers, 1991; Reeve et al., 2007). Parameter estimates generated for each item in the model include the item's discrimination (how well the item differentiates among people with varying levels of the underlying health domain) and item difficulty (the level of health that a person must have in order to endorse a specific item response). Inspection of item difficulty parameters highlights gaps in the measurement of the health domain, when there are no or too few items that provide information about respondents with a specific level of health.

PROMIS IRT methods support the development of fixed-length, short forms and CATs, both of which significantly reduce respondent burden without compromising measurement precision. Another advantage is that IRT permits statistical linking between child and adult item banks (assuming that the item banks indeed measure the same concept), so that measurements on a given domain can be placed on the same scale across the life course. Several pediatric–adult linkage studies are underway within the PROMIS cooperative group.

Application of PROMIS Measures

Items in PROMIS fixed-length, short forms (from 4 to 8 items per domain) are chosen from an item bank based on the item's measurement characteristics. Short forms use the most informative items to achieve satisfactory measurement precision while minimizing respondent and administrative burden (Cella, Gershon, Lai, & Choi, 2007). The available PROMIS short forms have been designed to provide an equal level of precision across the entire domain. Such short forms are used in populations in which respondents may vary widely in the outcome of interest, and the score can accurately capture the level of health across a wide range. Short forms can also be customized to measure more precisely around a meaningful level of health at the expense of increased error at less critical levels. For example, if one wanted to measure outcomes of patients with chronic pain, items clustered around the high end of the

pain interference item bank would be chosen to provide the most discrimination of values.

With respondents required to answer just 4–8 items per item bank, PROMIS CATs produce efficient estimates of the level of self-reported health with very high precision. CATs use software algorithms to select optimal items on the basis of a respondent's sequence and overall patterns of responses. The challenges of pediatric assessment, which require large item sets for wide age ranges, may be particularly suited to the benefits of a CAT platform (Jacobusse, van Buuren, & Verkerk, 2006; Jacobusse & Buuren, 2007). The initial item that the CAT presents is typically in the mid-range of the domain concept. An estimate of the respondent's health is determined as well as the corresponding error in this estimate. Subsequent items chosen for administration refine the estimate and are chosen to match the estimated level of health. If the respondent endorses an item, a slightly more challenging item is presented next, and vice-versa. This technique quickly converges on the respondent's estimated level of health for a given domain. Stopping rules are based on specification of the desired level of measurement precision (reflected in the updated standard errors generated with each item response), number of items administered (maximum), a length of time, content coverage, when the estimated score is converging and minimal change is observed after each item iteration, or some combination of these criteria.

Opportunities for Researchers

PROMIS provides a web-based platform, called Assessment Center[SM] (AC), for implementing studies using PROs. AC can be accessed via the PROMIS web site (www.nihpromis. org). The costs for running Assessment Center are covered by grants and contracts with the National Institutes of Health, making it a free service. Instruments are copyrighted, but PROMIS allows researchers to use item banks and their short forms and CATs without a license fee. AC enables researchers to create study-specific websites that capture participant data securely. Studies can include PROMIS measures (short forms and CATs) within the Assessment Center library as well as custom instruments created or entered by the researcher. Study investigators own data, and there is no requirement for data sharing with PROMIS or AC. Any PROMIS measure can be downloaded for administration on paper or be included in an online study, which can be accessed with a personal or tablet computer.

The AC enables customization of items or instruments (e.g., format, randomization, skip patterns), real-time scoring of CATs, storage of protected health information in a separate, secure database, automated accrual reports, real-time data export, graphing of individual outcome scores, and ability to capture endorsement of online consent forms among many other features. Based on the user's specific institutional approval, electronic consent forms can be uploaded and private health information flagged to allow limited access by study personnel.

Until recently, data about health outcomes have mostly been obtained from parents or health care providers. PROMIS has capitalized on advances in child-reported measurement science (Bevans, Riley, Moon et al., 2010) and provides pediatric psychologists with a broad array of self-reported outcome tools that can be administered to children as young as age 8-years old as well as proxy forms for parents. It is no longer necessary to leave the voices of children out of psychological research and clinical practice. PROMIS Pediatric items are developmentally appropriate, so new instruments do not have to be used as the child grows. PROMIS instruments and administration methods enable efficient (few items needed to assess a given health outcome) and accurate (high reliability and validity) assessment of child-reported physical, mental, and social health.

The diverse group of constructs used to define health outcomes is another benefit of PROMIS for pediatric psychologists. Each item bank is theoretically grounded, and the items are developed in a standardized way, employing state-of-the-art mixed methods. PROMIS measures can be used across conditions, and enable between-study comparisons, because the outcomes are on the same scale. Moreover, ongoing research is determining whether pediatric and adult measures of the same health domain can be statistically linked such that different items may be used for children and adults but the scores that the item banks produce will be on the same scale.

An example of how PROMIS can provide value in research in pediatric psychology would be its use in studies on pediatric medically unexplained symptoms (PMUS). PMUS is a group of symptoms that are prevalent and expensive, but for which there is little effective treatment. Over 19 million children and adolescents in the United States suffer from PMUS each year (Eminson, 2007; Perquin, Hazebroek-Kampschreur, Hunfeld, Bohnen et al., 2000). The cost of PMUS is significant, both in lost function for the child and the parent who loses days at work, and in healthcare dollars (Campo, Jansen-McWilliams, Comer, & Kelleher, 1999; Perquin,

Hazebroek-Kampschreur, Hunfeld, van Suijlekom-Smit et al., 2000). Examples include abdominal pain, headache, dysuria, pelvic pain, syncope, fatigue, and arthralgias. Researchers have long theorized that this collection of symptoms may actually represent one or two syndromes whose mechanism lies in abnormalities in the interconnected biological systems for stress response, immune function, pain, and psychological state. One of the critical obstacles in testing this theory is the absence of a detailed, reliable, and valid pediatric outcome assessment system that is relevant across the myriad diagnoses that are assigned to these children. The use of PROMIS could substantially advance this important area in pediatric psychology research.

Conclusions

PROMIS is becoming a standard for patient-reported outcomes in research and clinical practice. The substantial investment and commitment by the NIH to this initiative, the unique science, and clinical cogency of instruments are all contributors to the potential of the system to become a common denominator in health assessment. Standardization advances scientific knowledge because common outcomes are used across studies thereby permitting syntheses. Evidence-based medicine is founded not only on the clinical trial, but also on the meta-analytic synthesis of results across multiple trials. This synthesis is greatly facilitated by the widespread use of common measures.

PROMIS is an ideal tool for pediatric psychologists interested in measuring child self-reported health in research or clinical practice. It provides efficient and precise measurement across a range of health attributes. All of the biopsychosocial research models motivating pediatric psychological research recognize that health is intrinsically multidimensional. Moreover, many pediatric psychological interventions are premised on the idea that coordinated psychological and physiological interventions are mutually reinforcing. Teasing out the subtleties of how psychological outcomes mediate physiological outcomes, and vice-versa, encourages a multidimensional approach to measurement.

Psychologists pioneered and continue to make major contributions to the measurement of human attributes. Further advances in measurement science are certainly likely, but at this point in history, perhaps most critical to advancing measurement science is the adoption of standardized, well-designed assessment tools across entire fields of health and psychological research.

Acknowledgments

Funding

References

Bevans, K. B., Riley, A. W., Moon, J., & Forrest, C. B. (2010). Conceptual and methodological advances in child-reported outcomes measurement. *Expert Review of Pharmacoeconomics & Outcomes Research, 10*(4), 385–396.

Bevans, K. B., Riley, A. W., & Forrest, C. B. (2010). Development of the healthy pathways child-report scales. *Quality of Life Research, 19*(8), 1195–1214.

Campo, J. V., Jansen-McWilliams, L., Comer, D. M., & Kelleher, K. J. (1999). Somatization in pediatric primary care: Association with psychopathology, functional impairment, and use of services. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*(9), 1093–1101.

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., . . . Fries, J. F. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care, 45*(5 Suppl. 1), S3–S11.

Cella, D., Gershon, R., Lai, J. S., & Choi, S. (2007). The future of outcomes measurement: Item banking, tailored short-forms, and computerized adaptive assessment. *Quality of Life Research, 16*(Suppl. 1), 133–141.

Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., . . . Hays, R.; PROMIS Cooperative Group. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *Journal of Clinical Epidemiology, 63*(11), 1179–1194.

Clancy, C., & Collins, F. S. (2010). Patient-Centered Outcomes Research Institute: The intersection of science and health care. *Science Translational Medicine, 2*(37), 37cm18.

DeWalt, D. A., Rothrock, N., Yount, S., & Stone, A. A. (2007). Evaluation of item candidates: The PROMIS qualitative item review. *Medical Care, 45*(5 Suppl. 1), S12–S21.

Eminson, D. M. (2007). Medically unexplained symptoms in children and adolescents. *Clinical Psychology Review, 27*(7), 855–871.

Fortune-Greeley, A. K., Flynn, K. E., Jeffery, D. D., Williams, M. S., Keefe, F. J., Reeve, B. B., . . . Weinfurt, K. P. (2009). Using cognitive interviews to evaluate items for measuring sexual functioning across cancer populations: Improvements and re-maining challenges. *Quality of Life Research, 18*(8), 1085–1093.

Gershon, R. C., Rothrock, N., Hanrahan, R., Bass, M., & Cella, D. (2010). The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *Journal of Applied Measurement, 11*(3), 304–314.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newburg Park, CA: Sage.

IQOLA. (2011). The international quality of life assess-ment (IQOL) project. Retrieved from http://www .iqola.org

Irwin, D. E., Varni, J. W., Yeatts, K., & DeWalt, D. A. (2009). Cognitive interviewing methodology in the development of a pediatric item bank: A patient re-ported outcomes measurement information system (PROMIS) study. *Health and Quality of Life Outcomes, 7*, 3.

Jacobusse, G., & Buuren, S. (2007). Computerized adap-tive testing for measuring development of young chil-dren. *Stat Medical, 26*(13), 2629–2638.

Jacobusse, G., van Buuren, S., & Verkerk, P. H. (2006). An interval scale for development of children aged 0-2 years. *Stat Medical, 25*(13), 2272–2283.

Klem, M., Saghafi, E., Abromitis, R., Stover, A., Dew, M. A., & Pilkonis, P. (2009). Building PROMIS item banks: Librarians as co-investigators. *Quality of Life Research, 18*(7), 881–888.

Magasi, S., Ryan, G., Revicki, D., Lenderking, W., Hays, R. D., Brod, M., . . . & Cella, D. (2011). Content validity of patient-reported outcome mea-sures: Perspectives from a PROMIS meeting. *Quality of Life Research*. Advance Online Publication. doi:10.1007/s11136-011-9990-8

MAPI. (2011). The MAPI linguistic validation process. Retrieved from http://www.mapi-institute.com /linguistic-validation

MOS. (2011). MOS translation criteria. *Medical Outcomes Trust Bulletin, July,* Retrieved from http://www .outcomes-trust.org/bulletin/0797blltn.htm

Perquin, C. W., Hazebroek-Kampschreur, A. A., Hunfeld, J. A., Bohnen, A. M., van Suijlekom-Smit, L. W., Passchier, J., & van der Wouden, J. C (2000). Pain in children and adolescents: A common experience. *Pain, 87*(1), 51–58.

Perquin, C. W., Hazebroek-Kampschreur, A. A., Hunfeld, J. A., van Suijlekom-Smit, L. W., Passchier, J., & van der Wouden, J. C. (2000). Chronic pain among children and adolescents: Physician consultation and medication use. *The Clinical Journal of Pain, 16*(3), 229–235.

Rahimi, K., Malhotra, A., Banning, A. P., & Jenkinson, C. (2010). Outcome selection and role of patient reported outcomes in contemporary cardiovascular trials: Systematic review. *BMJ, 341*, c5707.

Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., . . . Jean Kilroe, J.; the European KIDSCREEN Group. (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics & Outcomes Research, 5*(3), 353–364.

Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Power, M., . . . Kilroe, J; KIDSCREEN Group. (2008). The KIDSCREEN-52 quality of life measure for children and adolescents: Psychometric results from a cross-cultural survey in 13 European countries. *Value Health, 11*(4), 645–658.

Rebok, G., Riley, A., Forrest, C., Starfield, B., Green, B., Robertson, J., & Tambor, E. (2001). Elementary school-aged children's reports of their health: A cog-nitive interviewing study. *Quality of Life Research, 10*(1), 59–70.

Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., & Thissen, D (2007). Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care, 45*(5 Suppl 1), S22–S31.

Riley, A. W., Forrest, C. B., Rebok, G. W., Starfield, B., Green, B. F., Robertson, J. A., & Friello, P. (2004). The Child Report Form of the CHIP-Child Edition: Reliability and validity. *Medical Care, 42*(3), 221–231.

Riley, W. T., Rothrock, N., Bruce, B., Christodolou, C., Cook, K., Hahn, E. A., & Cella, D. (2010). Patient-reported outcomes measurement information system (PROMIS) domain names and definitions re-visions: Further evaluation of content validity in IRT-derived item banks. *Quality of Life Research, 19*(9), 1311–1321.

Samejima, F. (1997). Graded response model. In W. J. van der Linden, & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). New York, NY: Springer.

U.S. Department of Health and Human Services, F. a. D. A. (2010). Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. Retrieved from http://www.fda.gov/downloads/Drugs /GuidanceComplianceRegulatoryInformation /Guidances/UCM193282.pdf

Varni, J. W., Limbers, C. A., & Burwinkle, T. M. (2007). How young can children reliably and validly self-report their health-related quality of life?: An analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. *Health and Quality of Life Outcomes, 5*, 1.

Wild, D., Eremenco, S., Mear, I., Martin, M., Houchin, C., Gawlicki, M., ... Molsen, E. (2009). Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: The ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. *Value Health, 12*(4), 430–440.

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. and ISPOR Task Force for Translation and Cultural Adaptation. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health, 8*(2), 94–104.

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks: Sage Publications.