Rasch analysis of Chinese Cerebral Palsy Quality of Life for Children (C-CP QOL-Child)

By

Sam Ka Lam, Sam

A thesis submitted to
The Hong Kong Institute of Education
in partial fulfillment of the requirements for
the degree of Doctor of Education

2012

STATEMENT OF ORIGINALITY

I, Sam Ka Lam, Sam, hereby declare that I am the sole author of the thesis and the material presented in this thesis is my original work except those indicated in the acknowledgement. I further declare that I have followed the Institute's policies and regulations on Academic Honesty, Copy Right and Plagiarism in writing the Thesis and no material in this thesis has been published or submitted for a degree in this or other universities.

_____

SAM KA LAM, SAM
September 2012

Thesis Examination Panel Approval

Members of the Thesis Examination Panel approve the thesis of Sam Ka Lam, Sam, defended on 27/07/2012.

Supervisors

Prof Lo Sing Kai
Chair Professor (Principal Supervisor)
Faculty of Arts and Sciences
The Hong Kong Institute of Education

Dr Chen Shihui
Associate Professor (Associate Supervisor)
Department of Physical Education
The Hong Kong Institute of Education

Examiners

Prof Wang Wen Chung
Chair Professor (Internal Examiner)
Department of Psychological Studies
The Hong Kong Institute of Education

Prof Hsieh Ching-Lin
Professor (External Examiner)
School of Occupational Therapy
National Taiwan University

Approved on behalf of the Thesis Examination Panel:

_____

Chair, Thesis Examination Panel
Prof Cheung Hin Tat
Professor
Department of Linguistics and Modern Language Studies
The Hong Kong Institute of Education

ABSTRACT

Rasch analysis of Chinese Cerebral Palsy Quality of Life for Children (C-CP QOL-Child)

by SAM, Sam Ka Lam

The Hong Kong Institute of Education

Abstract

The Chinese Cerebral Palsy Quality of Life for Children (C CP QOL-Child) is the first instrument developed to measure quality of life of (QOL) children with cerebral palsy in Chinese speaking populations. The aim of the study was to examine the psychometric properties of C CP QOLChild using Item Response Theory Models. We were particularly interested to know how intervention strategies could be designed for individuals based on the item scores. 145 primary caregivers (mostly mothers; mean age: 39.2) and their children with cerebral palsy (age: 4-12) were invited to complete the C CP QOL-Child questionnaire. Data were analyzed using Rasch analysis. The results indicated that item difficulty estimates were aligned with person ability values, indicating that the items in the scale generally demonstrated an appropriate depth and width for measuring QOL of persons in the target population. The results in study 2 (Part II) also showed that after dropping the 8 items in the dimension pain and impact of disability in the 65-item parent proxy scale, the revised 57-item scale exhibits unidimensionality (separation index = 4.43, r = 0.95); hence the total score computed from the 57 items adequately reflects the level of QOL of the child as perceived by the caregiver. We further found that the Rasch item difficulty estimates demonstrated an overall item hierarchy; hence therapists can expect a pattern of performance by a child with CP that is based on the established order of item difficulty. Hence, the hierarchical structure identified in the study may be useful for designing tailor-made interventions with an aim of improving QOL.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my principal supervisor, Prof Lo Sing Kai, my associate supervisor, Dr Chen Shihui, whose encouragement, guidance and support from the initial to the final stage enabled me to develop an in-depth understanding of the issue. Thanks Prof Lo, I have acquired valuable insights throughout your instructions, not only in academic studies but also enthusiasm and vigor in life.

My sincere acknowledgement also goes to Prof Wang Wen Chung and Prof Hsieh Ching-Lin, the committee members. They have provided me lots of constructive suggestions and comments not only on the wordings and contents but also the overall framework of the study with their professional experience and extensive knowledge among the fields. Special thanks to Prof Wang Hui-Yi, without her support and network from Taiwan I could not get the job done.

I offer my regards and blessings to all of those who supported me in any respect during the completion of the study. They are: Chunxiao Li, Jacob Xu, Iris Zhou, Amy Jia, Fiona Ching, Andrew Smith, and all the staff in Graduate School.

Last but not least, I would like to thank my mum, my younger brother and my friends, for their encouragement and support during the journey. Thank you so much.

TABLE OF CONTENTS

Part I

Psychometric Properties of Chinese Cerebral Palsy Quality of Life for Children (C-CP QOL-Child)

CHAPTER 1

INTRODUCTION

Cerebral palsy (CP) is a common physical disability in children. The quality of life (QOL) of children with CP seems to be significantly lower than that of children without CP (Chen et al., 2011; Davis et al., 2009; Wang et al., 2010); yet, studies on evaluating QOL for children with CP were limited and some existing scales have been found to have substantial limitations, i.e., many of them measure the absence of health difficulties or limitations rather than well-being, although health is only a domain within the QOL construct. Moreover, those scales focus on physical health and functioning of children with CP without paying attention to other potentially important domains such as supportive physical environment, family health, and acceptance in the family and community (Arnaud et al., 2008; Carlon et al., 2010; Chen et al., 2011; Davis et al., 2009; Davis, Mackinnon, & Waters, 2011; Davis, Shelly, Waters, & Davern, 2010; Wang et al., 2010). The urge for developing a relative robust QOL measure is crucial, since it can be used to assess the effectiveness of treatment interventions on improving well-being for children with CP.

1.1. Structure overview

In order to develop a well-structured QOL measure, in this study, the structure of the Cerebral Palsy Quality of Life for Children (CP QOL-Child) scale was investigated through Rasch analysis, as for the clinical implementation purposes, an on-site computer program was then developed. Thus, this thesis has twofold, and can be divided into three parts: the first session, the first and second part is about the hand-on analysis of the Taiwan data. The purpose of the first part is to investigate the structure and the nature itself of the CP QOL-Child scale, unidimensional Rasch models were applied for item calibrations; the second part is a follow-up study of Part I, advanced models, i.e., multidimensional Rasch models, were applied for further analysis of the data, and based on the results, the scale was modified accordingly; the third part is an exploratory session, which is about the theories and applications of Computerized Adaptive Testing (CAT). Via investigating the practical operations, we attempted to develop an on-site CP QOL-Child CAT program for the clinical research prospective; but part III was not a completed study yet, which only included in the appendix for references in this stage, following up actions should be taken in the near future.

1.2. Outcome measure selection via a systematic review

Carlon, Shields, Yong, Gilmore, Sakzewski, and Boyd (2010) reviewed 776 papers from the 8 electronic databases [AMED (1985-October 2009), CINAHL (1982-October 2009), Cochrane Library (1966-October 2009), EMBASE (1988-October 2009), ERIC (1966-October 2009), Medline (1950-October 2009), PEDro (1929-October 2009), and PsychINFO (1967-October 2009)], and finally identified 5 common QOL measures for children with CP (see Table 1.1) (Carlon, et al., 2010). The definitional criteria were based on the scale construction, standardization, reliability, validity, responsiveness, clinical and overall utility (Carlon, et al., 2010). CP QOL-Child was the only measure which wholly fulfilled the definitional criteria of QOL, and was found to be one of the strongest measures of QOL in children with CP after taking into account the evidence for reliability and validity, the ease of access, the relatively quick completion time, and the free availability (Carlon, et al., 2010; Davis, et al., 2011; Wang, et al., 2010; Waters et al., 2007; Waters, Maher, Salmon, Reddihough, & Boyd, 2005).

Table 1.1

Five common QOL measures for children with CP

Measure/Scale	Description	Cost	Target Group (Age; year)	Completion Time (minutes)
Cerebral Palsy Quality of Life for Children (CP QOL-Child)	A condition specific outcome measure intended to gauge and assess QOL changes in children with cerebral palsy	No cost for questionnaire or manual. Registration requested.	Caregiver/ Parent (4-12) Child (9-12)	15-25
Caregiver priorities and child health index of life with disabilities (CPCHILD)	A measure of health status and wellbeing for children with severe cerebral palsy, developed to measure effectiveness of interventions	No cost for questionnaire or manual. Registration requested.	Caregiver/ Parent (5-12)	20-30
Care & comfort hypertonicity questionnaire (C&CHQ)	Designed to capture QOL improvements as a result of reduced tone post intrathecal baclofen (ITB) treatment	No cost for use of questionnaire. No manual available.	Caregiver/ Parent	-
DISABKIDS	Intended to measure HRQOL and assess burden of disease in children and adolescents	Non-funded, Government funded or non-commercial organization research: free Commercial studies: 1000€ - 5000€	Caregiver/ Parent (4-16) Child (4-16)	-
PedsQL 3.0 CP Module	Developed to assess HRQOL in a population of children and	Non-funded: free Funded academic: $600US per study Non-Commercial:	Caregiver/ Parent (2-18) Adolescent/	5

	adolescents with cerebral palsy	$1600US Commercial $5600-$20,600US	Child (5-18)

1.3. Chinese version of CP QOL-Child

Wang, Cheng, Hung, Ju, Lin, and Lo (2010) translated and revised the instrument and developed the Chinese version CP QOL-Child. The C-CP QOL-Child scale was then administered to 189 participants; and the reliability and validity were preliminary established using Classical Test Theory models (Wang, et al., 2010). The purpose of this study is to test more comprehensively the psychometric properties of C-CP QOL-Child and to validate the instrument for use in Chinese populations using Rasch models on the data already collected by the research team.

It has become more common in scale development and evaluation that CTT analyses are complemented using another approach, namely, Item Response Theory (IRT) such as the Rasch Analysis. This study proposes to adopt Rasch Analysis to test and evaluate the construction of C-CP QOL-Child scale by looking at patterns of item responses and drawing probabilistic inferences. More specifically, we will calculate the infit and outfit statistics of the items to assess the goodness of fit. We will also perform item map analysis to examine item difficulty against person ability. Principal component analysis will be carried out on

residuals to assess whether it is appropriate to combine the scores of the items into a single/ total score for interpretation purposes. The ultimate goal of this study is to establish solid psychometric properties of the Chinese version CP-QOL-Child scale for use in Chinese speaking communities.

1.4. Importance of the study

Many researchers have attempted to examine the wellbeing of children with CP (Bagley et al., 2007; Chen, et al., 2011; Davis, et al., 2009; Davis, et al., 2010; Dieruf et al., 2009; Houlihan, O'Donnell, Conaway, & Stevenson, 2004; Hoving et al., 2009; World Health Organization, 1993). However, there were very few suitable measures of QOL with excellent psychometric properties for children with CP until CP QOL-Child was developed. The problem is more serious for Chinese speaking populations as C-CP QOL-Child is the only measurement tool available for this purpose (Wang, et al., 2010). The significance of this study and the potential impact it can make are obvious, as more and more researchers and clinicians in Chinese communities are using QOL as a main outcome measure in assessing the effectiveness of different types of intervention strategies on children with CP.

1.5. Research questions

Since Rasch analysis provides a scaling methodology that enables the examination of the construction of true interval-scale, construct validation and also to evaluate whether the responses conformed to what would be expected, by looking at patterns of item responses and drawing probabilistic inferences, in this study, four questions will be address:

(1) do items separate persons appropriately and provide an acceptable measure quality?

(2) do items exhibit an appropriate goodness of fit within each subscale?

(3) do items contribute to a unidimensional construct within each subscale as well as the full scale?

(4) does child self report correlate to parent proxy report?

CHAPTER 2

LITERATURE REVIEW

2.1. Definition of Cerebral Palsy

Cerebral palsy (CP) is a clinical descriptive term comprehending a group of unprogressive, noninfectious motor conditions that cause physical disability in human development, predominantly by the various areas of body movement (Stanley & Blair, 2000; Swiggum, Hamilton, Gleeson, & Roddey, 2010; Tecklin, 2008). Cerebral adverts to the cerebrum, which is the affected region of the brain (though the disorder essentially the most comprises connections between the cortex and other parts of the brain such as the cerebellum), and palsy adverts to disorder of apparent movement (Levitt, 2010; Odding, Roebroeck, & Stam, 2006; Stanley & Blair, 2000).

CP is also the commonly used name for a group of clinical syndromes characterized by motor deficits due to non-progressive brain damage early in life (Odding, et al., 2006; Stanley & Blair, 2000; Tecklin, 2008); its incidence has been approximated as 2-2.5 per thousand live births (Stanley & Blair, 2000). Treatment of disquiets affiliated with these motor deficits may

demand a range of interferences such as physiotherapy, orthopedical surgery and medicaments for spasticity (Amichai, Harries, Dvir, Patish, & Copeliovitch, 2009; Levitt, 2010; Tecklin, 2008).

2.2. Definition of Quality of Life

The term Quality of Life (QOL) is employed to assess the comprehensive wellbeing of individuals and societies (Davis, et al., 2010; Dieruf, et al., 2009; World Health Organization, 1993); QOL admits not solely healthiness and employment, but the constructed environment, physiological and psychological wellness, education, leisure and recreation, and social belonging as well. QOL constitutes an of import conception in the area of global exploitation, for it permits exploitation to be dissected on a measure more spacious than standard of life (Arnaud, et al., 2008; Chen, et al., 2011; Davis, et al., 2010; Dieruf, et al., 2009; Waters, et al., 2005). Recently, QOL also helps to evaluate students' outcome, performance, or effectiveness in education sectors.

Researchers have been employing evidence-based outcome measures to channelize the interpreting of health status and effectiveness of intervention for children with CP (Bagley, et al., 2007; Carlon, et al., 2010; Streiner & Norman, 2008; Waclawski, 2010; Waters, et al.,

2005). One of the most commonly adopted models from the previous studies, a determinant health model, the International Classification of Function, Disability and Health (ICFDH), which has been fabricated for delineating the encroachment of health status on a measure of lifespan spheres admitting QOL (Bagley, et al., 2007; Levitt, 2010; Wake, Ba, & Reddihough, 2003; Waters, et al., 2007). Thence QOL has attained raising attention as a crucial health condition or intervention outcome measure. QOL has been characterized by the World Health Organization as individuals' perceptual experience of their perspective in everyday life, in the context of the culture, ethic and moral principle in which they live as well as in regard to their goals, expectations, standards and concerns (World Health Organization, 1993).

2.3. CP QOL-Child and Chinese version CP QOL-Child

While measuring QOL is an essential portion of evaluating the health status of children with CP, few scales that have been utilized in the previous have been encountered to have substantial limitations (Waters, et al., 2005). In recognizing such limitations, a measure, the Cerebral Palsy Quality of Life for Children (CP QOL-Child) was developed by a group of researchers and clinicians in consultation with children, caregivers and health professionals in recent years (Wang, et al., 2010; Waters, et al., 2007; Waters, et al., 2005). It is a condition-specific QOL cadence for children with CP in which can be accustomed to assess

self perceived well-being of several life domains of such children. It includes two parallel forms, the parent proxy form and the child self-report form. Former researchers have reported the parent proxy form appears to have identically great reliability and validity (Arnaud, et al., 2008; Davis, et al., 2011; Wang, et al., 2010). As well, the child self-report form has demonstrated satisfactory psychometric properties (Waters, et al., 2007).

As the original version of CP QOL-Child is a well-developed measurement scale, the Chinese version of this scale (Chinese CP QOL-Child) was constructed in subsequent times. In the study of Wang, Cheng, Hung, Ju, Lin, and Lo (2010), the methods of the Classical Test Theory (CTT) analyses were conducted for this Chinese version, and its reliability and validity appears to be acceptable (Wang, et al., 2010). Although support has been found for the psychometric properties of the Chinese CP QOL-Child using CTT approaches, two major conceptual limitations have been pointed out: the lack of an explicit ordered continuum of the test items that represent a unidimensional construct, and the lack of justification of rating scale data. In contrast to the CTT approach, the item response theory (IRT)-based Rasch analysis has gained as a potent application of psychometric testing for the development of new instruments and the refinement of existing ones (Avery, Russell, Raina, Walter, & Rosenbaum, 2003; Chien, Hsu, Tai, Guo, & Su, 2008; de Morton, Keating, & Davidson, 2008; Dougherty, Nichols, & Nichols, 2011; Prieto, Alonso, & Lamarca, 2003; Smith et al.,

2006; Tang, Wong, Chiu, & Ungvari, 2007). Rasch analysis provides a scaling methodology that enables the examination of the construction of true interval-scale, construct validation and also to evaluate whether the responses conformed to what would be expected, by looking at patterns of item responses and drawing probabilistic inferences (Bond & Fox, 2007; Dougherty, et al., 2011; Fischer & Frewer, 2009). Details of the measurement theories will be reviewed in the following sections.

2.4. Measurement Theory

In the fields of psychology and education, many variables of interest (e.g., social abilities, emotional health) cannot be observed directly. Latent variables (i.e., traits, dispositions, abilities) are measured from a sample of behaviors via questionnaires, rating scales, observation measures, responses to simulated problems, or test questions (Streiner & Norman, 2008; Waclawski, 2010; Wright, 1996). All of these measurements share the problem of being ordinal measures. This limits the type of analysis that can be done (Bond & Fox, 2007). For many years, classical test theory was applied to these types of measures in order to classify individuals meaningfully. This approach did not take into account the problem of the ordinality of the metric. More recent theories of measurement utilize latent trait or item response models (Bond & Fox, 2007; Prieto, et al., 2003). In this section, we are going to

briefly review the classical test theory, then discuss some of the item response models in advance.

2.5. Classical Test Theory

Classical test theory is a body of related psychometric theory that predicts outcomes of psychological testing such as the difficulty of items or the ability of test-takers. Generally speaking, the aim of classical test theory is to understand and improve the reliability of psychological tests.

Classical test theory may be regarded as roughly synonymous with true score theory. The term “classical” refers not only to the chronology of these models but also contrasts with the more recent psychometric theories, generally referred to collectively as item response theory, which sometimes bear the appellation “modern” as in “modern latent trait theory”.

Classical test theory assumes that each person has a true score, which would be obtained if there were no errors in measurement. A person's true score is defined as the expected number-correct score over an infinite number of independent administrations of the test.

Unfortunately, test users never observe a person's true score, only an observed score. It is assumed that observed score = true score plus some error.

Benefits obtainable through the application of classical test theory include:

1. Smaller sample sizes required for analyses (a particularly valuable advantage for field testing).

2. Simpler mathematical analyses compared to item response theory.

3. Model parameter estimation is conceptually straightforward.

4. Analyses do not require strict goodness-of-fit studies to ensure a good fit of model to the test data.

Until recently, all social scales were developed using only classical test analysis. Classical test theory has some serious limitations. It assumes that errors are normally distributed among persons with constant variance and have an expected value of zero (Bond & Fox, 2007; Prieto, et al., 2003). In classical test theory the observed score is equal to the true score (the latent variable) plus error. In the application of classical test theory, the standard error of measurement is provided for the particular population rather than individual scores (Prieto, et al., 2003). The score obtained on a test applies only to that test or to items on a parallel form with equivalent item properties. The item difficulties and discrimination are omitted from the

model and are justified by their impact on various group statistics (variances and reliabilities) and their relationship to other measures (Bond & Fox, 2007; Mitchell et al., 2011; Prieto, et al., 2003).

The indicators of item difficulty and discrimination in classical test theory (i.e., p values and biserial correlations) are influenced greatly by the sample distributions (Prieto, et al., 2003). The estimate of item difficulty is the proportion of correct response (p-value) made by the individuals who were in the sample. If an item is administered to a group that is talented in the area being assessed, the p-value might be very high. If the sample is less able, the p-value is lower.

In a similar manner, the level of the trait assigned to any given person (i.e., person ability) is dependent upon the difficulty of the sample of items included in the assessment. Person ability in classical test theory is based on the proportion of items answered correctly (or endorsed) by the individual (Bond & Fox, 2007; Prieto, et al., 2003). If the sample of items is relatively easy, the person ability will appear high on the trait. A third grader taking a first grade reading test would receive ability scores on that test that would be quite high, even if their reading ability in third grade books was low (Bond & Fox, 2007). Classical test theory tried to mitigate this problem by developing norms for different samples (e.g. by grade or

age). Tests would include items targeted for a particular age or grade and individuals would receive a score based on the specific normative group to which they were being compared.

For most measures of social skills, items are usually factor analyzed and scale scores are constructed based on the sum (or the mean) of the ratings given to each individual item. The sum is then converted to a standard score based on the normative sample for that instrument. Factor analysis in classical test theory assumes normally distributed, interval, or ratio variables (Wright, 1996). These assumptions are not met in the measurement of social and emotional traits. The observations or ratings on measures are ordinal or sometimes nominal data, thus violating the assumptions of linearity and normality (Bond & Fox, 2007; Linacre, 2002). These data appear to give meaningful results because, when data are complete, the relationship between the raw scores and the measure is ogival, and toward the center of the ogive (cumulative frequency distribution resembling the arch of that name), the relationship between score and measure is approximately linear. However, this relationship breaks down at the extreme scores, with extreme scores, the center of the ogive might not be identical, therefore we cannot assume that the measure is linear (Bond & Fox, 2007; Linacre, 2002; Prieto, et al., 2003; Wright, 1996).

In classical test theory, the standard error of measurement is applied to all scores in a particular population. However, standard error varies across populations and is usually much higher for individuals at the extremes than for those scoring at the mean. For example, many programs designed to increase social skills and prevent violence and mental illnesses are targeted toward children who are at the extreme scores on social and behavioral measures, however, since the scale itself cannot appropriately separate those children with extreme scores, it results inability to measure changes accurately among these children, and inhibits the ability to track the development of skills, then to evaluate the success of programs intended to increase skills cannot be achieved (Bond & Fox, 2007; Wright & Stone, 1979).

Limitations of classical test theory include:

1. Item difficulty and item discrimination are group/sample dependent

2. Scores obtained by CTT applications are tests dependent.

2.6. Item Response Theory

Item Response Theory (IRT) is widely used in test development. Though previously applied to the measurement of academic and cognitive ability, in recent years, item IRT (sometimes called latent trait models) has been applied to measures of personality, affect, and behavior.

IRT typically uses maximum likelihood estimation to estimate the probability of a correct response on an item based on the ability or trait level of a particular person and the characteristics of a particular item (Bond & Fox, 2007; Burro, Sartori, & Vidotto, 2011; Chien, et al., 2008; Pallant & Tennant, 2007; Tang, et al., 2007). Unlike the Classical Test Theory, IRT assumes a non-linear model with Bernoulli, multinomial, or Poisson sampling (Linacre, 2002; Wright & Mok, 2000; Wright & Stone, 1979). This allows the use of the type of data collected when studying social behaviors. IRT models provide item characteristic curves (ICCs) that "describe how the probability of responding to an item in a specific way changes as a function of the examinee's position on a latent trait variable" (Bond & Fox, 2007). An examinee would have a 50% probability of answering correctly or endorsing an item when their ability level is the same as the difficulty level. On items above their ability level they would have a decreased likelihood of responding correctly or endorsing those items. On items below their ability level, they would have an increased likelihood of responding correctly or endorsing that item (Bond & Fox, 2007; Linacre, 2002; Wright & Stone, 1979).

In Rasch models the log odds of the probability of a correct response is a function of the difference between the person's ability or the person's level of the trait and the difficulty of the item. In the simplest dichotomous model, this is expressed as

$$\log\left(\frac{P_{ni}}{1-P_{ni}}\right) = B_n - D_i$$

in which P_{ni} is the probability of person n correctly answering or endorsing item i, B_n is the measure of person n, and D_i is the difficulty of answering or endorsing item i, and log is the natural logarithm. The item discrimination is held constant across the items (Andrich, 1978; Wright & Stone, 1979).

Applying the Rasch model to the data allows us to construct a system of invariant linear measures, estimate the accuracy of the measures (standard errors), and determine the degree to which these measures and their errors are confirmed in the data using the fit statistics (Bond & Fox, 2007; Linacre, 2002; Wright & Stone, 1979).

The Rasch model assumes unidimensionality, that is, a single dimension is being measured (Bond & Fox, 2007). We can use Rasch models to determine what is measurable on a linear scale, which data are useful in describing the latent trait and which are not, how the respondents used the categories in the measure, and whether different groups of respondents utilized the categories of the measures in different ways. The major advantages of the Rasch

model are the item and sample invariance properties and the interval measurement scale (Avery, et al., 2003; Bond & Fox, 2007; Linacre, 2002; Wright & Stone, 1979).

2.7. Sample size issue in Rasch model

Unlike other IRT models, the Rasch model meets the standard of "sufficient statistics" (provide more information on item level) (Wright & Masters, 1982). Therefore, the model can estimated with small sample sizes. For example, by randomly selecting varied sample sizes from a data bank of responses on an achievement test by 1478 fourth graders and 1808 eighth graders, Forster (1976) determined that only 150-200 students were required in order to obtain stable item difficulty estimates (Forster, 1976). Pearson Product Moment correlations between the item difficulty estimates for different samples were used as the stability criterion and a mean of 0.9797 was obtained with sample sizes in the range of 150 to 200 using a Rasch model. Embretson (1999) and Weiss (1983) also obtained parameter stability with small samples (Embretson & Hershberger, 1999; Weiss & Bock, 1983).

2.8. Rasch models

Rasch measurement includes models that can be used with a variety of types of data. Dichotomous models, binomial trials, Poisson counts, partial credit, rating scale and ranks are commonly encountered Rasch models (Wright & Mok, 2000). A dichotomous model is used with data that has only two categories and single attempts to each item. Data that have only two categories but include multiple attempts are more suited to the Poisson counts or binomial trials model. If there is an upper limit to the number of attempts, the binomial trials model is the appropriate choice (Linacre, 2002; Wright & Mok, 2000; Wright & Stone, 1979).

When three or more categories are present, the rating scale, partial credit, or ranks model is applied. If the step difficulties (between categories) are assumed to be the same across items (e.g., the distance from 1 to 2 is the same for all items), then the rating scale model may be applied. The partial credit and Rank models do not require that the step difficulties are the same. As suggested the name, the Rank model is used when the observations are ranks (Wright & Mok, 2000). Computer programs (e.g., Winsteps, ConQuest) applying the Rasch models typically report person ability (or trait level) and item difficulty (or endorsability) in logits.

The logit scale includes negative scores that are counter intuitive to many people. Applying a linear transformation to the logit metric allows for greater interpretability of the scale while maintaining the interval scale. One suggested scale transformation is based on the least measureable distance. On a rating scale, the least measurable difference would be one more category on the rating scale (an increase of 1 in the sum of the raw scores on the scale). Wright and Stone (1979) recommend an average value for the least measurable distance would be 6 divided by the maximum score on the rating scale. This average value is an approximation (Wright & Stone, 1979). The least measurable distance is dependent upon the width of the distribution of item difficulties. Popular transformations of the logit scale include normative transformations (often assigned a mean of 100 and a standard deviation of 15), and the use one or more substantive points on the scale (e.g., cut-scores) to determine the mean and standard deviation of the transformed scale (Wright & Masters, 1982; Wright & Mok, 2000; Wright & Stone, 1979).

2.9. Evaluating Rasch models

Rasch programs produce a variety of statistics that allow one to examine a measure's reliability and validity. The root mean square standard error (RMSE) assesses the fit of the data to the model. The RMSE is computed over both persons and items. Both a Model RMSE

and Real RMSE are provided. The Model RMSE assumes that misfit in the data is due to the probabilistic nature of the model and reports the upper limit to the reliability of this set of items with this sample. The Real RMSE assumes that misfit in the data is due to inability of the data to meet the model specifications and is inflated for misfit when the infit statistic is greater than one [Real RMSE = Model RMSE * square root (infit mean square)]. This represents the "worst case" reliability or lower limit to the reliability (Wright & Masters, 1982; Wright & Mok, 2000).

Root Mean Square Standard Error for items is

$$\mathrm{RMSE_i} = \left(\frac{\sum_{i=1}^{L} s_i^2}{L} \right)^{-2}$$

where s_i is the calibration error for each item and L is the number of items (for additional explanation of estimation procedures and standard errors) (Wright & Masters, 1982).

Fit statistics for the overall model are presented in summary statistics. Both an information-weighted (infit) and an outlier sensitive (outfit) statistic are provided. The formulas for the fit statistics are:

Infit MNSQ

$$v_i = \frac{\sum_{n=1}^{N} z_{ni}^2 W_{ni}}{\sum_{n=1}^{N} W_{ni}}$$

Outfit MNSQ

$$u_i = \frac{\sum_{n=1}^{N} z_{ni}^2}{N}$$

where z_{ni} is the standardized item residual for person n on item i and W_{ni} is the variance of the residual. The residual is the difference between the estimated and actual response to an item (Wright & Masters, 1982).

The infit mean square (MNSQ) is sensitive to unexpected responses close to the person's trait level. The outfit mean square is more sensitive to unexpected response on items far from the person's trait level. The expected value for the mean square is 1.0. Values that are much higher than 1.0 indicate that there is noise in your data, that is "variation in the responses of persons to items that significantly deviate from general patterns" (Wright & Mok, 2000).

The Rasch person separation reliability statistic is directly analogous to the KR-20, Cronbach's Alpha, and the Generalizability Coefficient (Wright & Mok, 2000). It measures the internal consistency of each scale. Unlike Cronbach's Alpha, because the person separation reliability is computed both with and without extreme persons, it is not inflated by extreme scores (Linacre, 2002; Wright & Mok, 2000). Person separation is the ratio of the adjusted standard deviation to the root mean standard error

$$\frac{((\text{S. D. of Measure})^2 - (\text{RMSE})^2)}{\text{RMSE}}$$

Person separation reliability is the square of the separation statistic divided by one plus the separation squared (Wright & Mok, 2000).

Item difficulty and infit mean square of items are also particularly helpful in evaluating the reliability and validity of a scale. The item difficulty estimates the likelihood that an individual will rate highly on the skills or behaviors represented by each scale item. A map illustrating the ordering of item difficulties is produced. This ordering should be consistent with theoretical definitions of the trait (Andrich, 1978; Linacre, 2002; Wright & Masters, 1982). For example, more severe aggressive behaviors should be harder to endorse than less severe aggressive behaviors. This would be a measure of the face validity of the scale.

The infit mean square statistic indicates the extent to which the responses to an item are consistent with the hierarchical position of that item in the scale. Infit mean square statistics that are close to 1.0 indicate that individuals are responding to the item in a way that is consistent with the item's location in the scale (Embretson & Hershberger, 1999; Wright & Masters, 1982; Wright & Mok, 2000). For example, if a child is given a high rating on a particular item, he should receive high ratings on the easier to endorse items below it in the scale, but not necessarily receive positive ratings on the more difficult to endorse items above it. Thus, the infit mean square statistic captures the extent to which response patterns are consistent with the difficulty-based rank-ordering of the item obtained in a Rasch analysis. Infiit mean square statistics that are far from 1.0 indicate that there is something other than the trait influencing the responses on that item. It may be that the wording of the item is not understood in the same way by all respondents, or that another trait is influencing the responses on a given item. A high level of misfit means that the item is contributing poorly to the measurement of the defined trait (Bond & Fox, 2007; Linacre, 2002; Wright & Masters, 1982; Wright & Mok, 2000). Taken together, the reliability estimates, item difficulties, and fit statistics provide evidence for evaluating the validity and reliability of measures produced by the Rasch model.

2.10. Applications of Rasch Modeling

Rasch modeling has been used to examine the reliability and validity of different measures of psychometric characteristics (Fischer & Frewer, 2009; Gothwal, Wright, Lamoureux, & Pesudovs, 2009; Palisano et al., 1997), to test the dimensionality (Cheng, Wang, & Ho, 2009; Smith, et al., 2006; Wang, Chen, & Cheng, 2004; Wright, 1996; Yao & Boughton, 2007), to evaluate response scales (Burro, et al., 2011; de Morton, et al., 2008; Dougherty, et al., 2011; Gothwal, et al., 2009; Tang, et al., 2007), and to examine gender or group differences in manifestations of psychometric traits (Avery, et al., 2003; Damsg, rd, Fors, Anke, & e, 2008; Gothwal, et al., 2009). Clinical applications of Rasch models have also been discussed in the recent literatures, i.e., Gross motor function, sleep disorder, low back pain, vision function, etc (Avery, et al., 2003; Burro, et al., 2011; Chien, et al., 2008; Damsg, et al., 2008; de Morton, et al., 2008; Gothwal, et al., 2009; Mitchell, et al., 2011; Pallant & Tennant, 2007; Ryser, Wright, Aeschlimann, Mariacher-Gehler, & Stucki, 1999; Smith, et al., 2006; Tang, et al., 2007; Tennant, McKenna, & Hagell, 2004).

The Rasch model uses responses on each of the items to estimate the probability that items will be endorsed by respondents with different levels of the trait. As noted earlier, if the items are measuring the construct reliably, the hierarchy of item endorsability should be consistent

with theoretical assumptions about the construct. Inconsistency between theoretical suppositions about the construct and the ordering of the items may indicate a need to refine the theory about the trait, or alternatively, may indicate that the items are understood by respondents in ways that are different than those assumed by the item developers. Further examination of the fit of the items to the model and examination of a factor analysis of the residuals helps elucidate whether the items are unidimensional and whether they are understood by the respondents in ways that are similar across respondents.

2.11. Assessing dimensionality

Linacre (1998) contends that "Rasch analysis followed by a factor analysis of residuals was always more effective at both constructing measures and identifying multidimensionality than direct factor analysis of the original response-level data" (Linacre, 1998, p. 282). Additional dimensions in the data may be due to differences in response styles or the presence of more than one trait in the data. When the dimensions are correlated for the majority of the sample, a factor analysis of the original response level data will not reveal the fact that there is more than one dimension in the data. If these dimensions differ in their relationship to other variables, treating them as a single dimension could obscure relationships with other variables. Bond and Fox (2007) note that high inter-correlations among subscales are a

necessary but not sufficient condition to assert unidimensionality, it is also essential that the subscales relate identically or, at least, very similarly, to theoretically linked covariates (Bond & Fox, 2007).

2.12. Response scale categories

Rasch models allow researchers to examine the use of the rating scale categories (Andrich, 1978; Bond & Fox, 2007; Linacre, 2002; Wright & Masters, 1982; Wright & Mok, 2000). Rasch models can be used to identify the optimal rating scale categories, identify persons or items that misfit, and provide a means of measuring response style. Measures of constructs in the social and emotional domains often use Likert-type scales (Bond & Fox, 2007). For example, a five point scale from strongly disagree to strongly agree is frequently used as a response scale. For respondents, the psychological distance from disagree to neutral may be greater than then distance from neutral to agree and it may be even easier to move from agree to strongly agree. Rasch models estimate the distance between categories or steps. Fit statistics on the step calibrations are provided by computer programs that apply the Rasch model. Examination of the item and person separation and fit statistics under alternative scoring allows researchers to investigate how the scales were interpreted by the respondents and to adopt the most reliable response scale.

2.13. Response style

Respondents may have different styles of responding. When individuals do not hold strong opinions about an issue, there is a tendency to answer with more agreeable response (Bond & Fox, 2007; Wright & Mok, 2000). Rasch models can be used to create measures of acquiescence. Response styles may vary according to group membership and differences in response styles can be investigated through the creation of measures of acquiescence. Measures of response style help in differentiating category utilization styles from response to content (Andrich, 1978; Bond & Fox, 2007; Wright & Masters, 1982; Wright & Mok, 2000).

Rasch is a typical form of IRT models, IRT models are a family of probabilistic test models that directly estimate and calibrate item and person ability parameters from test response data (Bond & Fox, 2007). The Rasch model uses a single statistical formula that integrates the item difficulty and person ability estimates, test scores, and scale calibrations (Wright & Mok, 2000). The Rasch model assumes that the construct under investigation is unidimensional and that the distribution of the item difficulty and person ability estimates approximates a logistic distribution (Bond & Fox, 2007; Wright & Mok, 2000).

Rasch models can be used to develop linear interval scales that measure change. The main features of Rasch models include the following: 1) difficulty and ability are estimated using data from the entire matrix. When the data meet the requirements of the model, the estimates are not as dependent upon the sample of items or the sample of individuals as they are in Classical Test Theory; 2) Item difficulty and person abilities are placed on the same scale. Item difficulty and ability (trait level) are expressed as log odds. A person with ability equivalent to item difficulty will have a 50% chance of endorsing that particular item (or getting it correct in the case of a test); 3) Person abilities (trait levels) can be estimated (and interpreted) even when data are incomplete; 4) Rasch models assess the undimensionality of items. Misfitting items are identified. Misfitting items are items do not fit a linear interpretation of the construct. These items are influenced strongly enough by other factors that an increase in endorsement of that item does not clearly indicate an increase in the ability level of the latent trait; 5) Rasch models provide a linear interval scale when the assumptions of the model are met. Standard errors are provided for each score.

With the increased interpretability due to the invariance of item structure and the linear interval scale, Rasch is the recommended measurement model for analysis of the psychological scale.

CHAPTER 3

METHODOLOGY

3.1. Sample

A convenience sample of children with CP and their primary caregivers was recruited from rehabilitation departments, early intervention centers, and special education schools in Taiwan.

The data collection process involved two phases:

Phase I: child and caregiver subjects were recruited from South Taiwan in 2009 (Table 3.1);

Phase II: child subjects were recruited from North Taiwan in 2011 (Table 3.2).

Five hospitals, three rehabilitation clinics, three early intervention centers and two special education schools were participated finally in this study.

There were two inclusion criteria for children:

(1) children aged 4 to 12 years, and diagnosed with CP by a pediatric neurologist; and

(2) children capable of providing self-report or could understand the explanations of the questionnaire contents.

Children suffering from neurodegenerative diseases or psychiatric illness were excluded.

The inclusion criterion for primary caregiver was:

(1) capable of completing the questionnaires without any assistance.

The responses of 145 caregivers and 44 children with CP in phase I, and 87 children with CP in phase II were used for analysis.

Demographic data collected on primary caregivers included age, gender, marital status, educational level, employment status, as well as number (and relationship) of family members living with the child. Data collected on the children included age, gender, and intensity of rehabilitation intervention (times per week). Ethics approval was given by Kaohsiung Medical University; informed consent was obtained prior to data collection.

Table 3.3 shows the demographic characteristics (combined Phase I and II) of all the participants.

Table 3.1

Demographic Details: Phase I (in the year2009)

Primary caregiver (n=145)		Their child with CP (n=44)	
Demographics	*Frequency (%)*	*Demographics*	*Frequency (%)*
Age, year: mean (range)	39.2(26-63)	Age, year: mean (range)	8.1(4-12)
Relationship to the child		Gender	
Father	23(15.9)	Male	84(57.9)
Mother	112(77.2)	Female	61(42.1)
Grandmother	4(2.8)	GMFCS[a] levels	
Babysitter	6(4.1)	I	25(17.2)
Education		II	30(20.7)
Primary school	6(4.1)	III	31(21.4)
Junior high school	14(9.7)	IV	31(21.4)
Senior high school	61(42.1)	V	28(19.3)
Graduate	62(42.7)	Neuromotor classification	
Graduate school	2(1.4)	Hemiplegia	26(17.9)
Marital statue		Spastic diplegia	49(33.8)
Married	131(90.3)	Spastic quadriplegia	42(29.0)
Divorce	11(7.6)	Athetoid	18(12.4)
Unmarried	2(1.4)	Others	10(6.9)
Spouse death	1(0.7)	Sibling live with: mean (range)	1.1(0-4)
Family revenue/expenditure			
Abundance	9(6.2)		
Balance	106(73.1)		
Insufficiency	30(20.7)		

[a] Gross Motor Function Classification System

Table 3.2

Demographic Details: Phase II (in the year2011)

Child with CP (n=87)	
Demographics	*Frequency (%)*
Age, year: mean (range)	10.8(8-12)
Gender	
Male	33(37.5)
Female	54(62.5)
GMFCS[a] levels	
I	16(18.8)
II	25(28.2)
III	24(27.2)
IV	15(17.6)
V	7(8.2)
Neuromotor classification	
Hemiplegia	13(14.9)
Spastic diplegia	24(27.7)
Spastic quadriplegia	39(44.8)
Athetoid	6(6.8)
Others	5(5.8)

[a] Gross Motor Function Classification System

Table 3.3

Demographic Statistics

Primary caregiver (n=145)		Child with CP (n=131)	
Demographics	*Frequency (%)*	*Demographics*	*Frequency (%)*
Age, year: mean (range)	39.2(26-63)	Age, year: mean (range)	9.9(4-12)
Gender		Gender	
Male	23 (15.9)	Male	49(37.5)
Female	122 (84.1)	Female	82(62.5)
Relationship to the child		GMFCS[a] levels	
Father	23(15.9)	I	24(18.3)
Mother	112(77.2)	II	34(26.0)
Grandmother	4(2.8)	III	33(25.2)
Babysitter	6(4.1)	IV	24(18.3)
Education		V	16(12.2)
Primary school	6(4.1)	Neuromotor classification	
Junior high school	14(9.7)	Hemiplegia	21(16.0)
Senior high school	61(42.1)	Spastic diplegia	40(30.5)
Graduate	62(42.7)	Spastic quadriplegia	50(38.2)
Graduate school	2(1.4)	Athetoid	12(9.2)
Marital statue		Others	8(6.1)
Married	131(90.3)		
Divorce	11(7.6)		
Unmarried	2(1.4)		
Spouse death	1(0.7)		
Family revenue/expenditure			
Abundance	9(6.2)		
Balance	106(73.1)		
Insufficiency	30(20.7)		

[a] Gross Motor Function Classification System

3.2. Measure

The primary caregiver proxy form of the CP QOL-Child was developed for primary caregivers of children with CP aged 4–12 years (65 items).The child self-report form was for children with CP aged 9–12 years (52 items).

The instrument measures seven domains/subscales of QOL for a child with CP:

(1) social well-being and acceptance;

(2) functioning;

(3) participation and physical health;

(4) emotional well-being; and

(5) pain and impact of disability.

The other two test domains are only included on the primary caregiver form, namely,

(6) access to services and

(7) family health.

Table 3.4 shows the subscales and its contributing items

Table 3.4

The subscales and items of CP QOL-Child

Subscale	Item of child self-report form	Item of parent proxy form
Swb	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15
Fun	8, 22, 23, 24, 27, 29, 34, 35, 37, 38, 39, 40	8, 22, 23, 24, 27, 29, 34, 35, 37, 38, 39, 40
Part	9, 16, 17, 18, 19, 20, 21, 25, 26, 30, 36	9, 16, 17, 18, 19, 20, 21, 25, 26, 30, 36
Emb	2, 28, 31, 32, 33, 52	2, 28, 31, 32, 33, 52
Acc	-	41, 42, 43, 53, 54, 55, 56, 57, 58, 59, 60, 61
Pain	44, 45, 46, 47, 48, 49, 50, 51	44, 45, 46, 47, 48, 49, 50, 51
Fam	-	62, 63, 64, 65

Swb : social well-being and acceptance
Fun : functioning
Part : participation and physical health
Ewb : emotional well-being
Acc : access to services
Pain : pain and impact of disability
Fam : family health

When using the CP QOL-Child to measure QOL, both the primary caregiver form and the child form are used. The stem of the test item is “How do you think your child feels about. . .?” or “How do you feel about. . .?”. This type of item stem is used because it does not measure the children’s condition or their functioning; it assesses how they feel about their condition.

The CP QOL-Child uses a Likert scale to measure happiness. All the items except one are rated on a nine-point scale ranging from 1 to 9. One item from the domain of pain and impact of disability is rated on a five-point scale.

The neuromotor pattern of movements and severity of motor disability of the children with CP were assessed by their treating physical therapist using the neuromotor classification of CP and the Gross Motor Function Classification System (GMFCS), respectively. The GMFCS levels range from Level I to Level V, with greater numerical level indicating more severe motor disability (Stanley & Blair, 2000; Tecklin, 2008).

3.3. Procedure

The CP QOL-Child was originally developed by a group of researchers and clinicians in consultation with parents, children and health professionals in Australia; please refer to the paper “Development of a condition-specific measure of quality of life for children with cerebral palsy: Empirical thematic data reported by parents and children” (Waters et al., 2005). For the purpose of adopting CP QOL-Child into Chinese context, the translation work was done by the following steps:

A six-member expert panel discussed issues related to translating both the primary caregiver form and the child self-report form of the CP QOL-Child, into Chinese. A formal permission to translate the English version into Chinese was obtained from the original authors.

The working steps for translation were:

(1) translation of each item into Chinese by two native Chinese-speaking people who were university lecturers teaching English;

(2) piloting of the translated items by inviting six parents of children with cerebral palsy to fill the preliminary Chinese CP QOL-Child;

(3) a panel discussion to revise the problematic items;

(4) exacting examination of the revised Chinese items;

(5) backward-translation of the items into English;

(6) examination of any major differences in content between the backward-translated English items and the original English items;

(7) final examination of the Chinese CP QOL-Child.

During the translation process, the researchers were mindful that the translated content should be easily understood by parents and children with CP; and that the terminologies used were appropriate cross-regionally in Mandarin-speaking communities.

We contacted parents of children aged 9-12 with CP in the rehabilitation departments of five hospitals, three rehabilitation clinics, three early intervention centers and two special education schools. Primary caregivers and their children were invited to complete the appropriate questionnaire*. Moreover, caregivers or the researcher, if necessary, were allowed to assist the child in completing the questionnaire by explaining it to them. Ethics approval was given by Kaohsiung Medical University; informed consent was obtained prior to data collection.

*The questionnaires were administrated by the health professionals, i.e., Occupational Therapists, Physical Therapists or Medical doctors; regarding to the children, most of them were associated with mental retardation.

3.4. Data analysis

Data were analyzed using Winsteps 3.70.1.1 (MSEA Press, Chicago, USA), ConQuest 2.0 (ACER Press, Camberwell, Australia) and SPSS 19 (IBM Press, USA). Before conducting any analysis, eight items in “Pain” originally designed in a negative direction were

reverse-scored in order to give all items a score in a positive direction; hence, higher scores indicating happier status or better well-being.

Variable maps and an item fit statistic tables were constructed to investigate the item-person relationship and to examine the item data fit for the full scale as well as each subscales. The variable map shows the item distribution based on its difficulty and the relative person location which implies person ability The sufficiency of the fit of each item is assessed by the goodness of fit statistics of residuals, mean-square of measures, including information-weighted fit (Infit) and outlier usual unweighted fit (Outfit), which provide information about the responses given to items around the same difficulty of endorsement level. An appropriate range for an infit and outfit value in a Likert rating scale is between 0.5 and 1.5. An infit and outfit value of less than 0.5 implicates that the item does not provide adequate information beyond that provided by the rest of items within the scale. This can occur once several items in there are similar or highly correlated or one item is dependent on another. By contrast, an infit and outfit value of greater than 1.5 implicates that the item does not define the same construct as defined by the rest of the items, it is either a poorly constructed or understood item, or is ambiguously defined (Bond & Fox, 2007; Tang, et al., 2007; Wright & Mok, 2000).

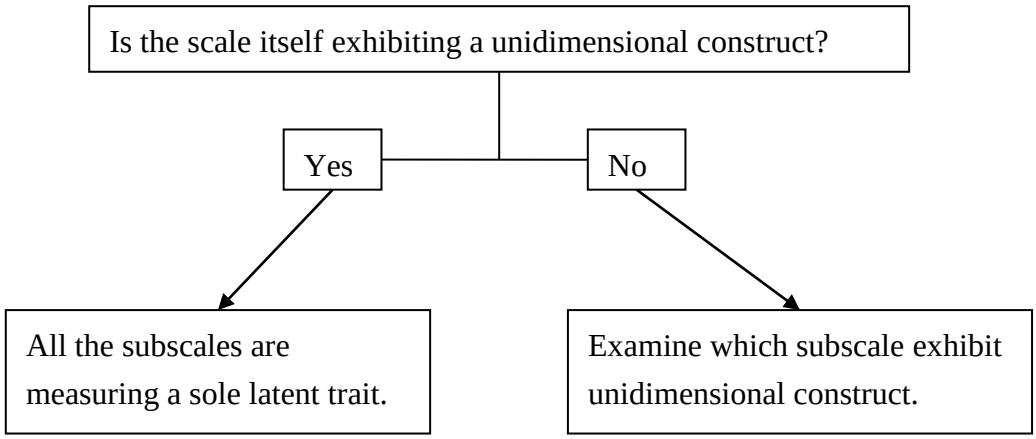
Item characteristic curves (ICCs) with Rasch ICC thresholds were figured to summarize the expected sample behavior on item level. The ICC demarcates by mean categories and relates the rating scale to the latent variable as one plotted line without explicit mention of the underlying probabilities.

A reliability test was used to estimate the reliability of person and item measures and the associated separation index for persons and items were also obtained. The person separation index indicates the ability to differentiate persons based on the measured variable, whereas the item separation index represents the ability to define a distinct hierarchy of items along the measured variable. The greater the separation, the more likely that the categorization of the items will be better separated and the differences between respondents will be better distinguished. The framework of the data analysis was shown in figure 3.1 (Page 53).

To adequately describe the data, multidimensional approach might need to be adopted in order to squeeze as much information as possible from all the data to provide measurements that are more precise.

Figure 3.1

Framework of the data analysis



CHAPTER 4

RESULTS

A total of 145 primary caregivers and 131 children with CP completed the C CP QOL-Child. Table 3.3 displays the demographic characteristics of the participants. The mean age of the caregivers was 39.2 years old; there were 122 (84.1%) females and 23 (15.9%) males. Most of them (77.2%) were the mother, followed by father (15.9%), babysitter (4.1%), and grandmother (2.8%). Sixty-two primary caregivers (42.7%) had tertiary education level, whereas 42.1% had senior high school level. The marital status of most of the caregivers (90.3%) was married. A majority (73.1%) reported that the household income and expenditure were balanced; 30 (20.7%) reported that the household income was insufficient. The mean age of the children was 9.9 years old; there were 49 (37.5%) boys and 82 (62.5%) girls. By using the GMFCS, 24 (18.3%) children were classified in level I, 34 (26%) children in level II, 33 (25.2%) children in level III, 24 (18.3%) children in level IV, and 16 (12.2%) were classified in level V. The children were also identified according to the neuromotor classification. A high proportion of the children had spastic quadriplegia (50 children, 38.2%) and spastic diplegia (40 children, 30.5%); 21 (16%) children had hemiplegia; 12 (9.2%) children had athetoid; and the other 8 (6.1%) children were identified as mixed type.

4.1. Child self report

For the child self report, after completing the calibration process, the results showed that, in item 1 (“the way you get along with people, generally?”), the item difficulty was -0.08, SE=0.07, the infit MNSQ was 0.85, SD=-1.1, the outfit MNSQ was 0.85, SD=-1.1, and the point-measure correlation was 0.56; in item 2 (“the way you get along with the person who looks after you?”), the item difficulty was -0.32, SE=0.07, the infit MNSQ was 0.89, SD=-0.6, the outfit MNSQ was 0.86, SD=-0.9, and the point-measure correlation was 0.49; in item 3 (“the way you get along with your brothers and sisters?”), the item difficulty was -0.14, SE=0.06, the infit MNSQ was 1.32, SD=1.8, the outfit MNSQ was 1.85, SD=4.2, and the point-measure correlation was 0.26; in item 4 (“The way you get along with other children at school?”), the item difficulty was -0.14, SE=0.06, the infit MNSQ was 0.95, SD=-0.3, the outfit MNSQ was 0.99, SD=0, and the point-measure correlation was 0.53; in item 5 (“the way you get along with other children outside of school?”), the item difficulty was 0, SE=0.06, the infit MNSQ was 0.89, SD=-0.7, the outfit MNSQ was 0.93, SD=-0.5, and the point-measure correlation was 0.57; in item 6 (“the way you get along with adults?”), the item difficulty was 0, SE=0.07, the infit MNSQ was 0.97, SD=-0.2, the outfit MNSQ was 0.97, SD=-0.2, and the point-measure correlation was 0.47; in item 7 (“the way you get along with your teachers and/or careers?”), the item difficulty was -0.34, SE=0.07, the infit MNSQ was 0.98, SD=-0.1, the outfit MNSQ was

1.04, SD=0.3, and the point-measure correlation was 0.46; in item 8 (“your ability to play on your own?”), the item difficulty was 0, SE=0.06, the infit MNSQ was 1.27, SD=2, the outfit MNSQ was 1.46, SD=3.2, and the point-measure correlation was 0.34; in item 9 (“your ability to play with friends?”), the item difficulty was -0.26, SE=0.05, the infit MNSQ was 0.95, SD=-0.3, the outfit MNSQ was 1, SD=0.1, and the point-measure correlation was 0.53; in item 10 (“going out on trips with your family?”), the item difficulty was -0.59, SE=0.07, the infit MNSQ was 0.94, SD=-0.5, the outfit MNSQ was 0.94, SD=-0.5, and the point-measure correlation was 0.48; in item 11 (“how you are accepted by your family?”), the item difficulty was -0.42, SE=0.07, the infit MNSQ was 0.92, SD=-0.4, the outfit MNSQ was 0.84, SD=-1.1, and the point-measure correlation was 0.5; in item 12 (“how you are accepted by other children at school?”), the item difficulty was -0.18, SE=0.06, the infit MNSQ was 0.9, SD=-0.7, the outfit MNSQ was 0.96, SD=-0.2, and the point-measure correlation was 0.54; in item 13 (“how you are accepted by other children outside of school?”), the item difficulty was -0.06, SE=0.06, the infit MNSQ was 0.79, SD=-1.8, the outfit MNSQ was 0.8, SD=-1.7, and the point-measure correlation was 0.64; in item 14 (“how you are accepted by adults?”), the item difficulty was -0.17, SE=0.07, the infit MNSQ was 0.84, SD=-0.9, the outfit MNSQ was 0.85, SD=-1, and the point-measure correlation was 0.56; in item 15 (“how you are accepted by people in general?”), the item difficulty was -0.04, SE=0.06, the infit MNSQ was 1.04, SD=0.4, the outfit MNSQ was 1.06, SD=0.6, and the point-measure correlation was 0.5; in item 16

("being able to do the things you want to do?"), the item difficulty was -0.14, SE=0.05, the infit MNSQ was 0.92, SD=-0.5, the outfit MNSQ was 0.89, SD=-0.7, and the point-measure correlation was 0.58; in item 17 ("your ability to participate at school?"), the item difficulty was -0.13, SE=0.06, the infit MNSQ was 0.76, SD=-2, the outfit MNSQ was 0.75, SD=-2.1, and the point-measure correlation was 0.67; in item 18 ("your ability to participate in recreational activities?"), the item difficulty was -0.2, SE=0.06, the infit MNSQ was 0.8, SD=-1.6, the outfit MNSQ was 0.79, SD=-1.7, and the point-measure correlation was 0.63; in item 19 ("your ability to participate in sporting activities?"), the item difficulty was 0.09, SE=0.05, the infit MNSQ was 0.9, SD=-0.8, the outfit MNSQ was 0.9, SD=-0.8, and the point-measure correlation was 0.61; in item 20 ("your ability to participate in social events outside of school?"), the item difficulty was -0.08, SE=0.05, the infit MNSQ was 0.85, SD=-1.2, the outfit MNSQ was 0.85, SD=-1.2, and the point-measure correlation was 0.61; in item 21 ("your ability to participate in your community?"), the item difficulty was 0.02, SE=0.06, the infit MNSQ was 0.92, SD=-0.7, the outfit MNSQ was 0.91, SD=-0.7, and the point-measure correlation was 0.59; in item 22 ("the way you communicate with people you know well?"), the item difficulty was -0.04, SE=0.07, the infit MNSQ was 0.84, SD=-1.2, the outfit MNSQ was 0.85, SD=-1.2, and the point-measure correlation was 0.57; in item 23 ("the way you communicate with people you don't know well?"), the item difficulty was 0.39, SE=0.06, the infit MNSQ was 1.2, SD=1.5, the outfit MNSQ was 1.28, SD=2, and the point-measure correlation was 0.34; in

item 24 (“the way other people communicate with you?”), the item difficulty was 0.06, SE=0.06, the infit MNSQ was 0.94, SD=-0.4, the outfit MNSQ was 0.94, SD=-0.4, and the point-measure correlation was 0.55; in item 25 (“your physical health?”), the item difficulty was -0.13, SE=0.06, the infit MNSQ was 0.81, SD=-1.8, the outfit MNSQ was 0.81, SD=-1.7, and the point-measure correlation was 0.65; in item 26 (“the way you get around?”), the item difficulty was -0.02, SE=0.05, the infit MNSQ was 0.74, SD=-2.4, the outfit MNSQ was 0.71, SD=-2.5, and the point-measure correlation was 0.69; in item 27 (“how you sleep?”), the item difficulty was -0.31, SE=0.06, the infit MNSQ was 0.85, SD=-1.1, the outfit MNSQ was 0.81, SD=-1.4, and the point-measure correlation was 0.57; in item 28 (“the way you look?”), the item difficulty was -0.23, SE=0.07, the infit MNSQ was 0.83, SD=-1.1, the outfit MNSQ was 0.83, SD=-1.2, and the point-measure correlation was 0.58; in item 29 (“your ability to keep up academically with your peers?”), the item difficulty was 0.23, SE=0.05, the infit MNSQ was 1.11, SD=1, the outfit MNSQ was 1.36, SD=2.7, and the point-measure correlation was 0.49; in item 30 (“your ability to keep up physically with your peers?”), the item difficulty was 0.4, SE=0.05, the infit MNSQ was 0.94, SD=-0.5, the outfit MNSQ was 1.02, SD=0.2, and the point-measure correlation was 0.61; in item 31 (“your life in general?”), the item difficulty was -0.23, SE=0.06, the infit MNSQ was 0.66, SD=-2.6, the outfit MNSQ was 0.64, SD=-3, and the point-measure correlation was 0.72; in item 32 (“yourself?”), the item difficulty was 0.1, SE=0.06, the infit MNSQ was 0.67, SD=-3.1, the outfit MNSQ was 0.68, SD=-3, and the

point-measure correlation was 0.73; in item 33 (“your future?”), the item difficulty was 0.1, SE=0.05, the infit MNSQ was 0.77, SD=-2, the outfit MNSQ was 0.77, SD=-2, and the point-measure correlation was 0.68; in item 34 (“your opportunities in life?”), the item difficulty was 0.07, SE=0.05, the infit MNSQ was 0.74, SD=-2.4, the outfit MNSQ was 0.73, SD=-2.4, and the point-measure correlation was 0.7; in item 35 (“the way you use your arms?”), the item difficulty was 0.04, SE=0.05, the infit MNSQ was 0.89, SD=-0.9, the outfit MNSQ was 0.9, SD=-0.8, and the point-measure correlation was 0.61; in item 36 (“the way you use your legs?”), the item difficulty was 0.33, SE=0.05, the infit MNSQ was 0.88, SD=-1, the outfit MNSQ was 0.88, SD=-1, and the point-measure correlation was 0.65; in item 37 (“the way you use your hands?”), the item difficulty was 0.1, SE=0.05, the infit MNSQ was 0.94, SD=-0.5, the outfit MNSQ was 0.97, SD=-0.2, and the point-measure correlation was 0.59; in item 38 (“your ability to dress yourself?”), the item difficulty was 0.22, SE=0.05, the infit MNSQ was 0.89, SD=-0.9, the outfit MNSQ was 0.91, SD=-0.7, and the point-measure correlation was 0.63; in item 39 (“your ability to eat or drink independently?”), the item difficulty was -0.09, SE=0.05, the infit MNSQ was 0.83, SD=-1.4, the outfit MNSQ was 0.82, SD=-1.3, and the point-measure correlation was 0.61; in item 40 (“your ability to use the toilet by yourself?”), the item difficulty was 0.18, SE=0.05, the infit MNSQ was 1.02, SD=0.2, the outfit MNSQ was 1.09, SD=0.7, and the point-measure correlation was 0.56; in item 41 (“the special equipment you have at home?”), the item difficulty was 0.13, SE=0.07, the infit MNSQ

was 1.17, SD=1.1, the outfit MNSQ was 1.28, SD=1.7, and the point-measure correlation was 0.43; in item 42 (“the special equipment you have at your school?”), the item difficulty was 0.08, SE=0.07, the infit MNSQ was 0.91, SD=-0.5, the outfit MNSQ was 0.97, SD=-0.1, and the point-measure correlation was 0.57; in item 43 (“the special equipment that is available in the community?”), the item difficulty was 0.37, SE=0.06, the infit MNSQ was 1, SD=0.1, the outfit MNSQ was 1.03, SD=0.2, and the point-measure correlation was 0.57; in item 44 (“are you bothered by hospital visits?”), the item difficulty was 0.12, SE=0.05, the infit MNSQ was 1.56, SD=4.2, the outfit MNSQ was 1.61, SD=4, and the point-measure correlation was 0.36; in item 45 (“are you bothered when you miss school for health reasons?”), the item difficulty was 0.19, SE=0.04, the infit MNSQ was 1.99, SD=6.9, the outfit MNSQ was 2.5, SD=7.8, and the point-measure correlation was 0.22; in item 46 (“are you bothered by being handled by other people?”), the item difficulty was 0.12, SE=0.04, the infit MNSQ was 1.83, SD=5.9, the outfit MNSQ was 2.19, SD=7, and the point-measure correlation was 0.23; in item 47 (“do you worry about who will take care of you in the future?”), the item difficulty was 1.08, SE=0.07, the infit MNSQ was 1.42, SD=3.4, the outfit MNSQ was 1.39, SD=2.9, and the point-measure correlation was 0.16; in item 48 (“are you concerned about having cerebral palsy?”), the item difficulty was 0.29, SE=0.04, the infit MNSQ was 2.01, SD=7, the outfit MNSQ was 4.84, SD=9.9, and the point-measure correlation was 0.23; in item 49 (“how much pain do you have?”), the item difficulty was -0.07, SE=0.05, the infit MNSQ was 1.15, SD=1.3, the outfit

MNSQ was 1.12, SD=0.8, and the point-measure correlation was 0.5; in item 50 ("how do you feel about the amount of pain you have?"), the item difficulty was 0.01, SE=0.05, the infit MNSQ was 1.23, SD=1.8, the outfit MNSQ was 1.22, SD=1.5, and the point-measure correlation was 0.48; in item 51 ("how much discomfort do you experience?"), the item difficulty was -0.04, SE=0.05, the infit MNSQ was 1.18, SD=1.5, the outfit MNSQ was 1.22, SD=1.5, and the point-measure correlation was 0.48; in item 52 ("how happy are you?"), the item difficulty was -0.28, SE=0.06, the infit MNSQ was 0.72, SD=-2.2, the outfit MNSQ was 0.69, SD=-2.5, and the point-measure correlation was 0.68 (Table 4.1a, Page 64-65).

Table 4.1a

C CP-QOL Child: Child Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
1	-0.08	0.07	0.85	-1.1	0.85	-1.1	0.56
2	-0.32	0.07	0.89	-0.6	0.86	-0.9	0.49
3	-0.14	0.06	1.32	1.8	1.85	4.2	0.26
4	-0.14	0.06	0.95	-0.3	0.99	0	0.53
5	0	0.06	0.89	-0.7	0.93	-0.5	0.57
6	0	0.07	0.97	-0.2	0.97	-0.2	0.47
7	-0.34	0.07	0.98	-0.1	1.04	0.3	0.46
8	0	0.06	1.27	2	1.46	3.2	0.34
9	-0.26	0.05	0.95	-0.3	1	0.1	0.53
10	-0.59	0.07	0.94	-0.5	0.94	-0.5	0.48
11	-0.42	0.07	0.92	-0.4	0.84	-1.1	0.5
12	-0.18	0.06	0.9	-0.7	0.96	-0.2	0.54
13	-0.06	0.06	0.79	-1.8	0.8	-1.7	0.64
14	-0.17	0.07	0.84	-0.9	0.85	-1	0.56
15	-0.04	0.06	1.04	0.4	1.06	0.6	0.5
16	-0.14	0.05	0.92	-0.5	0.89	-0.7	0.58
17	-0.13	0.06	0.76	-2	0.75	-2.1	0.67
18	-0.2	0.06	0.8	-1.6	0.79	-1.7	0.63
19	0.09	0.05	0.9	-0.8	0.9	-0.8	0.61
20	-0.08	0.05	0.85	-1.2	0.85	-1.2	0.61
21	0.02	0.06	0.92	-0.7	0.91	-0.7	0.59
22	-0.04	0.07	0.84	-1.2	0.85	-1.2	0.57
23	0.39	0.06	1.2	1.5	1.28	2	0.34
24	0.06	0.06	0.94	-0.4	0.94	-0.4	0.55
25	-0.13	0.06	0.81	-1.8	0.81	-1.7	0.65
26	-0.02	0.05	0.74	-2.4	0.71	-2.5	0.69
27	-0.31	0.06	0.85	-1.1	0.81	-1.4	0.57
28	-0.23	0.07	0.83	-1.1	0.83	-1.2	0.58
29	0.23	0.05	1.11	1	1.36	2.7	0.49
30	0.4	0.05	0.94	-0.5	1.02	0.2	0.61
31	-0.23	0.06	0.66	-2.6	0.64	-3	0.72
32	0.1	0.06	0.67	-3.1	0.68	-3	0.73
33	0.1	0.05	0.77	-2	0.77	-2	0.68
34	0.07	0.05	0.74	-2.4	0.73	-2.4	0.7

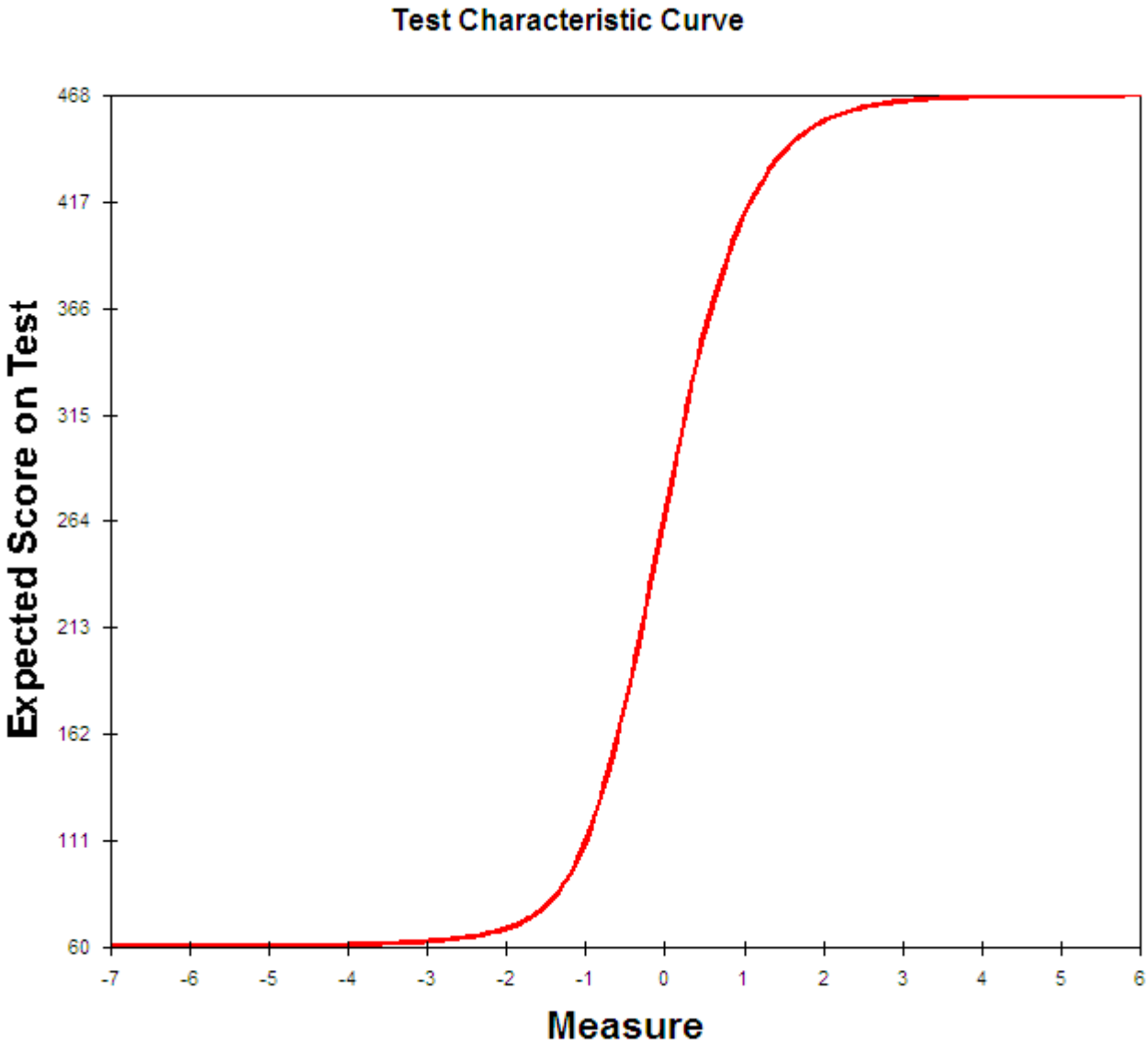
Table 4.1a *(Continued)*

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
35	0.04	0.05	0.89	-0.9	0.9	-0.8	0.61
36	0.33	0.05	0.88	-1	0.88	-1	0.65
37	0.1	0.05	0.94	-0.5	0.97	-0.2	0.59
38	0.22	0.05	0.89	-0.9	0.91	-0.7	0.63
39	-0.09	0.05	0.83	-1.4	0.82	-1.3	0.61
40	0.18	0.05	1.02	0.2	1.09	0.7	0.56
41	0.13	0.07	1.17	1.1	1.28	1.7	0.43
42	0.08	0.07	0.91	-0.5	0.97	-0.1	0.57
43	0.37	0.06	1	0.1	1.03	0.2	0.57
44	0.12	0.05	1.56	4.2	1.61	4	0.36
45	0.19	0.04	1.99	6.9	2.5	7.8	0.22
46	0.12	0.04	1.83	5.9	2.19	7	0.23
47	1.08	0.07	1.42	3.4	1.39	2.9	0.16
48	0.29	0.04	2.01	7	4.84	9.9	0.23
49	-0.07	0.05	1.15	1.3	1.12	0.8	0.5
50	0.01	0.05	1.23	1.8	1.22	1.5	0.48
51	-0.04	0.05	1.18	1.5	1.22	1.5	0.48
52	-0.28	0.06	0.72	-2.2	0.69	-2.5	0.68

The most difficult item within the child scale was item 47 (“do you worry about who will take care of you in the future?”) [1.08, SE=0.07], which was more than two standard deviation above the item mean. The easiest item within the child scale was item 10 (“going out on trips with your family?”) [-0.59, SE=0.07], which was approximately two standard deviation below the item mean (see Appendix, Figure A1)

Figure 4.1

Characteristic Curve of C CP-QOL Child: Child Scale



According to the test characteristic curve of the child scale, the expected score on test ranged from 60 to 468, the expected mean score was 264, and the measure (in logit) ranged from -4 to 4 approximately (Figure 4.1).

After conducting the step calibration item by item in the child scale, the results showed that, in item 1 (“the way you get along with people, generally?”), the Rasch half-point thresholds ranged from -1.65 to 1.76 (-1.65, -0.95, -0.49, 0.03, 0.79 and 1.76 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 1 (Figure C1), the expected scoring range was from 3 to 9 and the expected mean score was 6.5; in item 2 (“the way you get along with the person who looks after you?”), the Rasch half-point thresholds ranged from -1.51 to 1.44 (-1.51, -1.05, -0.82, -0.63, -0.42, -0.13, 0.44 and 1.44 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 2 (Figure C2), the expected scoring range was from 1 to 9 and the expected mean score was 7; in item 3 (“the way you get along with your brothers and sisters?”), the Rasch half-point thresholds ranged from -1.08 to 1.31 (-1.08, -0.75, -0.56, -0.4, -0.23, 0.02, 0.45 and 1.31 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 3 (Figure C3), the expected scoring range was from 0 to 9 and the expected mean score was 5.5; in item 4 (“The way you get along with other children at school?”), the Rasch half-point thresholds ranged from -1.5 to 1.62 (-1.5, -0.91, -0.63, -0.4, -0.15, 0.2, 0.75 and 1.62 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 4 (Figure C4), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 5 (“the way you get along with other children outside of

school?"), the Rasch half-point thresholds ranged from -1.31 to 1.86 (-1.31, -0.86, -0.61, -0.37, -0.08, 0.36, 0.97 and 1.86 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 5 (Figure C5), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 6 ("the way you get along with adults?"), the Rasch half-point thresholds ranged from -1.59 to 1.74 (-1.59, -0.88, -0.37, 0.19, 0.85 and 1.74 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 6 (Figure C6), the expected scoring range was from 3 to 9 and the expected mean score was 6.5; in item 7 ("the way you get along with your teachers and/or careers?"), the Rasch half-point thresholds ranged from -1.39 to 1.32 (-1.39, -1.07, -0.88, -0.7, -0.48, -0.12, 0.45 and 1.32 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 7 (Figure C7), the expected scoring range was from 1 to 9 and the expected mean score was 7; in item 8 ("your ability to play on your own?"), the Rasch half-point thresholds ranged from -1.45 to 1.81 (-1.45, -0.89, -0.57, -0.29, 0.01, 0.37, 0.89 and 1.81 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 8 (Figure C8), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 9 ("your ability to play with friends?"), the Rasch half-point thresholds ranged from -1.96 to 1.37 (-1.96, -1.09, -0.69, -0.39, -0.1, 0.22, 0.64 and 1.37 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and

Step 8 respectively), and according to the item characteristic curve of item 9 (Figure C9), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 10 ("going out on trips with your family?"), the Rasch half-point thresholds ranged from -2.88 to 0.99 (-2.88, -0.91, -0.27, 0.23 and 0.99 in Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 10 (Figure C10), the expected scoring range was from 4 to 9 and the expected mean score was 7; in item 11 ("how you are accepted by your family?"), the Rasch half-point thresholds ranged from -1.39 to 1.02 (-1.39, -1.04, -0.84, -0.68, -0.49, -0.24, 0.2 and 1.02 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 11 (Figure C11), the expected scoring range was from 1 to 9 and the expected mean score was 7; in item 12 ("how you are accepted by other children at school?"), the Rasch half-point thresholds ranged from -1.53 to 1.46 (-1.53, -0.88, -0.6, -0.38, -0.16, 0.14, 0.61 and 1.46 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 12 (Figure C12), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 13 ("how you are accepted by other children outside of school?"), the Rasch half-point thresholds ranged from -1.62 to 1.77 (-1.62, -0.98, -0.63, -0.32, 0, 0.39, 0.91 and 1.77 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 13 (Figure C13), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 14

("how you are accepted by adults?"), the Rasch half-point thresholds ranged from -1.54 to 2.04 (-1.54, -1.1, -0.85, -0.63, -0.36, 0.09, 0.89 and 2.04 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 14 (Figure C14), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 15 ("how you are accepted by people in general?"), the Rasch half-point thresholds ranged from -1.64 to 1.75 (-1.64, -0.97, -0.61, -0.3, 0.06, 0.46, 0.94 and 1.75 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 15 (Figure C15), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 16 ("being able to do the things you want to do?"), the Rasch half-point thresholds ranged from -1.22 to 1.18 (-1.22, -0.78, -0.54, -0.33, -0.12, 0.14, 0.5 and 1.18 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 16 (Figure C16), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 17 ("your ability to participate at school?"), the Rasch half-point thresholds ranged from -1.51 to 1.53 (-1.51, -0.96, -0.66, -0.39, -0.09, 0.27, 0.74 and 1.53 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 17 (Figure C17), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 18 ("your ability to participate in recreational activities?"), the Rasch half-point thresholds ranged from -1.82 to 1.64 (-1.82, -1.09, -0.75, -0.45, -0.12, 0.29, 0.82

and 1.64 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 18 (Figure C18), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 19 ("your ability to participate in sporting activities?"), the Rasch half-point thresholds ranged from -1.17 to 1.61 (-1.17, -0.72, -0.45, -0.18, 0.13, 0.49, 0.9 and 1.61 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 19 (Figure C19), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 20 ("your ability to participate in social events outside of school?"), the Rasch half-point thresholds ranged from -1.67 to 1.64 (-1.67, -0.96, -0.59, -0.29, 0.03, 0.4, 0.86 and 1.64 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 20 (Figure C20), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 21 ("your ability to participate in your community?"), the Rasch half-point thresholds ranged from -1.73 to 1.82 (-1.73, -1, -0.59, -0.2, 0.21, 0.62, 1.07 and 1.82 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 21 (Figure C21), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 22 ("the way you communicate with people you know well?"), the Rasch half-point thresholds ranged from -1.51 to 1.75 (-1.51, -0.88, -0.45, 0.05, 0.75 and 1.75 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 22

(Figure C22), the expected scoring range was from 3 to 9 and the expected mean score was 6.5; in item 23 ("the way you communicate with people you don't know well?"), the Rasch half-point thresholds ranged from -1.27 to 2.51 (-1.27, -0.71, -0.38, -0.04, 0.41, 0.97, 1.61 and 2.51 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 23 (Figure C23), the expected scoring range was from 1 to 9 and the expected mean score was 4.5; in item 24 ("the way other people communicate with you?"), the Rasch half-point thresholds ranged from -1.6 to 1.97 (-1.6, -1.02, -0.67, -0.3, 0.18, 0.69, 1.2 and 1.97 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 24 (Figure C24), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 25 ("your physical health?"), the Rasch half-point thresholds ranged from -2.99 to 1.73 (-2.99, -0.8, -0.28, 0.09, 0.48, 0.96 and 1.73 in Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 25 (Figure C25), the expected scoring range was from 2 to 9 and the expected mean score was 5.5; in item 26 ("the way you get around?"), the Rasch half-point thresholds ranged from -1.59 to 1.65 (-1.59, -0.86, -0.48, -0.19, 0.08, 0.39, 0.83 and 1.65 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 26 (Figure C26), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 27 ("how you sleep?"), the Rasch half-point thresholds ranged from -1.58 to

1.34 (-1.58, -1.11, -0.84, -0.6, -0.33, 0.03, 0.52 and 1.34 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 27 (Figure C27), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 28 ("the way you look?"), the Rasch half-point thresholds ranged from -1.53 to 1.7 (-1.53, -1.11, -0.87, -0.64, -0.35, 0.1, 0.76 and 1.7 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 28 (Figure C28), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 29 ("your ability to keep up academically with your peers?"), the Rasch half-point thresholds ranged from -0.98 to 1.76 (-0.98, -0.48, -0.23, -0.01, 0.23, 0.54, 0.97 and 1.76 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 29 (Figure C29), the expected scoring range was from 1 to 9 and the expected mean score was 4.5; in item 30 ("your ability to keep up physically with your peers?"), the Rasch half-point thresholds ranged from -0.94 to 1.98 (-0.94, -0.38, -0.09, 0.17, 0.44, 0.77, 1.22 and 1.98 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 30 (Figure C30), the expected scoring range was from 1 to 9 and the expected mean score was 4; in item 31 ("your life in general?"), the Rasch half-point thresholds ranged from -1.66 to 1.63 (-1.66, -1.09, -0.8, -0.55, -0.26, 0.15, 0.74 and 1.63 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve

of item 31 (Figure C31), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 32 (“yourself?”), the Rasch half-point thresholds ranged from -1.39 to 1.68 (-1.39, -0.66, -0.18, 0.29, 0.83 and 1.68 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 32 (Figure C32), the expected scoring range was from 3 to 9 and the expected mean score was 6; in item 33 (“your future?”), the Rasch half-point thresholds ranged from -1.3 to 1.79 (-1.3, -0.74, -0.43, -0.16, 0.14, 0.5, 0.96 and 1.79 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 33 (Figure C33), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 34 (“your opportunities in life?”), the Rasch half-point thresholds ranged from -1.44 to 1.73 (-1.44, -0.78, -0.43, -0.14, 0.17, 0.52, 0.95 and 1.73 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 34 (Figure C34), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 35 (“the way you use your arms?”), the Rasch half-point thresholds ranged from -1.55 to 1.73 (-1.55, -0.88, -0.49, -0.15, 0.18, 0.53, 0.96 and 1.73 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 35 (Figure C35), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 36 (“the way you use your legs?”), the Rasch half-point thresholds ranged from -1.22 to 2 (-1.22, -0.53, -0.15, 0.14, 0.42, 0.74, 1.18 and 2 in Step 1, Step 2, Step

3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 36 (Figure C36), the expected scoring range was from 1 to 9 and the expected mean score was 4; in item 37 ("the way you use your hands?"), the Rasch half-point thresholds ranged from -1.35 to 1.72 (-1.35, -0.72, -0.37, -0.1, 0.17, 0.48, 0.91 and 1.72 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 37 (Figure C37), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 38 ("your ability to dress yourself?"), the Rasch half-point thresholds ranged from -1.17 to 1.72 (-1.17, -0.51, -0.19, 0.05, 0.29, 0.57, 0.97 and 1.72 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 38 (Figure C38), the expected scoring range was from 1 to 9 and the expected mean score was 4.5; in item 39 ("your ability to eat or drink independently?"), the Rasch half-point thresholds ranged from -1.28 to 1.3 (-1.28, -0.69, -0.44, -0.25, -0.06, 0.18, 0.54 and 1.3 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 39 (Figure C39), the expected scoring range was from 0 to 9 and the expected mean score was 5.5; in item 40 ("your ability to use the toilet by yourself?"), the Rasch half-point thresholds ranged from -1.03 to 1.58 (-1.03, -0.5, -0.23, -0.01, 0.22, 0.49, 0.87 and 1.58 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 40 (Figure C40), the expected scoring range was from

0 to 9 and the expected mean score was 4.5; in item 41 ("the special equipment you have at home?"), the Rasch half-point thresholds ranged from -1.36 to 1.95 (-1.36, -0.74, -0.42, -0.15, 0.15, 0.54, 1.08 and 1.95 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 41 (Figure C41), the expected scoring range was from 1 to 9 and the expected mean score was 3; in item 42 ("the special equipment you have at your school?"), the Rasch half-point thresholds ranged from -1.3 to 1.93 (-1.3, -0.76, -0.49, -0.24, 0.04, 0.43, 1 and 1.93 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 42 (Figure C42), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 43 ("the special equipment that is available in the community?"), the Rasch half-point thresholds ranged from -0.89 to 2.16 (-0.89, -0.42, -0.16, 0.06, 0.31, 0.65, 1.18 and 2.16 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 43 (Figure C43), the expected scoring range was from 0 to 9 and the expected mean score was 2.5; in item 44 ("are you bothered by hospital visits?"), the Rasch half-point thresholds ranged from -1.12 to 1.27 (-1.12, -0.53, -0.21, 0.05, 0.28, 0.5, 0.77 and 1.27 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 44 (Figure C44), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 45 ("are you bothered when you miss school for health reasons?"), the Rasch half-point thresholds ranged

from -0.88 to 1.2 (-0.88, -0.35, -0.08, 0.12, 0.31, 0.51, 0.75 and 1.2 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 45 (Figure C45), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 46 ("are you bothered by being handled by other people?"), the Rasch half-point thresholds ranged from -1.04 to 1.27 (-1.04, -0.48, -0.2, 0.03, 0.24, 0.47, 0.74 and 1.27 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 46 (Figure C46), the expected scoring range was from 1 to 9 and the expected mean score was 3; in item 47 ("do you worry about who will take care of you in the future?"), the Rasch half-point thresholds ranged from -1.24 to 2.55 (-1.24, -0.22, 0.48, 1.29, 1.76, 2, 2.22 and 2.55 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 47 (Figure C47), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 48 ("are you concerned about having cerebral palsy?"), the Rasch half-point thresholds ranged from -0.63 to 1.23 (-0.63, -0.19, 0.03, 0.21, 0.38, 0.55, 0.78 and 1.23 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 48 (Figure C48), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 49 ("how much pain do you have?"), the Rasch half-point thresholds ranged from -1.22 to 1.09 (-1.22, -0.71, -0.42, -0.19, 0.04, 0.27, 0.56 and 1.09 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and

Step 8 respectively), and according to the item characteristic curve of item 49 (Figure C49), the expected scoring range was from 1 to 9 and the expected mean score was 2.5; in item 50 ("how do you feel about the amount of pain you have?"), the Rasch half-point thresholds ranged from -1.12 to 1.12 (-1.12, -0.62, -0.34, -0.1, 0.14, 0.37, 0.64 and 1.12 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 50 (Figure C50), the expected scoring range was from 1 to 9 and the expected mean score was 2.5; in item 51 ("how much discomfort do you experience?"), the Rasch half-point thresholds ranged from -1.28 to 1.12 (-1.28, -0.72, -0.4, -0.13, 0.11, 0.35, 0.63 and 1.12 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 51 (Figure C51), the expected scoring range was from 1 to 9 and the expected mean score was 2.5; in item 52 ("how happy are you?"), the Rasch half-point thresholds ranged from -1.59 to 1.39 (-1.59, -1.1, -0.82, -0.58, -0.29, 0.11, 0.62 and 1.39 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 52 (Figure C52), the expected scoring range was from 1 to 9 and the expected mean score was 6.5 (Table 4.1b, Page 79-80).

Table 4.1b

Rasch half-point thresholds of C CP-QOL Child: Child Scale

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
1			-1.65	-0.95	-0.49	0.03	0.79	1.76
2	-1.51	-1.05	-0.82	-0.63	-0.42	-0.13	0.44	1.44
3	-1.08	-0.75	-0.56	-0.4	-0.23	0.02	0.45	1.31
4	-1.5	-0.91	-0.63	-0.4	-0.15	0.2	0.75	1.62
5	-1.31	-0.86	-0.61	-0.37	-0.08	0.36	0.97	1.86
6			-1.59	-0.88	-0.37	0.19	0.85	1.74
7	-1.39	-1.07	-0.88	-0.7	-0.48	-0.12	0.45	1.32
8	-1.45	-0.89	-0.57	-0.29	0.01	0.37	0.89	1.81
9	-1.96	-1.09	-0.69	-0.39	-0.1	0.22	0.64	1.37
10				-2.88	-0.91	-0.27	0.23	0.99
11	-1.39	-1.04	-0.84	-0.68	-0.49	-0.24	0.2	1.02
12	-1.53	-0.88	-0.6	-0.38	-0.16	0.14	0.61	1.46
13	-1.62	-0.98	-0.63	-0.32	0	0.39	0.91	1.77
14	-1.54	-1.1	-0.85	-0.63	-0.36	0.09	0.89	2.04
15	-1.64	-0.97	-0.61	-0.3	0.06	0.46	0.94	1.75
16	-1.22	-0.78	-0.54	-0.33	-0.12	0.14	0.5	1.18
17	-1.51	-0.96	-0.66	-0.39	-0.09	0.27	0.74	1.53
18	-1.82	-1.09	-0.75	-0.45	-0.12	0.29	0.82	1.64
19	-1.17	-0.72	-0.45	-0.18	0.13	0.49	0.9	1.61
20	-1.67	-0.96	-0.59	-0.29	0.03	0.4	0.86	1.64
21	-1.73	-1	-0.59	-0.2	0.21	0.62	1.07	1.82
22			-1.51	-0.88	-0.45	0.05	0.75	1.75
23	-1.27	-0.71	-0.38	-0.04	0.41	0.97	1.61	2.51
24	-1.6	-1.02	-0.67	-0.3	0.18	0.69	1.2	1.97
25		-2.99	-0.8	-0.28	0.09	0.48	0.96	1.73
26	-1.59	-0.86	-0.48	-0.19	0.08	0.39	0.83	1.65
27	-1.58	-1.11	-0.84	-0.6	-0.33	0.03	0.52	1.34
28	-1.53	-1.11	-0.87	-0.64	-0.35	0.1	0.76	1.7
29	-0.98	-0.48	-0.23	-0.01	0.23	0.54	0.97	1.76
30	-0.94	-0.38	-0.09	0.17	0.44	0.77	1.22	1.98
31	-1.66	-1.09	-0.8	-0.55	-0.26	0.15	0.74	1.63
32			-1.39	-0.66	-0.18	0.29	0.83	1.68
33	-1.3	-0.74	-0.43	-0.16	0.14	0.5	0.96	1.79
34	-1.44	-0.78	-0.43	-0.14	0.17	0.52	0.95	1.73

Table 4.1b *(Continued)*

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
35	-1.55	-0.88	-0.49	-0.15	0.18	0.53	0.96	1.73
36	-1.22	-0.53	-0.15	0.14	0.42	0.74	1.18	2
37	-1.35	-0.72	-0.37	-0.1	0.17	0.48	0.91	1.72
38	-1.17	-0.51	-0.19	0.05	0.29	0.57	0.97	1.72
39	-1.28	-0.69	-0.44	-0.25	-0.06	0.18	0.54	1.3
40	-1.03	-0.5	-0.23	-0.01	0.22	0.49	0.87	1.58
41	-1.36	-0.74	-0.42	-0.15	0.15	0.54	1.08	1.95
42	-1.3	-0.76	-0.49	-0.24	0.04	0.43	1	1.93
43	-0.89	-0.42	-0.16	0.06	0.31	0.65	1.18	2.16
44	-1.12	-0.53	-0.21	0.05	0.28	0.5	0.77	1.27
45	-0.88	-0.35	-0.08	0.12	0.31	0.51	0.75	1.2
46	-1.04	-0.48	-0.2	0.03	0.24	0.47	0.74	1.27
47	-1.24	-0.22	0.48	1.29	1.76	2	2.22	2.55
48	-0.63	-0.19	0.03	0.21	0.38	0.55	0.78	1.23
49	-1.22	-0.71	-0.42	-0.19	0.04	0.27	0.56	1.09
50	-1.12	-0.62	-0.34	-0.1	0.14	0.37	0.64	1.12
51	-1.28	-0.72	-0.4	-0.13	0.11	0.35	0.63	1.12
52	-1.59	-1.1	-0.82	-0.58	-0.29	0.11	0.62	1.39

The item content, categorized by different domains, was shown as blow (each question begins with “How do you fell about…”):

Table 4.1c

CP-QOL Child: Child Scale grouped by domain

Social Wellbeing and Acceptance
1 the way you get along with people, generally?
3 the way you get along with your brothers and sisters?
4 The way you get along with other children at school?
5 the way you get along with other children outside of school?
6 the way you get along with adults?
7 the way you get along with your teachers and/or careers?
10 going out on trips with your family?
11 how you are accepted by your family?
12 how you are accepted by other children at school?
13 how you are accepted by other children outside of school?
14 how you are accepted by adults?
15 how you are accepted by people in general?
Functioning
8 your ability to play on your own?
22 the way you communicate with people you know well?
23 the way you communicate with people you don't know well?
24 the way other people communicate with you?
27 how you sleep?
29 your ability to keep up academically with your peers?
34 your opportunities in life?
35 the way you use your arms?
37 the way you use your hands?
38 your ability to dress yourself?
39 your ability to eat or drink independently?
40 your ability to use the toilet by yourself?
Participation and Physical Health
9 your ability to play with friends?
16 being able to do the things you want to do?
17 your ability to participate at school?
18 your ability to participate in recreational activities?
19 your ability to participate in sporting activities?
20 your ability to participate in social events outside of school?
21 your ability to participate in your community?
25 your physical health?
26 the way you get around?

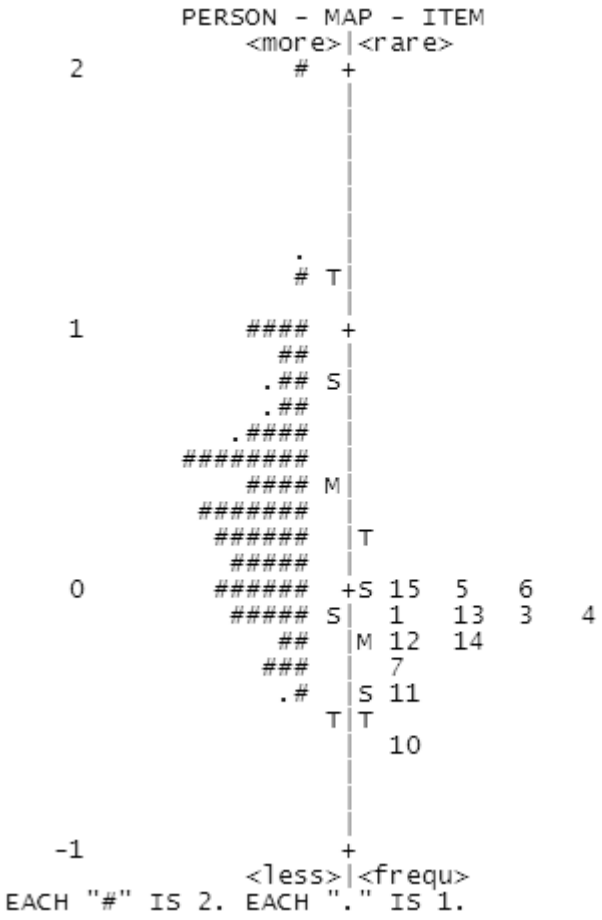
Table 4.1c *(Continued)*

Participation and Physical Health
30 your ability to keep up physically with your peers?
36 the way you use your legs?
Emotional Wellbeing
2 the way you get along with the person who looks after you?
28 the way you look?
31 your life in general?
32 yourself?
33 your future?
52 how happy are you?
Pain and Impact of Disability
44 are you bothered by hospital visits?
45 are you bothered when you miss school for health reasons?
46 are you bothered by being handled by other people?
47 do you worry about who will take care of you in the future?
48 are you concerned about having cerebral palsy?
49 how much pain do you have?
50 how do you feel about the amount of pain you have?
51 how much discomfort do you experience?

4.1.1. Subscales of child self-report

Figure 4.1.1a

Social Wellbeing and Acceptance Variable Map



Note:
1 the way you get along with people, generally?
3 the way you get along with your brothers and sisters?
4 the way you get along with other children at school?
5 the way you get along with other children outside of school?
6 the way you get along with adults?
7 the way you get along with your teachers and/or careers?
10 going out on trips with your family?
11 how you are accepted by your family?
12 how you are accepted by other children at school?
13 how you are accepted by other children outside of school?
14 how you are accepted by adults?
15 how you are accepted by people in general?

The person-item map (Figure 4.1.1a, Page 83) shows that the Social Wellbeing and Acceptance scale was relatively "easy" for this sample. The mean for the sample is two standard deviations above the mean for the items. The item difficulties cluster together.

The easiest item (10, "going out on trips with your family?") is approximately two standard deviations below the item mean. The most difficult items (15, "how you are accepted by people in general?"; 5, "the way you get along with other children outside of school?"; 6, "the way you get along with adults?") are approximately one standard deviation above the item mean.

In the preceding table, item 12 ("how you are accepted by other children at school?") and 14 ("how you are accepted by adults?") were just fallen into the mean range of the scale. The item mean, however, was far lower than the person mean, and the wide range of person distribution revealed that there are only limited effects on distinguishing this group of people within the scale.

More specifically, for the alignment of person-item distributions, around 65% person was above the item limit. These information have two meanings, the first one is that there are so many (65%) high "ability" people (or high satisfactory level in terms of social wellbeing and acceptance) within this specific group; the second one is that this subscale itself cannot

adequately measure those (65%) high “ability” people, possibly, design and insert some high “difficulty” items or amend the content of the redundant items (more than one in the same difficulty level) is recommended.

Table 4.1.1a

Social Wellbeing and Acceptance Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
3	-0.14	0.06	1.32	1.8	1.85	4.2	0.26
15	-0.04	0.06	1.04	0.4	1.06	0.6	0.5
7	-0.34	0.07	0.98	-0.1	1.04	0.3	0.46
4	-0.14	0.06	0.95	-0.3	0.99	0	0.53
6	0	0.07	0.97	-0.2	0.97	-0.2	0.47
12	-0.18	0.06	0.9	-0.7	0.96	-0.2	0.54
10	-0.59	0.07	0.94	-0.5	0.94	-0.5	0.48
5	0	0.06	0.89	-0.7	0.93	-0.5	0.57
11	-0.42	0.07	0.92	-0.4	0.84	-1.1	0.5
1	-0.08	0.07	0.85	-1.1	0.85	-1.1	0.56
14	-0.17	0.07	0.84	-0.9	0.85	-1	0.56
13	-0.06	0.06	0.79	-1.8	0.8	-1.7	0.64

The item difficulty logits (Item Score in Table 4.1.1a) ranged from -0.59 to 0 on calibrations of the Social Wellbeing and Acceptance scale. The easiest item was item 10 (“going out on trips with your family?”) [-0.59, SE=0.07]. The most difficult item was item 6 (“the way you get along with adults?”) [0, SE=0.07].

On the examination of the fit statistics of the Social Wellbeing and Acceptance scale, the infit and outfit statistics for items ranged from 0.79 to 1.85. The lowest MNSQ was the infit for item 13 ("how you are accepted by other children outside of school?") [0.79, SD=1.8]. The highest MNSQ was the outfit for item 3 ("how you are accepted by other children outside of school?") [1.85, SD=4.2]

Within this scale, item 7 ("the way you get along with your teachers and/or careers?"), 4 ("the way you get along with other children at school?"), 6 ("the way you get along with adults?"), 12 ("how you are accepted by other children at school?"), 10 ("going out on trips with your family?"), 5 ("the way you get along with other children outside of school?"), and 11 ("how you are accepted by your family?") were fallen into an optimal fit level (infit/outfit, 1±0.1).

Item 3 ("the way you get along with your brothers and sisters?") [outfit, 1.85] was out of the Likert scale critical range (i.e., 0.5-1.5). Item 3, had a slightly low score to total correlation (r=0.26). In order to properly fit the criteria of unidimensionality, item 3 may recommended to be separated from other items in the scale and interpreted accordingly.

Figure 4.1.1b

Social Wellbeing and Acceptance Item Characteristic Curves

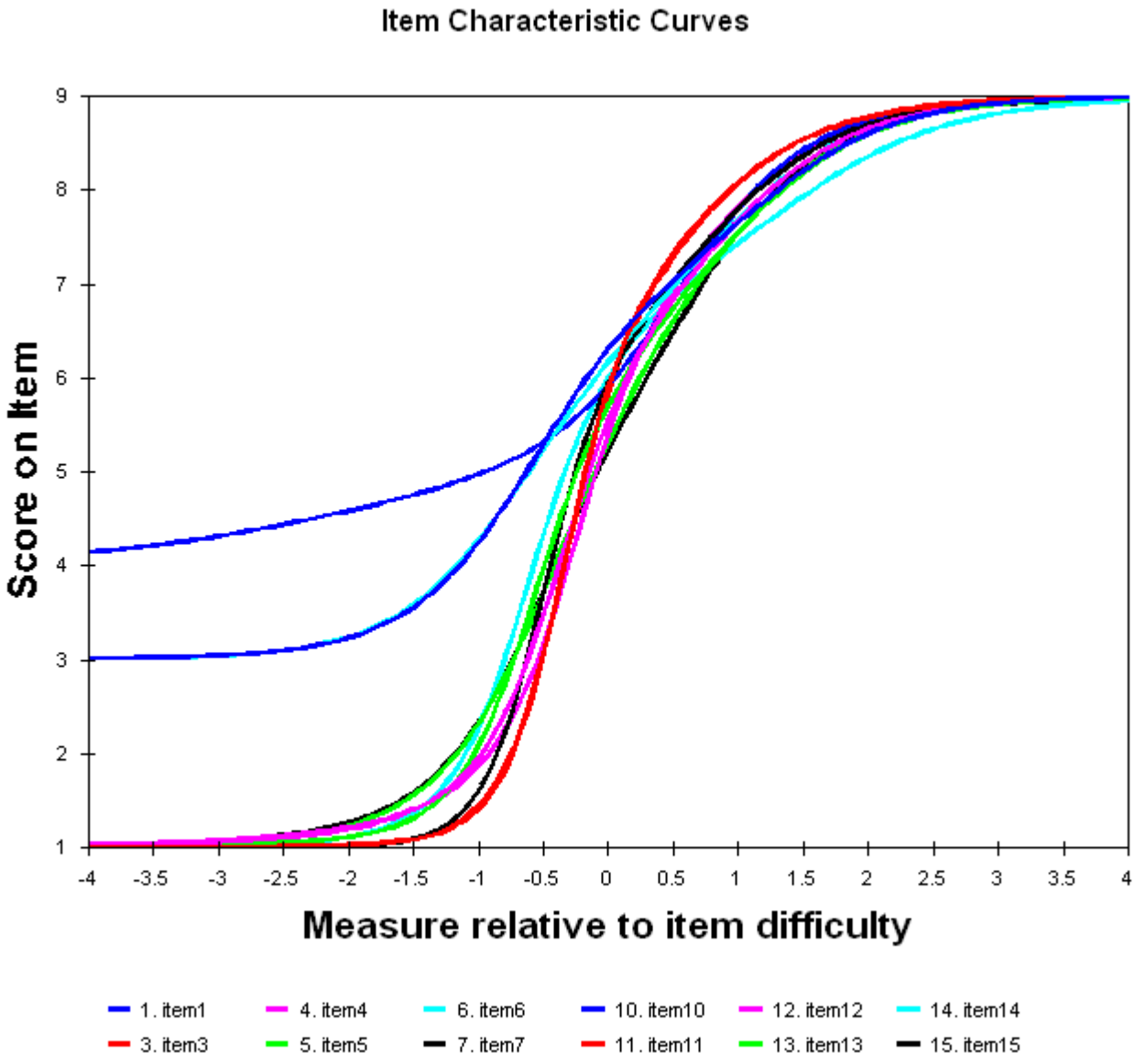


Table 4.1.1b

Social Wellbeing and Acceptance Scale: Rasch half-point thresholds

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
3	-1.08	-0.75	-0.56	-0.4	-0.23	0.02	0.45	1.31
15	-1.64	-0.97	-0.61	-0.3	0.06	0.46	0.94	1.75
7	-1.39	-1.07	-0.88	-0.7	-0.48	-0.12	0.45	1.32
4	-1.5	-0.91	-0.63	-0.4	-0.15	0.2	0.75	1.62
6			-1.59	-0.88	-0.37	0.19	0.85	1.74
12	-1.53	-0.88	-0.6	-0.38	-0.16	0.14	0.61	1.46
10				-2.88	-0.91	-0.27	0.23	0.99
5	-1.31	-0.86	-0.61	-0.37	-0.08	0.36	0.97	1.86
11	-1.39	-1.04	-0.84	-0.68	-0.49	-0.24	0.2	1.02
1			-1.65	-0.95	-0.49	0.03	0.79	1.76
14	-1.54	-1.1	-0.85	-0.63	-0.36	0.09	0.89	2.04
13	-1.62	-0.98	-0.63	-0.32	0	0.39	0.91	1.77

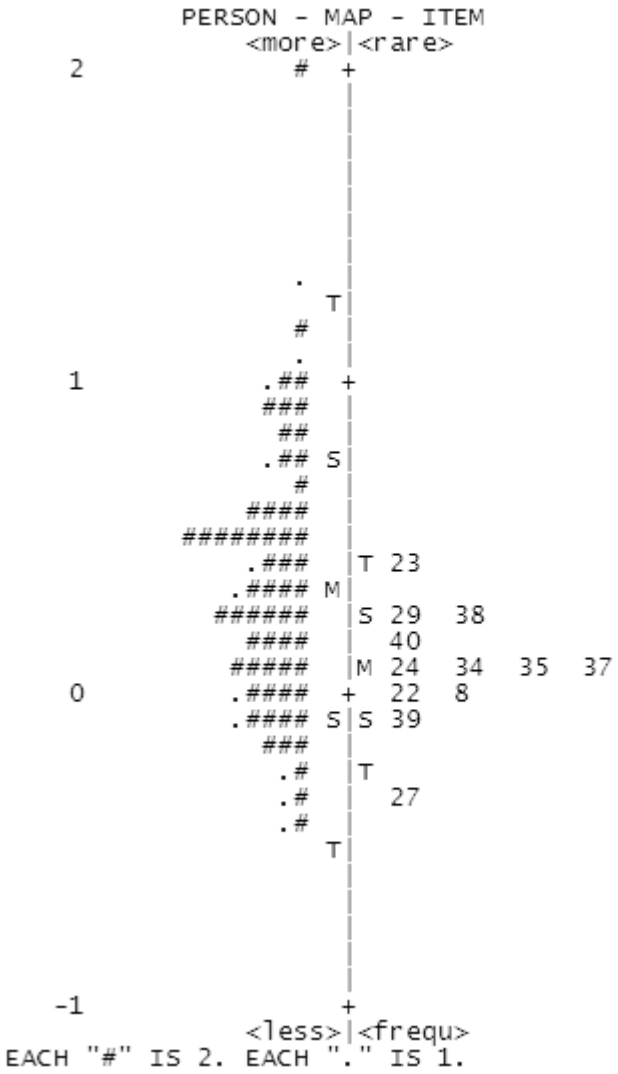
On Rasch half-point step calibrations of the Social Wellbeing and Acceptance scale (Figure 4.1.1b; Table 4.1.1b), the Rasch half-point thresholds for items ranged from -1.64 to 2.04. The lowest threshold was the step 1 for item 15 (“how you are accepted by people in general?”). The highest threshold was the step 8 for item 14 (“how you are accepted by adults?”).

According to Table 4.1.1b, item 6 and item 1 had missed values in step 1 and step 2, it was due to the category 1 and 2 for these two items received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 6 and item 1 may recommended to be assigned a 7-point Likert option instead.

Item 10 had missed values in step 1, 2 and 3; it was due to the category 1, 2 and 3 for item 10 received zero respondent rates (no endorsement). In order to simplify the scale and shorten the completion time, item 10 may recommended to be assigned a 6-point Likert option instead.

Figure 4.1.2a

Functioning Variable Map



Note: 8 your ability to play on your own?

22 the way you communicate with people you know well?

23 the way you communicate with people you don't know well?

24 the way other people communicate with you?

27 how you sleep?

29 your ability to keep up academically with your peers?

34 your opportunities in life?

35 the way you use your arms?

37 the way you use your hands?

38 your ability to dress yourself?

39 your ability to eat or drink independently?

40 your ability to use the toilet by yourself?

The person-item map (Figure 4.1.2a, Page 90) shows that the Functioning scale was relatively "easy" for this sample. The mean for the sample is one and a half standard deviations above the mean for the items. The item difficulties cluster together.

The easiest item (27, "how you sleep?") is approximately two standard deviations below the item mean. The most difficult item (23, "the way you communicate with people you don't know well?") is approximately two standard deviations above the item mean.

In the preceding table, item 24 ("how you are accepted by other children at school?"), 34 ("your opportunities in life?"), 35 ("the way you use your arms?"), and 37 ("the way you use your hands?") were just fallen into the mean range of the scale. The item mean, however, was lower than the person mean, and the wide range of person distribution limited the item effects on distinguishing this group of people within the scale.

In details, for the alignment of person-item distributions, around 35% person was above the item limit. These information have two meanings, the first one is that there are so many (35%) high "ability" people (or high satisfactory level in terms of functioning) within this group of samples; the second one is that this subscale itself cannot adequately measure those (35%)

high “ability” people, possibly, design and insert some high “difficulty” items or amend the content of the redundant items is recommended.

Table 4.1.2a

Functioning Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
8	0	0.06	1.27	2	1.46	3.2	0.34
29	0.23	0.05	1.11	1	1.36	2.7	0.49
23	0.39	0.06	1.2	1.5	1.28	2	0.34
40	0.18	0.05	1.02	0.2	1.09	0.7	0.56
37	0.1	0.05	0.94	-0.5	0.97	-0.2	0.59
24	0.06	0.06	0.94	-0.4	0.94	-0.4	0.55
38	0.22	0.05	0.89	-0.9	0.91	-0.7	0.63
35	0.04	0.05	0.89	-0.9	0.9	-0.8	0.61
27	-0.31	0.06	0.85	-1.1	0.81	-1.4	0.57
22	-0.04	0.07	0.84	-1.2	0.85	-1.2	0.57
39	-0.09	0.05	0.83	-1.4	0.82	-1.3	0.61
34	0.07	0.05	0.74	-2.4	0.73	-2.4	0.7

The item difficulty logits (Item Score in Table 4.1.2a) ranged from -0.31 to 0.39 on calibrations of the Functioning scale. The easiest item was item 27 (“how you sleep?”) [-0.31, SE=0.06]. The most difficult item was item 23 (“the way you communicate with people you don't know well?”) [0.39, SE=0.06].

On the examination of the fit statistics of the Functioning scale, the infit and outfit statistics for items ranged from 0.73 to 1.46. The lowest MNSQ was the outfit for item 34 (“your

opportunities in life?") [0.73, SD=-2.4]. The highest MNSQ was the outfit for item 8 ("your ability to play on your own?") [1.46, SD=3.2]

Within this scale, item 40 ("your ability to use the toilet by yourself?"), 37 ("the way you use your hands?"), 24 ("the way other people communicate with you?"), 38 ("your ability to dress yourself?"), and 35 ("the way you use your arms?") were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.1.2b

Functioning Item Characteristic Curves

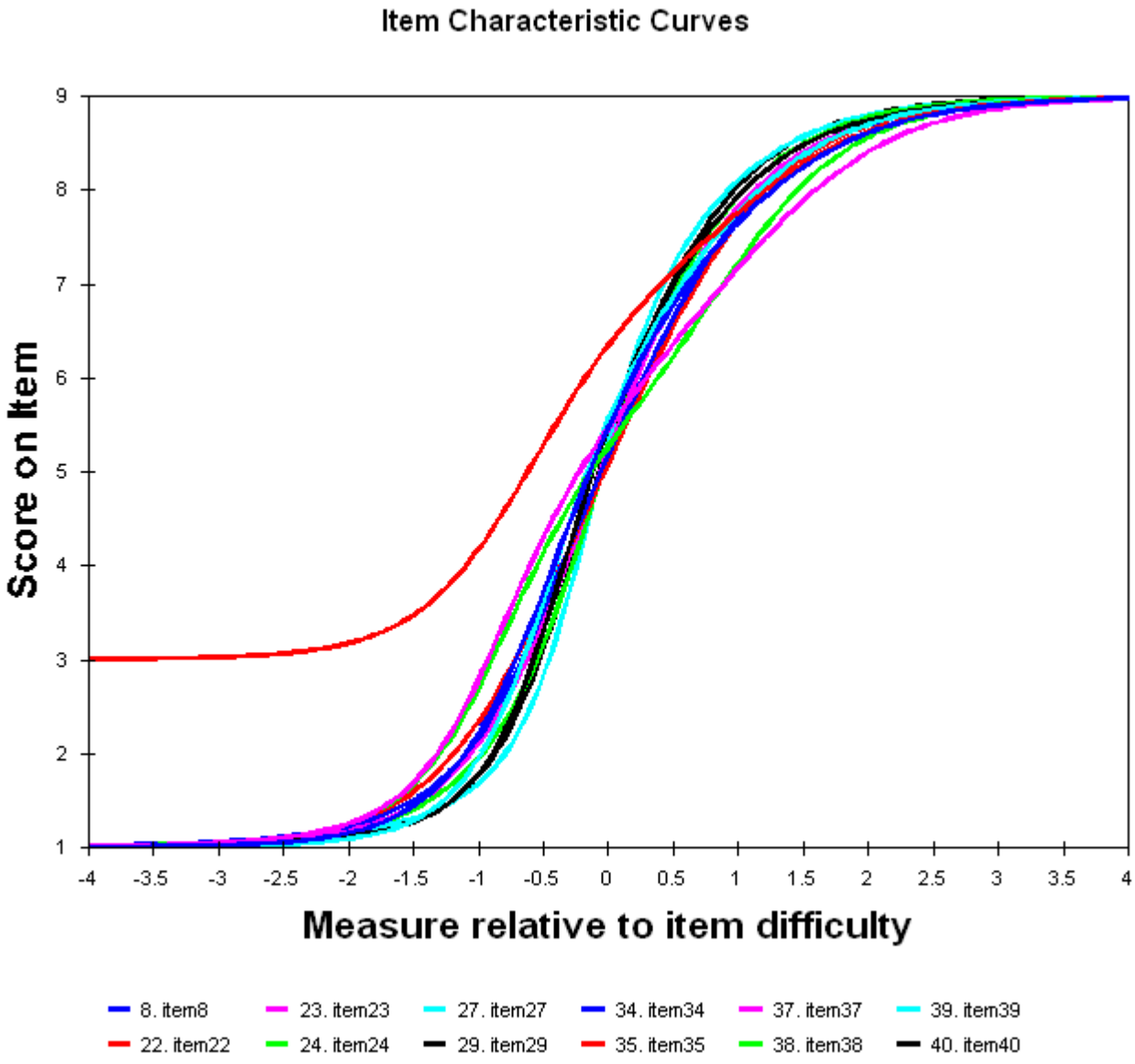


Table 4.1.2b

Functioning Scale: Rasch half-point thresholds

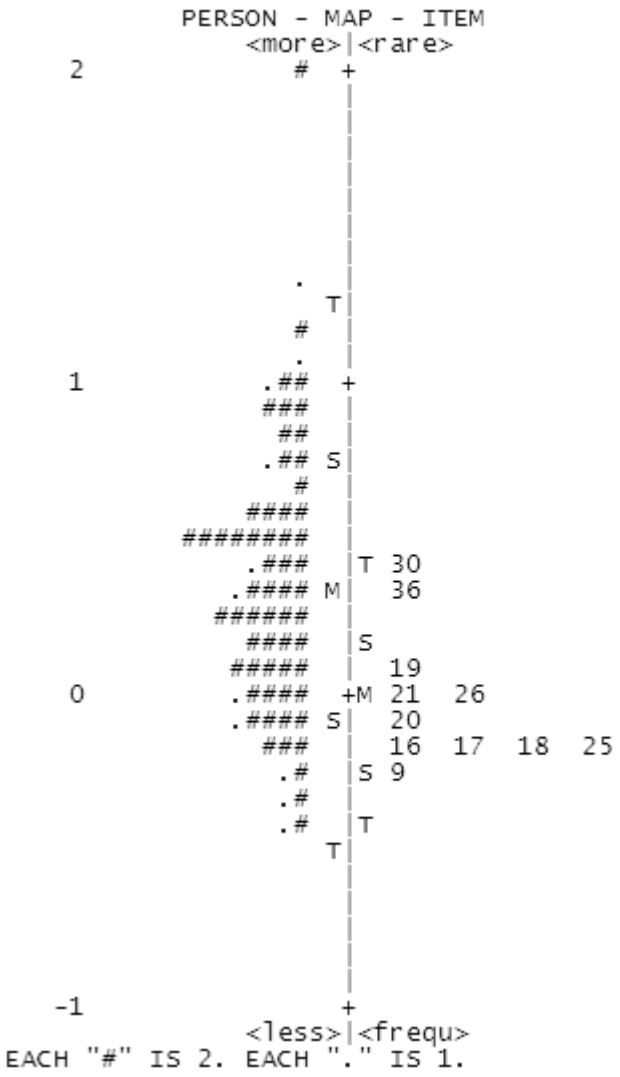
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
8	-1.45	-0.89	-0.57	-0.29	0.01	0.37	0.89	1.81
29	-0.98	-0.48	-0.23	-0.01	0.23	0.54	0.97	1.76
23	-1.27	-0.71	-0.38	-0.04	0.41	0.97	1.61	2.51
40	-1.03	-0.5	-0.23	-0.01	0.22	0.49	0.87	1.58
37	-1.35	-0.72	-0.37	-0.1	0.17	0.48	0.91	1.72
24	-1.6	-1.02	-0.67	-0.3	0.18	0.69	1.2	1.97
38	-1.17	-0.51	-0.19	0.05	0.29	0.57	0.97	1.72
35	-1.55	-0.88	-0.49	-0.15	0.18	0.53	0.96	1.73
27	-1.58	-1.11	-0.84	-0.6	-0.33	0.03	0.52	1.34
22			-1.51	-0.88	-0.45	0.05	0.75	1.75
39	-1.28	-0.69	-0.44	-0.25	-0.06	0.18	0.54	1.3
34	-1.44	-0.78	-0.43	-0.14	0.17	0.52	0.95	1.73

On Rasch half-point step calibrations of the Functioning scale (Figure 4.1.2b; Table 4.1.2b), the Rasch half-point thresholds for items ranged from -1.58 to 2.51. The lowest threshold was the step 1 for item 27 (“how you sleep?”). The highest threshold was the step 8 for item 23 (“the way you communicate with people you don't know well?”).

According to Table 4.1.2b, item 22 had missed values in step 1 and step 2, it was due to the category 1 and 2 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 22 may recommended to be assigned a 7-point Likert option instead.

Figure 4.1.3a

Participation and Physical Health Variable Map



Note:
- 9 your ability to play with friends?
- 16 being able to do the things you want to do?
- 17 your ability to participate at school?
- 18 your ability to participate in recreational activities?
- 19 your ability to participate in sporting activities?
- 20 your ability to participate in social events outside of school?
- 21 your ability to participate in your community?
- 25 your physical health?
- 26 the way you get around?
- 30 your ability to keep up physically with your peers?
- 36 the way you use your legs?

The person-item map (Figure 4.1.3a, Page 96) shows that the Participation and Physical Health scale was relatively “easy” for this sample. The mean for the sample is one and a half standard deviations above the mean for the items. The item difficulties cluster together.

The easiest item (9, “your ability to play with friends?”) is approximately one standard deviation below the item mean. The most difficult item (30, “your ability to keep up physically with your peers?”) is approximately two standard deviations above the item mean.

In the preceding table, item 21 (“your ability to participate in your community?”) and 26 (“the way you get around?”) were just fallen into the mean range of the scale. The item mean, however, was lower than the person mean, and the wide range of person distribution limited the item effects on distinguishing this group of people within the scale.

Furthermore, for the alignment of person-item distributions, around 40% person was above the item limit. These information have two meanings, the first one is that there are so many (40%) high “ability” people (or high satisfactory level in terms of participation and physical health) within this specific group; the second one is that this subscale itself cannot adequately measure those (40%) high “ability” people, possibly, design and insert some high “difficulty” items or amend the content of the redundant items is recommended.

Table 4.1.3a

Participation and Physical Health Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
30	0.4	0.05	0.94	-0.5	1.02	0.2	0.61
9	-0.26	0.05	0.95	-0.3	1	0.1	0.53
16	-0.14	0.05	0.92	-0.5	0.89	-0.7	0.58
21	0.02	0.06	0.92	-0.7	0.91	-0.7	0.59
19	0.09	0.05	0.9	-0.8	0.9	-0.8	0.61
36	0.33	0.05	0.88	-1	0.88	-1	0.65
20	-0.08	0.05	0.85	-1.2	0.85	-1.2	0.61
25	-0.13	0.06	0.81	-1.8	0.81	-1.7	0.65
18	-0.2	0.06	0.8	-1.6	0.79	-1.7	0.63
17	-0.13	0.06	0.76	-2	0.75	-2.1	0.67
26	-0.02	0.05	0.74	-2.4	0.71	-2.5	0.69

The item difficulty logits (Item Score in Table 4.1.3a) ranged from -0.26 to 0.4 on calibrations of the Participation and Physical Health scale. The easiest item was item 9 ("your ability to play with friends?") [-0.26, SE=0.05]. The most difficult item was item 30 ("your ability to keep up physically with your peers?") [0.4, SE=0.05].

On the examination of the fit statistics of the Participation and Physical Health scale, the infit and outfit statistics for items ranged from 0.71 to 1.02. The lowest MNSQ was the outfit for item 26 ("the way you get around?") [0.71, SD=-2.5]. The highest MNSQ was the outfit for item 30 ("your ability to keep up physically with your peers?") [1.02, SD=0.2]

Within this scale, item 30 (“your ability to keep up physically with your peers?”), 9 (“your ability to play with friends?”), 16 (“being able to do the things you want to do?”), 21 (“your ability to participate in your community?”), and 19 (“your ability to participate in sporting activities?”) were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.1.3b

Participation and Physical Health Item Characteristic Curves

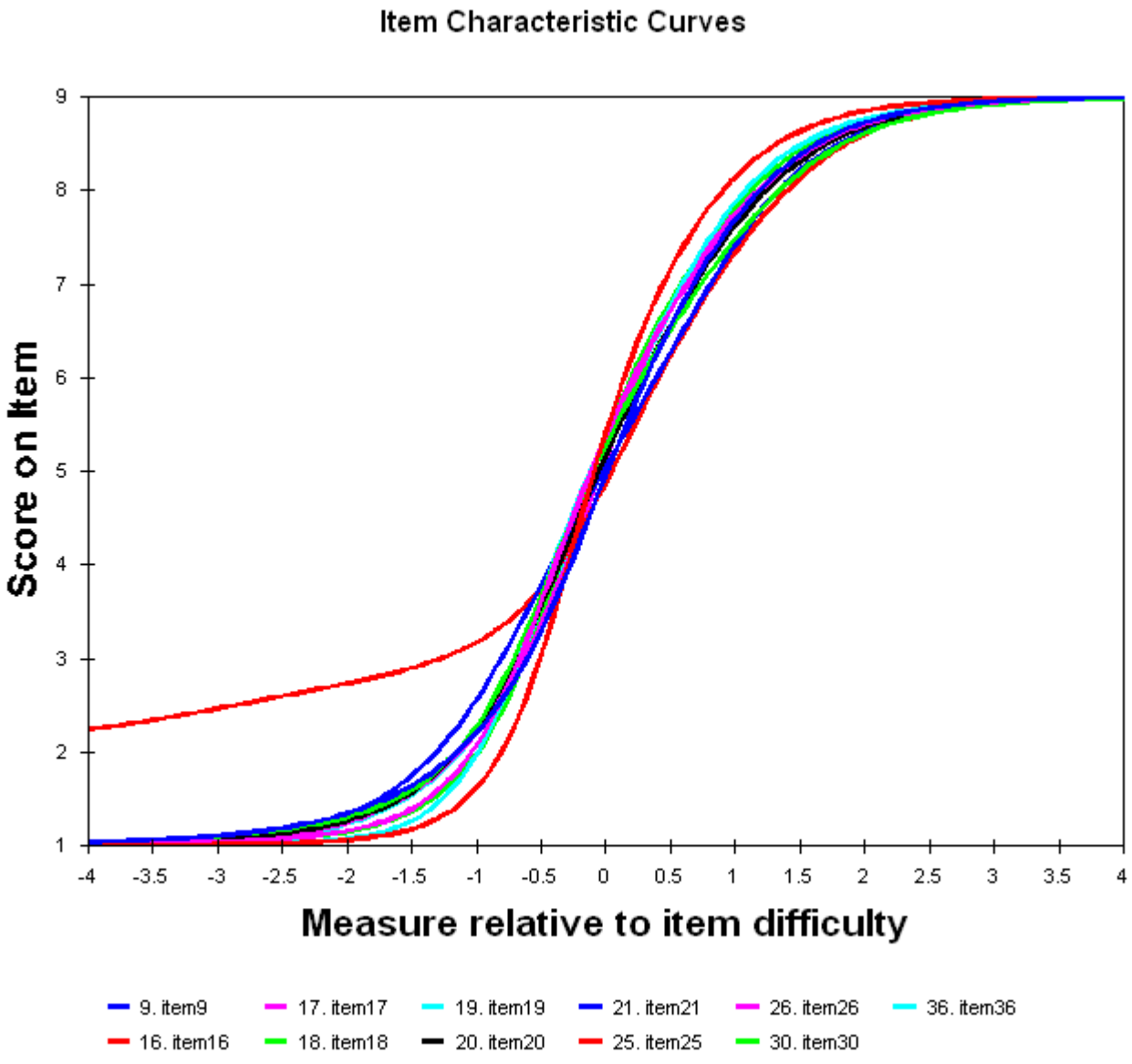


Table 4.1.3b

Participation and Physical Health Scale: Rasch half-point thresholds

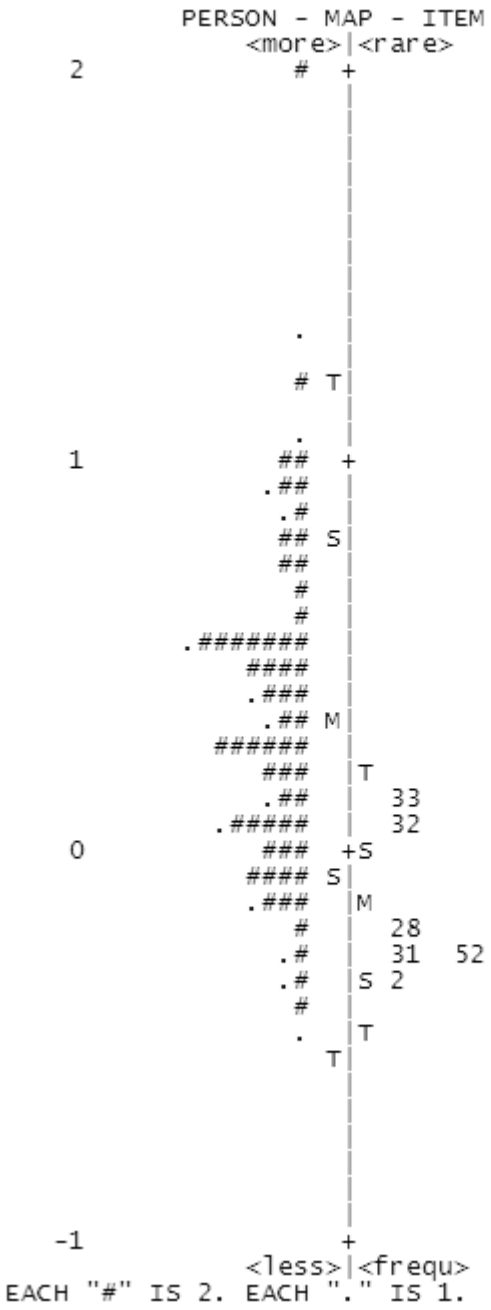
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
30	-0.94	-0.38	-0.09	0.17	0.44	0.77	1.22	1.98
9	-1.96	-1.09	-0.69	-0.39	-0.1	0.22	0.64	1.37
16	-1.22	-0.78	-0.54	-0.33	-0.12	0.14	0.5	1.18
21	-1.73	-1	-0.59	-0.2	0.21	0.62	1.07	1.82
19	-1.17	-0.72	-0.45	-0.18	0.13	0.49	0.9	1.61
36	-1.22	-0.53	-0.15	0.14	0.42	0.74	1.18	2
20	-1.67	-0.96	-0.59	-0.29	0.03	0.4	0.86	1.64
25		-2.99	-0.8	-0.28	0.09	0.48	0.96	1.73
18	-1.82	-1.09	-0.75	-0.45	-0.12	0.29	0.82	1.64
17	-1.51	-0.96	-0.66	-0.39	-0.09	0.27	0.74	1.53
26	-1.59	-0.86	-0.48	-0.19	0.08	0.39	0.83	1.65

On Rasch half-point step calibrations of the Participation and Physical Health scale (Figure 4.1.3b; Table 4.1.3b), the Rasch half-point thresholds for items ranged from -1.96 to 2. The lowest threshold was the step 1 for item 9 (“your ability to play with friends?”). The highest threshold was the step 8 for item 36 (“the way you use your legs?”).

According to Table 4.1.3b, item 25 had missed values in step 1, it was due to the category 1 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 25 may recommended to be assigned an 8-point Likert option instead.

Figure 4.1.4a

Emotional Wellbeing Variable Map



Note: 2 the way you get along with the person who looks after you?

28 the way you look?

31 your life in general?

32 yourself?

33 your future?

52 how happy are you?

The person-item map (Figure 4.1.4a, Page 101) shows that the Emotional Wellbeing scale was relatively “easy” for this sample. The mean for the sample is two standard deviations above the mean for the items. The item difficulties cluster together.

The easiest item (2, “the way you get along with the person who looks after you?”) is approximately one standard deviation below the item mean. The most difficult item (33, “your future?”) is approximately one and a half standard deviations above the item mean.

The item mean was far lower than the person mean, and the wide range of person distribution revealed that there are only limited effects on distinguishing this group of people within the scale.

Moreover, for the alignment of person-item distributions, around 60% person was above the item limit. These information have two meanings, the first one is that there are so many (60%) high “ability” people (or high satisfactory level in terms of emotional wellbeing) within this group of samples; the second one is that this subscale itself cannot adequately measure those (60%) high “ability” people, possibly, design and insert some high “difficulty” items or amend the content of the redundant items is recommended.

Table 4.1.4a

Emotional Wellbeing Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
2	-0.32	0.07	0.89	-0.6	0.86	-0.9	0.49
28	-0.23	0.07	0.83	-1.1	0.83	-1.2	0.58
33	0.1	0.05	0.77	-2	0.77	-2	0.68
52	-0.28	0.06	0.72	-2.2	0.69	-2.5	0.68
32	0.1	0.06	0.67	-3.1	0.68	-3	0.73
31	-0.23	0.06	0.66	-2.6	0.64	-3	0.72

The item difficulty logits (Item Score in Table 4.1.4a) ranged from -0.32 to 0.1 on calibrations of the Emotional Wellbeing scale. The easiest item was item 2 ("the way you get along with the person who looks after you?") [-0.32, SE=0.07]. The most difficult item was item 32 ("yourself?") [0.1, SE=0.06].

On the examination of the fit statistics of the Emotional Wellbeing scale, the infit and outfit statistics for items ranged from 0.64 to 0.89. The lowest MNSQ was the outfit for item 31 ("your life in general?") [0.64, SD=-3]. The highest MNSQ was the infit for item 2 ("the way you get along with the person who looks after you?") [0.89, SD=-0.6]

Figure 4.1.4b

Emotional Wellbeing Item Characteristic Curves

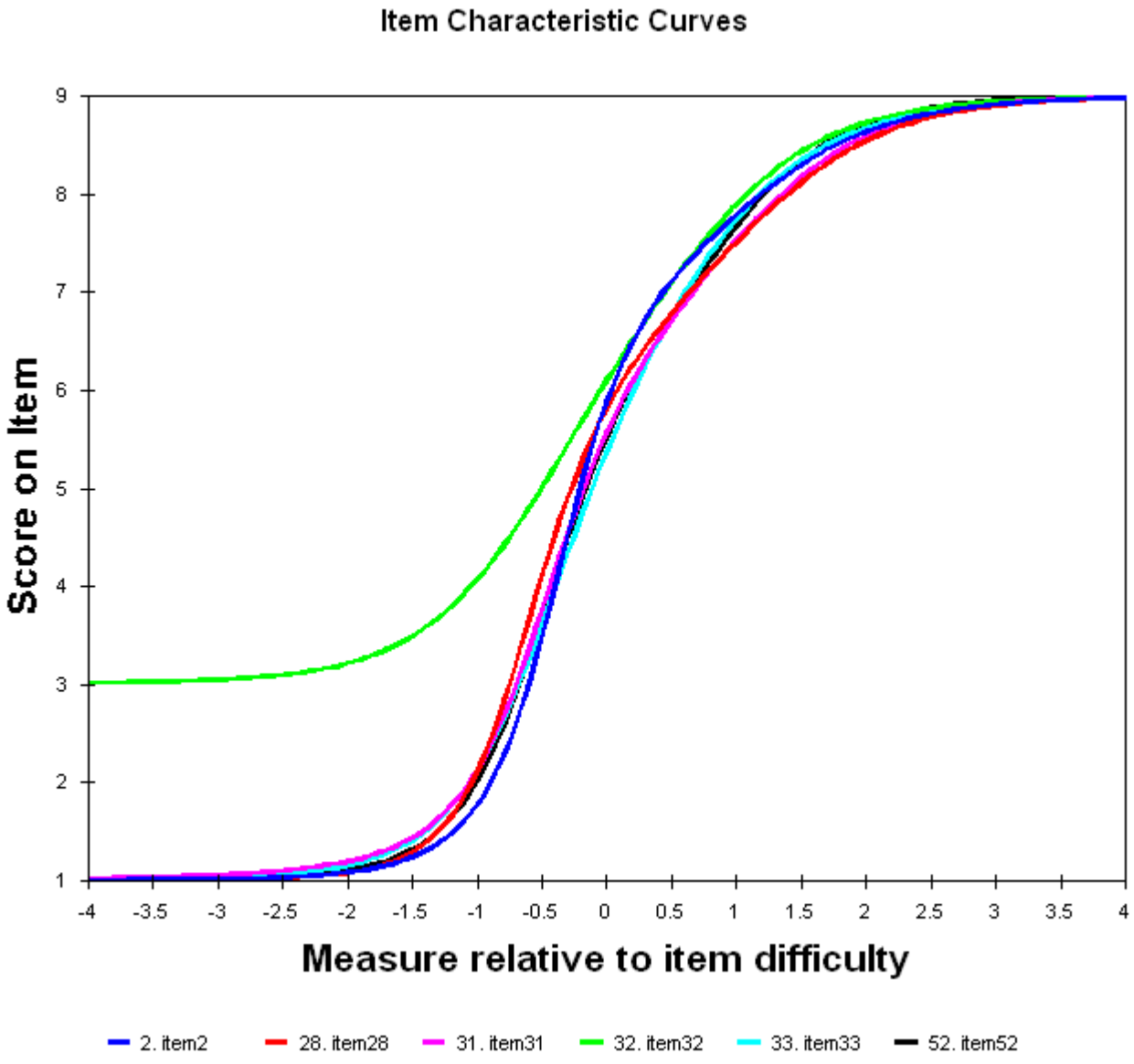


Table 4.1.4b

Emotional Wellbeing Scale: Rasch half-point thresholds

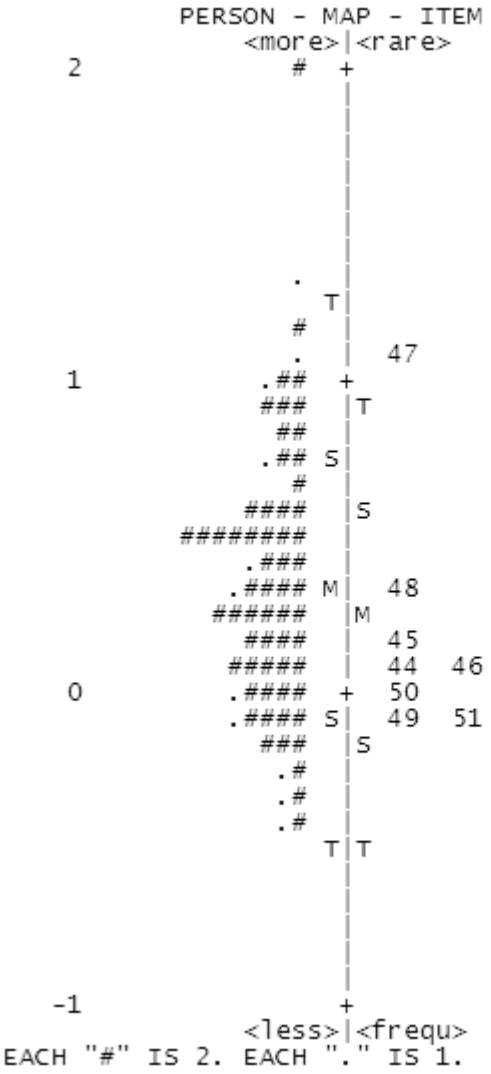
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
2	-1.51	-1.05	-0.82	-0.63	-0.42	-0.13	0.44	1.44
28	-1.53	-1.11	-0.87	-0.64	-0.35	0.1	0.76	1.7
33	-1.3	-0.74	-0.43	-0.16	0.14	0.5	0.96	1.79
52	-1.59	-1.1	-0.82	-0.58	-0.29	0.11	0.62	1.39
32			-1.39	-0.66	-0.18	0.29	0.83	1.68
31	-1.66	-1.09	-0.8	-0.55	-0.26	0.15	0.74	1.63

On Rasch half-point step calibrations of the Emotional Wellbeing scale (Figure 4.1.4b; Table 4.1.4b), the Rasch half-point thresholds for items ranged from -1.66 to 1.79. The lowest threshold was the step 1 for item 31 ("your life in general?"). The highest threshold was the step 8 for item 33 ("your future?").

According to Table 4.1.4b, item 32 had missed values in step 1 and step 2, it was due to the category 1 and 2 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 32 may recommended to be assigned a 7-point Likert option instead.

Figure 4.1.5a

Pain and Impact of Disability Variable Map



Note: 44 are you bothered by hospital visits?

45 are you bothered when you miss school for health reasons?

46 are you bothered by being handled by other people?

47 do you worry about who will take care of you in the future?

48 are you concerned about having cerebral palsy?

49 how much pain do you have?

50 how do you feel about the amount of pain you have?

51 how much discomfort do you experience?

The person-item map (Figure 4.1.5a, Page 106) shows that the Pain and Impact of Disability scale was relatively “easy” for this sample. The mean for the sample is one fourth standard deviation above the mean for the items. The item difficulties except item 47, cluster together.

The easiest items (49, “how much pain do you have?”; 51, “how much discomfort do you experience?”) are approximately one standard deviation below the item mean. The most difficult item (47, “do you worry about who will take care of you in the future?”) is approximately two standard deviations above the item mean.

In details, for the alignment of person-item distributions, around 15% person was below the item limit. These information have two meanings, the first one is that there are some (15%) low “ability” people (or low satisfactory level in terms of pain and impact of disability) within this specific group; the second one is that this subscale itself cannot adequately measure those (15%) low “ability” people, possibly, design and insert some low “difficulty” items or amend the content of the redundant items is recommended. Noticeably, there is a big gap between item 47 and item 48; design and insert some items in order to fill this gap is also recommended

Table 4.1.5a

Pain and Impact of Disability Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
48	0.29	0.04	2.01	7	4.84	9.9	0.23
45	0.19	0.04	1.99	6.9	2.5	7.8	0.22
46	0.12	0.04	1.83	5.9	2.19	7	0.23
44	0.12	0.05	1.56	4.2	1.61	4	0.36
47	1.08	0.07	1.42	3.4	1.39	2.9	0.16
50	0.01	0.05	1.23	1.8	1.22	1.5	0.48
51	-0.04	0.05	1.18	1.5	1.22	1.5	0.48
49	-0.07	0.05	1.15	1.3	1.12	0.8	0.5

The item difficulty logits (Item Score in Table 4.1.5a) ranged from -0.07 to 1.08 on calibrations of the Pain and Impact of Disability scale. The easiest item was item 49 (“how much pain do you have?”) [-0.07, SE=0.05]. The most difficult item was item 47 (“do you worry about who will take care of you in the future?”) [1.08, SE=0.07].

On the examination of the fit statistics of the Pain and Impact of Disability scale, the infit and outfit statistics for items ranged from 1.12 to 4.84. The lowest MNSQ was the outfit for item 49 (“how much pain do you have?”) [1.12, SD=0.8]. The highest MNSQ was the outfit for item 48 (“are you concerned about having cerebral palsy?”) [4.84, SD=9.9]

Item 48 (“are you concerned about having cerebral palsy?”) [infit, 2.01, outfit, 4.84], 45 (“are you bothered when you miss school for health reasons?”) [infit, 1.99, outfit, 2.5], 46 (“are you

bothered by being handled by other people?") [infit, 1.83, outfit, 2.19], 44 ("are you bothered by hospital visits?") [infit, 1.56, outfit, 1.61] were out of the Likert scale critical range (i.e., 0.5-1.5). Item 48, 45, 46 and 47, had a slightly low score to total correlation (r=0.23, 0.22, 0.23, 0.16). In order to properly fit the criteria of unidimensionality, item 48, 45 and 46 may recommended to be separated from other items in the scale and interpreted accordingly.

Figure 4.1.5b

Pain and Impact of Disability Item Characteristic Curves

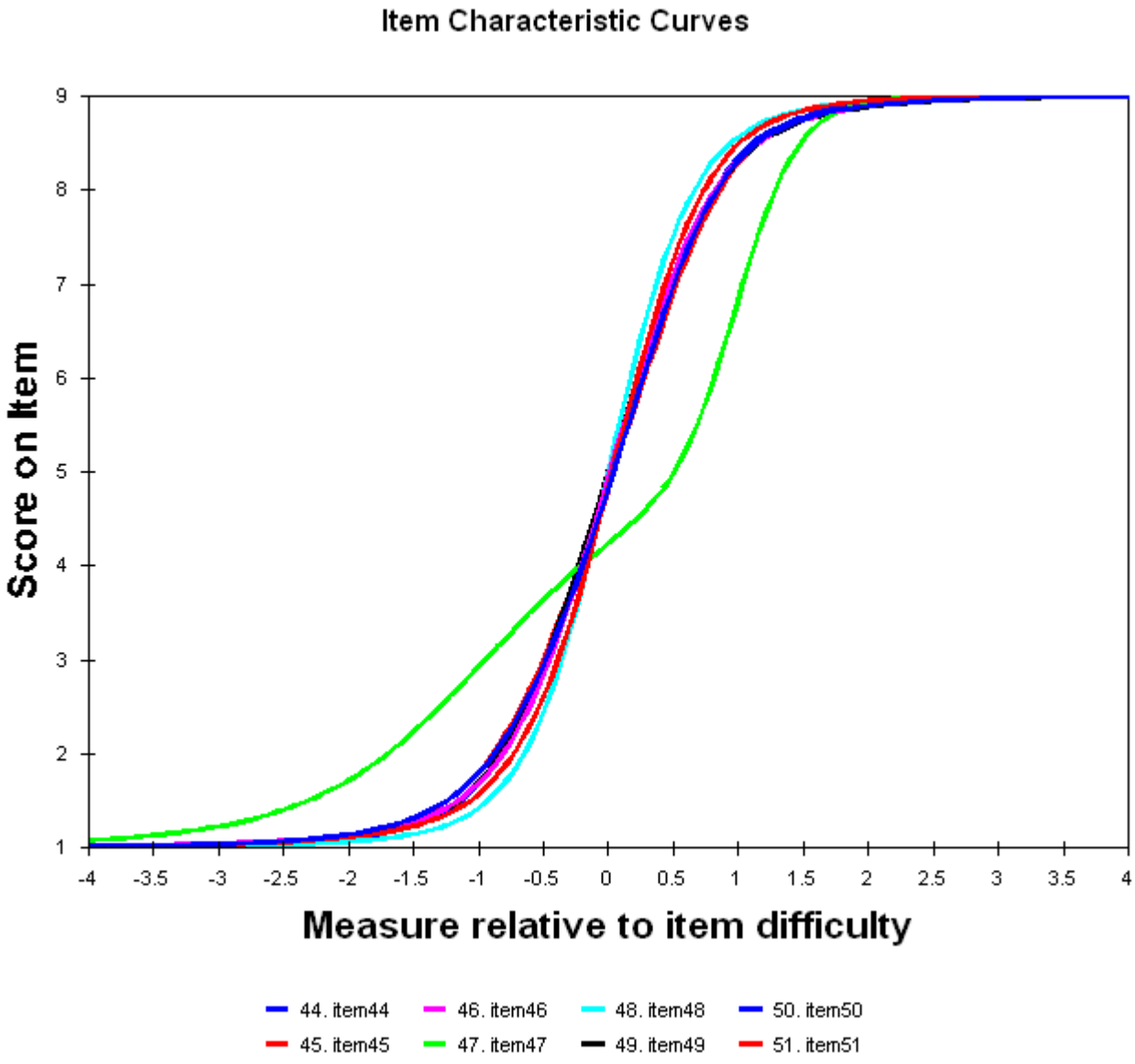


Table 4.1.5b

Pain and Impact of Disability Scale: Rasch half-point thresholds

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
48	-0.63	-0.19	0.03	0.21	0.38	0.55	0.78	1.23
45	-0.88	-0.35	-0.08	0.12	0.31	0.51	0.75	1.2
46	-1.04	-0.48	-0.2	0.03	0.24	0.47	0.74	1.27
44	-1.12	-0.53	-0.21	0.05	0.28	0.5	0.77	1.27
47	-1.24	-0.22	0.48	1.29	1.76	2	2.22	2.55
50	-1.12	-0.62	-0.34	-0.1	0.14	0.37	0.64	1.12
51	-1.28	-0.72	-0.4	-0.13	0.11	0.35	0.63	1.12
49	-1.22	-0.71	-0.42	-0.19	0.04	0.27	0.56	1.09

On Rasch half-point step calibrations of the Pain and Impact of Disability scale (Figure 4.1.5b; Table 4.1.5b), the Rasch half-point thresholds for items ranged from -1.28 to 2.55. The lowest threshold was the step 1 for item 51 (“how much discomfort do you experience?”). The highest threshold was the step 8 for item 47 (“do you worry about who will take care of you in the future?”).

4.2. Parent proxy report

For the parent proxy report, after completing the calibration process, the results showed that, in item 1 (“the way they get along with people, generally?”), the item difficulty was -0.25, SE=0.08, the infit MNSQ was 0.95, SD=-0.3, the outfit MNSQ was 0.97, SD=-0.2, and the point-measure correlation was 0.46; in item 2 (“the way they get along with you?”), the item difficulty was -0.45, SE=0.08, the infit MNSQ was 1.01, SD=0.1, the outfit MNSQ was 1.03, SD=0.3, and the point-measure correlation was 0.37; in item 3 (“the way they get along with their brothers and sisters?”), the item difficulty was -0.18, SE=0.08, the infit MNSQ was 1.04, SD=0.4, the outfit MNSQ was 1.04, SD=0.4, and the point-measure correlation was 0.36; in item 4 (“the way they get along with other children at preschool or school?”), the item difficulty was -0.23, SE=0.08, the infit MNSQ was 0.96, SD=-0.3, the outfit MNSQ was 0.97, SD=-0.2, and the point-measure correlation was 0.43; in item 5 (“the way they get along with other children outside of preschool or school?”), the item difficulty was -0.23, SE=0.07, the infit MNSQ was 0.91, SD=-0.6, the outfit MNSQ was 0.9, SD=-0.7, and the point-measure correlation was 0.53; in item 6 (“the way they get along with adults?”), the item difficulty was -0.33, SE=0.08, the infit MNSQ was 0.88, SD=-0.7, the outfit MNSQ was 0.87, SD=-0.9, and the point-measure correlation was 0.51; in item 7 (“the way they get along with their teachers and/or careers?”), the item difficulty was -0.45, SE=0.07, the infit MNSQ was 0.98, SD=-0.1, the outfit MNSQ was 1.02, SD=0.2, and the point-measure correlation was 0.43; in item 8

("their ability to play on their own?"), the item difficulty was -0.21, SE=0.06, the infit MNSQ was 0.95, SD=-0.3, the outfit MNSQ was 1, SD=0, and the point-measure correlation was 0.53; in item 9 ("their ability to play with friends?"), the item difficulty was -0.24, SE=0.06, the infit MNSQ was 0.88, SD=-1, the outfit MNSQ was 0.88, SD=-1, and the point-measure correlation was 0.58; in item 10 ("going out on trips with the family?"), the item difficulty was -0.62, SE=0.07, the infit MNSQ was 0.98, SD=-0.1, the outfit MNSQ was 0.94, SD=-0.4, and the point-measure correlation was 0.47; in item 11 ("how they are accepted by their family?"), the item difficulty was -0.6, SE=0.07, the infit MNSQ was 1.36, SD=2.2, the outfit MNSQ was 1.32, SD=2.1, and the point-measure correlation was 0.23; in item 12 ("how they are accepted by other children at preschool or school?"), the item difficulty was -0.16, SE=0.08, the infit MNSQ was 0.99, SD=0, the outfit MNSQ was 0.99, SD=0, and the point-measure correlation was 0.43; in item 13 ("how they are accepted by other children outside of preschool or school?"), the item difficulty was -0.1, SE=0.07, the infit MNSQ was 0.88, SD=-0.9, the outfit MNSQ was 0.86, SD=-1.1, and the point-measure correlation was 0.56; in item 14 ("how they are accepted by adults?"), the item difficulty was -0.32, SE=0.07, the infit MNSQ was 1.04, SD=0.3, the outfit MNSQ was 0.98, SD=-0.1, and the point-measure correlation was 0.43; in item 15 ("how they are accepted by people in general?"), the item difficulty was -0.03, SE=0.07, the infit MNSQ was 0.82, SD=-1.4, the outfit MNSQ was 0.78, SD=-1.8, and the point-measure correlation was 0.61; in item 16 ("being able to do the things they want to do?"),

the item difficulty was -0.05, SE=0.05, the infit MNSQ was 0.94, SD=-0.4, the outfit MNSQ was 0.88, SD=-0.8, and the point-measure correlation was 0.6; in item 17 ("their ability to participate at preschool or school?"), the item difficulty was -0.04, SE=0.06, the infit MNSQ was 0.7, SD=-2.1, the outfit MNSQ was 0.71, SD=-2.1, and the point-measure correlation was 0.67; in item 18 ("their ability to participate in recreational activities?"), the item difficulty was -0.02, SE=0.05, the infit MNSQ was 0.74, SD=-2.1, the outfit MNSQ was 0.72, SD=-2.2, and the point-measure correlation was 0.69; in item 19 ("their ability to participate in sporting activities?"), the item difficulty was 0.14, SE=0.05, the infit MNSQ was 0.91, SD=-0.7, the outfit MNSQ was 0.9, SD=-0.8, and the point-measure correlation was 0.57; in item 20 ("their ability to participate in social events outside of preschool or school?"), the item difficulty was 0.02, SE=0.05, the infit MNSQ was 0.78, SD=-1.8, the outfit MNSQ was 0.79, SD=-1.6, and the point-measure correlation was 0.66; in item 21 ("their ability to participate in their community?"), the item difficulty was 0.16, SE=0.05, the infit MNSQ was 0.86, SD=-1.1, the outfit MNSQ was 0.88, SD=-0.9, and the point-measure correlation was 0.61; in item 22 ("the way they communicate with people they know well?"), the item difficulty was -0.19, SE=0.06, the infit MNSQ was 0.93, SD=-0.4, the outfit MNSQ was 0.96, SD=-0.2, and the point-measure correlation was 0.52; in item 23 ("the way they communicate with people they don't know well?"), the item difficulty was 0.31, SE=0.06, the infit MNSQ was 0.97, SD=-0.2, the outfit MNSQ was 1.05, SD=0.4, and the point-measure correlation was 0.51; in item 24 ("the way other

people communicate with them?"), the item difficulty was 0.16, SE=0.06, the infit MNSQ was 0.78, SD=-1.6, the outfit MNSQ was 0.8, SD=-1.4, and the point-measure correlation was 0.61; in item 25 ("their physical health?"), the item difficulty was -0.15, SE=0.06, the infit MNSQ was 0.83, SD=-1.5, the outfit MNSQ was 0.85, SD=-1.3, and the point-measure correlation was 0.61; in item 26 ("the way they get around?"), the item difficulty was -0.04, SE=0.05, the infit MNSQ was 0.78, SD=-1.9, the outfit MNSQ was 0.74, SD=-2.2, and the point-measure correlation was 0.68; in item 27 ("how they sleep?"), the item difficulty was -0.45, SE=0.07, the infit MNSQ was 0.97, SD=-0.1, the outfit MNSQ was 0.98, SD=-0.1, and the point-measure correlation was 0.48; in item 28 ("the way they look?"), the item difficulty was -0.4, SE=0.07, the infit MNSQ was 0.75, SD=-1.7, the outfit MNSQ was 0.75, SD=-1.9, and the point-measure correlation was 0.6; in item 29 ("their ability to keep up academically with their peers?"), the item difficulty was 0.31, SE=0.06, the infit MNSQ was 1.06, SD=0.6, the outfit MNSQ was 1.13, SD=1.1, and the point-measure correlation was 0.48; in item 30 ("their ability to keep up physically with their peers?"), the item difficulty was 0.46, SE=0.05, the infit MNSQ was 0.97, SD=-0.3, the outfit MNSQ was 0.98, SD=-0.1, and the point-measure correlation was 0.56; in item 31 ("their life in general?"), the item difficulty was -0.19, SE=0.07, the infit MNSQ was 0.74, SD=-1.9, the outfit MNSQ was 0.76, SD=-1.7, and the point-measure correlation was 0.65; in item 32 ("themselves?"), the item difficulty was -0.04, SE=0.06, the infit MNSQ was 0.75, SD=-1.7, the outfit MNSQ was 0.78, SD=-1.5, and the

point-measure correlation was 0.63; in item 33 (“their future?”), the item difficulty was 0.37, SE=0.06, the infit MNSQ was 0.92, SD=-0.6, the outfit MNSQ was 0.92, SD=-0.5, and the point-measure correlation was 0.55; in item 34 (“their opportunities in life?”), the item difficulty was 0.36, SE=0.06, the infit MNSQ was 0.9, SD=-0.7, the outfit MNSQ was 0.92, SD=-0.5, and the point-measure correlation was 0.55; in item 35 (“the way they use their arms?”), the item difficulty was 0.1, SE=0.05, the infit MNSQ was 0.88, SD=-1, the outfit MNSQ was 0.9, SD=-0.8, and the point-measure correlation was 0.6; in item 36 (“the way they use their legs?”), the item difficulty was 0.33, SE=0.05, the infit MNSQ was 0.84, SD=-1.6, the outfit MNSQ was 0.85, SD=-1.4, and the point-measure correlation was 0.65; in item 37 (“the way they use their hands?”), the item difficulty was 0.13, SE=0.05, the infit MNSQ was 0.8, SD=-1.7, the outfit MNSQ was 0.8, SD=-1.7, and the point-measure correlation was 0.65; in item 38 (“their ability to dress themselves?”), the item difficulty was 0.4, SE=0.05, the infit MNSQ was 0.93, SD=-0.6, the outfit MNSQ was 0.94, SD=-0.5, and the point-measure correlation was 0.61; in item 39 (“their ability to eat or drink independently?”), the item difficulty was -0.03, SE=0.05, the infit MNSQ was 0.84, SD=-1.3, the outfit MNSQ was 0.85, SD=-1.1, and the point-measure correlation was 0.65; in item 40 (“their ability to use the toilet by themselves?”), the item difficulty was 0.38, SE=0.05, the infit MNSQ was 0.92, SD=-0.7, the outfit MNSQ was 0.95, SD=-0.4, and the point-measure correlation was 0.6; in item 41 (“the special equipment they have at home?”), the item difficulty was -0.08, SE=0.08,

the infit MNSQ was 0.92, SD=-0.5, the outfit MNSQ was 0.92, SD=-0.5, and the point-measure correlation was 0.53; in item 42 ("the special equipment they have at their school?"), the item difficulty was 0.05, SE=0.07, the infit MNSQ was 0.96, SD=-0.2, the outfit MNSQ was 0.99, SD=0, and the point-measure correlation was 0.51; in item 43 ("the special equipment that is available in the community?"), the item difficulty was 0.24, SE=0.06, the infit MNSQ was 1.01, SD=0.1, the outfit MNSQ was 1.08, SD=0.6, and the point-measure correlation was 0.52; in item 44 ("is your child bothered by hospital visits?"), the item difficulty was 0.02, SE=0.04, the infit MNSQ was 1.5, SD=4, the outfit MNSQ was 1.95, SD=6.5, and the point-measure correlation was 0.34; in item 45 ("is your child bothered when they miss school for health reasons?"), the item difficulty was 0.08, SE=0.04, the infit MNSQ was 1.73, SD=5.5, the outfit MNSQ was 2.1, SD=6.9, and the point-measure correlation was 0.32; in item 46 ("is your child bothered by being handled by other people?"), the item difficulty was -0.14, SE=0.05, the infit MNSQ was 1.34, SD=2.8, the outfit MNSQ was 1.41, SD=3.2, and the point-measure correlation was 0.4; in item 47 ("does your child worry about who will take care of them in the future?"), the item difficulty was 0.01, SE=0.07, the infit MNSQ was 1.27, SD=2.6, the outfit MNSQ was 1.43, SD=3.7, and the point-measure correlation was 0.2; in item 48 ("is your child concerned about having cerebral palsy?"), the item difficulty was 0.17, SE=0.04, the infit MNSQ was 1.8, SD=6.1, the outfit MNSQ was 1.94, SD=6.4, and the point-measure correlation was 0.29; in item 49 ("how much pain does your child have?"), the

item difficulty was -0.01, SE=0.05, the infit MNSQ was 1.47, SD=3.9, the outfit MNSQ was 1.8, SD=5.8, and the point-measure correlation was 0.34; in item 50 (“how does your child feel about the amount of pain they have?”), the item difficulty was 0.1, SE=0.04, the infit MNSQ was 1.42, SD=3.5, the outfit MNSQ was 1.63, SD=4.8, and the point-measure correlation was 0.4; in item 51 (“how much discomfort does your child experience?”), the item difficulty was 0.05, SE=0.04, the infit MNSQ was 1.44, SD=3.6, the outfit MNSQ was 1.72, SD=5.2, and the point-measure correlation was 0.38; in item 52 (“how happy is your child?”), the item difficulty was -0.35, SE=0.07, the infit MNSQ was 0.85, SD=-1, the outfit MNSQ was 0.87, SD=-0.9, and the point-measure correlation was 0.59; in item 53 (“your child's access to treatment?”), the item difficulty was -0.23, SE=0.07, the infit MNSQ was 0.97, SD=-0.2, the outfit MNSQ was 0.91, SD=-0.7, and the point-measure correlation was 0.51; in item 54 (“your child's access to therapy?”), the item difficulty was -0.17, SE=0.06, the infit MNSQ was 0.98, SD=-0.1, the outfit MNSQ was 0.9, SD=-0.7, and the point-measure correlation was 0.54; in item 55 (“your child's access to specialised medical or surgical care?”), the item difficulty was 0.06, SE=0.05, the infit MNSQ was 1, SD=0, the outfit MNSQ was 0.97, SD=-0.2, and the point-measure correlation was 0.55; in item 56 (“your ability to get advice from a paediatrician?”), the item difficulty was -0.04, SE=0.06, the infit MNSQ was 1.04, SD=0.4, the outfit MNSQ was 1.05, SD=0.5, and the point-measure correlation was 0.52; in item 57 (“your access to respite care?”), the item difficulty was 0.31, SE=0.11, the infit MNSQ was

0.95, SD=-0.2, the outfit MNSQ was 0.95, SD=-0.1, and the point-measure correlation was 0.49; in item 58 (“the amount of respite care you receive?”), the item difficulty was 0.3, SE=0.1, the infit MNSQ was 0.95, SD=-0.1, the outfit MNSQ was 0.97, SD=-0.1, and the point-measure correlation was 0.51; in item 59 (“how easy it is to get respite?”), the item difficulty was 0.39, SE=0.1, the infit MNSQ was 1.08, SD=0.4, the outfit MNSQ was 1.11, SD=0.5, and the point-measure correlation was 0.42; in item 60 (“your child's access to community services and facilities?”), the item difficulty was 0.55, SE=0.05, the infit MNSQ was 0.9, SD=-1, the outfit MNSQ was 0.9, SD=-0.9, and the point-measure correlation was 0.65; in item 61 (“your child's access to extra help with learning at preschool or school?”), the item difficulty was 0.27, SE=0.05, the infit MNSQ was 1, SD=0, the outfit MNSQ was 1, SD=0, and the point-measure correlation was 0.57; in item 62 (“your physical health?”), the item difficulty was 0.03, SE=0.06, the infit MNSQ was 1.04, SD=0.4, the outfit MNSQ was 1.01, SD=0.1, and the point-measure correlation was 0.48; in item 63 (“your work situation?”), the item difficulty was 0.23, SE=0.06, the infit MNSQ was 1.04, SD=0.4, the outfit MNSQ was 1.03, SD=0.3, and the point-measure correlation was 0.5; in item 64 (“your family's financial situation?”), the item difficulty was 0.4, SE=0.05, the infit MNSQ was 1.2, SD=1.6, the outfit MNSQ was 1.22, SD=1.7, and the point-measure correlation was 0.42; in item 65 (“how happy are you?”), the item difficulty was 0.15, SE=0.05, the infit MNSQ was 0.95, SD=-0.3, the outfit MNSQ was 0.94, SD=-0.5, and the point-measure correlation was 0.59(Table 4.2a).

Table 4.2a

C CP-QOL Child: Parent Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
1	-0.25	0.08	0.95	-0.3	0.97	-0.2	0.46
2	-0.45	0.08	1.01	0.1	1.03	0.3	0.37
3	-0.18	0.08	1.04	0.4	1.04	0.4	0.36
4	-0.23	0.08	0.96	-0.3	0.97	-0.2	0.43
5	-0.23	0.07	0.91	-0.6	0.9	-0.7	0.53
6	-0.33	0.08	0.88	-0.7	0.87	-0.9	0.51
7	-0.45	0.07	0.98	-0.1	1.02	0.2	0.43
8	-0.21	0.06	0.95	-0.3	1	0	0.53
9	-0.24	0.06	0.88	-1	0.88	-1	0.58
10	-0.62	0.07	0.98	-0.1	0.94	-0.4	0.47
11	-0.6	0.07	1.36	2.2	1.32	2.1	0.23
12	-0.16	0.08	0.99	0	0.99	0	0.43
13	-0.1	0.07	0.88	-0.9	0.86	-1.1	0.56
14	-0.32	0.07	1.04	0.3	0.98	-0.1	0.43
15	-0.03	0.07	0.82	-1.4	0.78	-1.8	0.61
16	-0.05	0.05	0.94	-0.4	0.88	-0.8	0.6
17	-0.04	0.06	0.7	-2.1	0.71	-2.1	0.67
18	-0.02	0.05	0.74	-2.1	0.72	-2.2	0.69
19	0.14	0.05	0.91	-0.7	0.9	-0.8	0.57
20	0.02	0.05	0.78	-1.8	0.79	-1.6	0.66
21	0.16	0.05	0.86	-1.1	0.88	-0.9	0.61
22	-0.19	0.06	0.93	-0.4	0.96	-0.2	0.52
23	0.31	0.06	0.97	-0.2	1.05	0.4	0.51
24	0.16	0.06	0.78	-1.6	0.8	-1.4	0.61
25	-0.15	0.06	0.83	-1.5	0.85	-1.3	0.61
26	-0.04	0.05	0.78	-1.9	0.74	-2.2	0.68
27	-0.45	0.07	0.97	-0.1	0.98	-0.1	0.48
28	-0.4	0.07	0.75	-1.7	0.75	-1.9	0.6
29	0.31	0.06	1.06	0.6	1.13	1.1	0.48
30	0.46	0.05	0.97	-0.3	0.98	-0.1	0.56
31	-0.19	0.07	0.74	-1.9	0.76	-1.7	0.65
32	-0.04	0.06	0.75	-1.7	0.78	-1.5	0.63
33	0.37	0.06	0.92	-0.6	0.92	-0.5	0.55

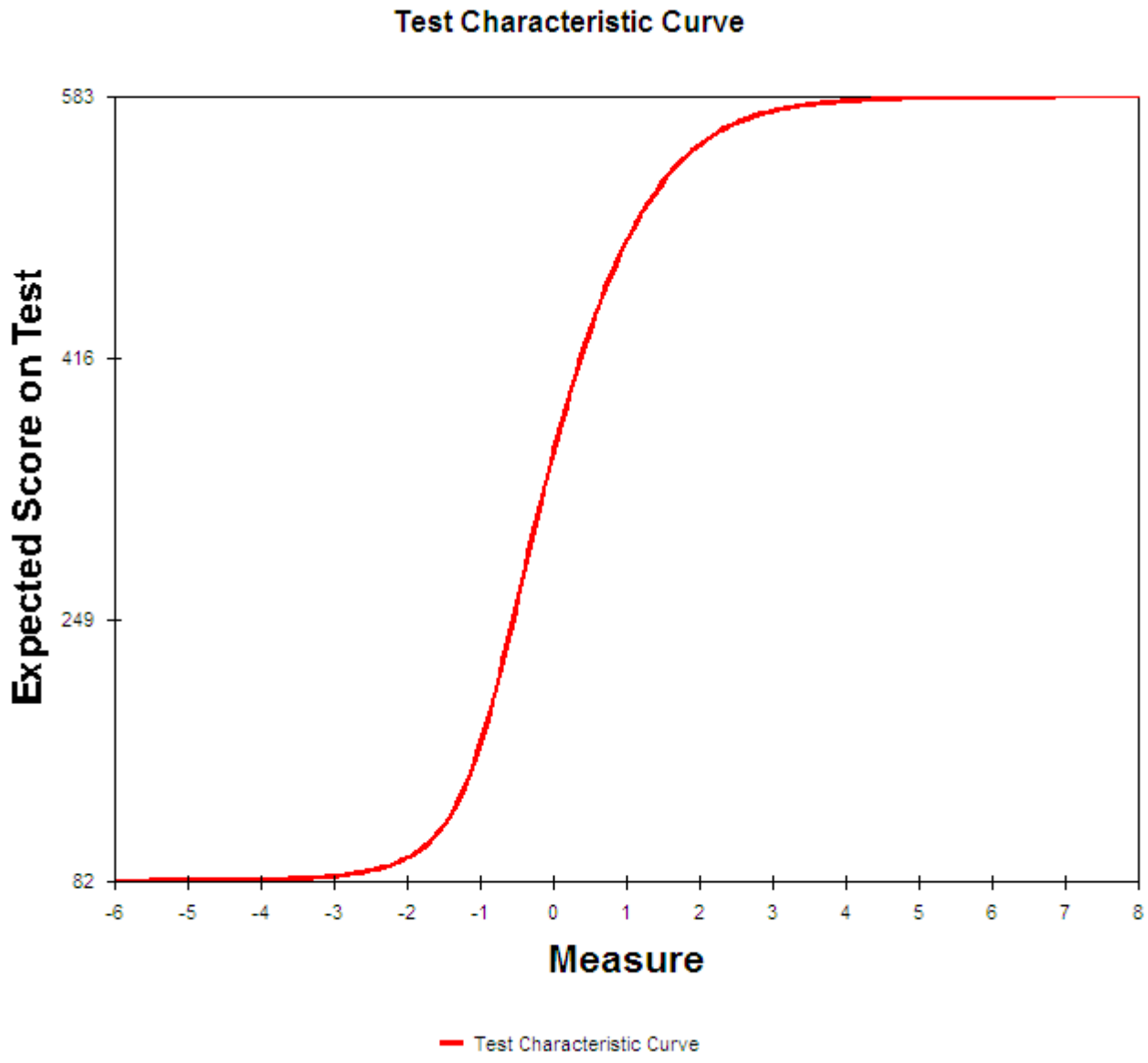
Table 4.2a *(Continued)*

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
34	0.36	0.06	0.9	-0.7	0.92	-0.5	0.55
35	0.1	0.05	0.88	-1	0.9	-0.8	0.6
36	0.33	0.05	0.84	-1.6	0.85	-1.4	0.65
37	0.13	0.05	0.8	-1.7	0.8	-1.7	0.65
38	0.4	0.05	0.93	-0.6	0.94	-0.5	0.61
39	-0.03	0.05	0.84	-1.3	0.85	-1.1	0.65
40	0.38	0.05	0.92	-0.7	0.95	-0.4	0.6
41	-0.08	0.08	0.92	-0.5	0.92	-0.5	0.53
42	0.05	0.07	0.96	-0.2	0.99	0	0.51
43	0.24	0.06	1.01	0.1	1.08	0.6	0.52
44	0.02	0.04	1.5	4	1.95	6.5	0.34
45	0.08	0.04	1.73	5.5	2.1	6.9	0.32
46	-0.14	0.05	1.34	2.8	1.41	3.2	0.4
47	0.01	0.07	1.27	2.6	1.43	3.7	0.2
48	0.17	0.04	1.8	6.1	1.94	6.4	0.29
49	-0.01	0.05	1.47	3.9	1.8	5.8	0.34
50	0.1	0.04	1.42	3.5	1.63	4.8	0.4
51	0.05	0.04	1.44	3.6	1.72	5.2	0.38
52	-0.35	0.07	0.85	-1	0.87	-0.9	0.59
53	-0.23	0.07	0.97	-0.2	0.91	-0.7	0.51
54	-0.17	0.06	0.98	-0.1	0.9	-0.7	0.54
55	0.06	0.05	1	0	0.97	-0.2	0.55
56	-0.04	0.06	1.04	0.4	1.05	0.5	0.52
57	0.31	0.11	0.95	-0.2	0.95	-0.1	0.49
58	0.3	0.1	0.95	-0.1	0.97	-0.1	0.51
59	0.39	0.1	1.08	0.4	1.11	0.5	0.42
60	0.55	0.05	0.9	-1	0.9	-0.9	0.65
61	0.27	0.05	1	0	1	0	0.57
62	0.03	0.06	1.04	0.4	1.01	0.1	0.48
63	0.23	0.06	1.04	0.4	1.03	0.3	0.5
64	0.4	0.05	1.2	1.6	1.22	1.7	0.42
65	0.15	0.05	0.95	-0.3	0.94	-0.5	0.59

The most difficult item within the parent scale was item 60 (“your child's access to community services and facilities?”) [0.55, SE=0.05], which was approximately two standard deviation above the item mean. The easiest item within the child scale was item 10 (“going out on trips with the family?”) [-0.62, SE=0.07], which was approximately two standard deviation below the item mean (see Appendix, Figure B1)

Figure 4.2

Characteristic Curve of C CP-QOL Child: Parent Scale



According to the test characteristic curve of the parent scale, the expected score on test ranged from 82 to 583, the expected mean score was 332.5, and the measure (in logit) ranged from -3.5 to 5 approximately (Figure 4.2).

After conducting the step calibration item by item in the parent scale, the results showed that, in item 1 (“the way they get along with people, generally?”), the Rasch half-point thresholds ranged from -2.03 to 1.81 (-2.03, -1.3, -0.75, -0.07, 0.78 and 1.81 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 1 (Figure D1), the expected scoring range was from 3 to 9 and the expected mean score was 6.5; in item 2 (“the way they get along with you?”), the Rasch half-point thresholds ranged from -1.95 to 1.08 (-1.95, -0.92, -0.03 and 1.08 in Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 2 (Figure D2), the expected scoring range was from 5 to 9 and the expected mean score was 7.5; in item 3 (“the way they get along with their brothers and sisters?”), the Rasch half-point thresholds ranged from -1.49 to 1.17 (-1.49, -0.58, 0.13 and 1.17 in Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 3 (Figure D3), the expected scoring range was from 5 to 9 and the expected mean score was 7.5; in item 4 (“the way they get along with other children at preschool or school?”), the Rasch half-point thresholds ranged from -2.29 to 1.99 (-2.29, -1.46, -0.73, 0.07, 0.93 and 1.99 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 4 (Figure D4), the expected scoring range was from 3 to 9 and the expected mean score was 6.5; in item 5 (“the way they get along with other children outside of preschool or school?”), the Rasch half-point thresholds ranged from -2.04 to 2.12 (-2.04, -1.38, -1, -0.67, -0.31, 0.24, 1.08 and 2.12 in

Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 5 (Figure D5), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 6 ("the way they get along with adults?"), the Rasch half-point thresholds ranged from -1.64 to 2.07 (-1.64, -1.33, -1.15, -0.96, -0.7, -0.07, 0.92 and 2.07 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 6 (Figure D6), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 7 ("the way they get along with their teachers and/or careers?"), the Rasch half-point thresholds ranged from -2.35 to 1.68 (-2.35, -1.42, -0.85, -0.27, 0.51 and 1.68 in Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 7 (Figure D7), the expected scoring range was from 3 to 9 and the expected mean score was 7; in item 8 ("their ability to play on their own?"), the Rasch half-point thresholds ranged from -1.7 to 1.91 (-1.7, -1.14, -0.85, -0.6, -0.32, 0.12, 0.84 and 1.91 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 8 (Figure D8), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 9 ("their ability to play with friends?"), the Rasch half-point thresholds ranged from -2.04 to 1.67 (-2.04, -1.06, -0.65, -0.3, 0.13, 0.73 and 1.67 in Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 9 (Figure D9), the expected scoring range was from 2 to 9 and the expected mean score was

6.5; in item 10 (“going out on trips with the family?”), the Rasch half-point thresholds ranged from -1.69 to 0.98 (-1.69, -1.21, -1.01, -0.85, -0.69, -0.46, -0.02 and 0.98 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 10 (Figure D10), the expected scoring range was from 1 to 9 and the expected mean score was 7.5; in item 11 (“how they are accepted by their family?”), the Rasch half-point thresholds ranged from -1.68 to 1.04 (-1.68, -1.3, -1.09, -0.9, -0.7, -0.41, 0.09 and 1.04 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 11 (Figure D11), the expected scoring range was from 1 to 9 and the expected mean score was 7.5; in item 12 (“how they are accepted by other children at preschool or school?”), the Rasch half-point thresholds ranged from -1.59 to 2.36 (-1.59, -1.22, -0.99, -0.76, -0.44, 0.15, 1.05 and 2.36 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 12 (Figure D12), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 13 (“how they are accepted by other children outside of preschool or school?”), the Rasch half-point thresholds ranged from -2.08 to 2.48 (-2.08, -1.42, -1, -0.61, -0.15, 0.49, 1.37 and 2.48 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 13 (Figure D13), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 14 (“how they are accepted by adults?”), the Rasch half-point thresholds ranged from -1.79 to 2.09 (-1.79,

-1.33, -1.08, -0.85, -0.57, -0.06, 0.91 and 2.09 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 14 (Figure D14), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 15 ("how they are accepted by people in general?"), the Rasch half-point thresholds ranged from -2.09 to 2.76 (-2.09, -1.37, -0.97, -0.6, -0.12, 0.59, 1.54 and 2.76 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 15 (Figure D15), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 16 ("being able to do the things they want to do?"), the Rasch half-point thresholds ranged from -1.12 to 1.51 (-1.12, -0.7, -0.49, -0.32, -0.12, 0.14, 0.6 and 1.51 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 16 (Figure D16), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 17 ("their ability to participate at preschool or school?"), the Rasch half-point thresholds ranged from -1.43 to 2.15 (-1.43, -0.98, -0.73, -0.5, -0.21, 0.28, 1.08 and 2.15 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 17 (Figure D17), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 18 ("their ability to participate in recreational activities?"), the Rasch half-point thresholds ranged from -1.34 to 1.81 (-1.34, -0.82, -0.56, -0.34, -0.09, 0.29, 0.9 and 1.81 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according

to the item characteristic curve of item 18 (Figure D18), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 19 ("their ability to participate in sporting activities?"), the Rasch half-point thresholds ranged from -1.21 to 2.03 (-1.21, -0.73, -0.46, -0.23, 0.06, 0.49, 1.12 and 2.03 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 19 (Figure D19), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 20 ("their ability to participate in social events outside of preschool or school?"), the Rasch half-point thresholds ranged from -1.32 to 1.91 (-1.32, -0.83, -0.57, -0.33, -0.06, 0.37, 0.99 and 1.91 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 20 (Figure D20), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 21 ("their ability to participate in their community?"), the Rasch half-point thresholds ranged from -1.14 to 2.02 (-1.14, -0.68, -0.44, -0.21, 0.07, 0.5, 1.1 and 2.02 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 21 (Figure D21), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 22 ("the way they communicate with people they know well?"), the Rasch half-point thresholds ranged from -1.31 to 1.81 (-1.31, -0.95, -0.76, -0.59, -0.4, -0.09, 0.61 and 1.81 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 22 (Figure D22), the expected scoring range was from 1 to 9 and

the expected mean score was 6.5; in item 23 (“the way they communicate with people they don't know well?”), the Rasch half-point thresholds ranged from -1.26 to 3.18 (-1.26, -0.74, -0.43, -0.13, 0.3, 1.1 and 3.18 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6 and Step 7 respectively), and according to the item characteristic curve of item 23 (Figure D23), the expected scoring range was from 1 to 8 and the expected mean score was 5; in item 24 (“the way other people communicate with them?”), the Rasch half-point thresholds ranged from -1.34 to 3.07 (-1.34, -0.88, -0.6, -0.32, 0.1, 0.93 and 3.07 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6 and Step 7 respectively), and according to the item characteristic curve of item 24 (Figure D24), the expected scoring range was from 1 to 8 and the expected mean score was 5.5; in item 25 (“their physical health?”), the Rasch half-point thresholds ranged from -2.22 to 2.35 (-2.22, -1.43, -0.9, -0.49, -0.11, 0.37, 1.15 and 2.35 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 25 (Figure D25), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 26 (“the way they get around?”), the Rasch half-point thresholds ranged from -1.51 to 1.82 (-1.51, -0.86, -0.55, -0.3, -0.05, 0.28, 0.83 and 1.82 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 26 (Figure D26), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 27 (“how they sleep?”), the Rasch half-point thresholds ranged from -2.13 to 1.55 (-2.13, -1.24, -0.88, -0.58, -0.19, 0.44 and 1.55 in Step 2,

Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 27 (Figure D27), the expected scoring range was from 2 to 9 and the expected mean score was 7; in item 28 ("the way they look?"), the Rasch half-point thresholds ranged from -1.76 to 1.81 (-1.76, -1.33, -1.09, -0.87, -0.6, -0.17, 0.63 and 1.81 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 28 (Figure D28), the expected scoring range was from 1 to 9 and the expected mean score was 7; in item 29 ("their ability to keep up academically with their peers?"), the Rasch half-point thresholds ranged from -1.39 to 2.71 (-1.39, -0.82, -0.49, -0.16, 0.24, 0.76, 1.44 and 2.71 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 29 (Figure D29), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 30 ("their ability to keep up physically with their peers?"), the Rasch half-point thresholds ranged from -1.4 to 2.72 (-1.4, -0.7, -0.3, 0.06, 0.49, 1.04, 1.71 and 2.72 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 30 (Figure D30), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 31 ("their life in general?"), the Rasch half-point thresholds ranged from -1.96 to 2.53 (-1.96, -1.36, -1.03, -0.73, -0.39, 0.15, 1.15 and 2.53 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 31 (Figure D31), the expected scoring range was from

1 to 9 and the expected mean score was 6.5; in item 32 (“themselves?”), the Rasch half-point thresholds ranged from -1.35 to 2.22 (-1.35, -0.99, -0.77, -0.57, -0.3, 0.2, 1.07 and 2.22 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 32 (Figure D32), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 33 (“their future?”), the Rasch half-point thresholds ranged from -1.36 to 2.8 (-1.36, -0.85, -0.53, -0.19, 0.3, 0.96, 1.69 and 2.8 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 33 (Figure D33), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 34 (“their opportunities in life?”), the Rasch half-point thresholds ranged from -1.34 to 2.84 (-1.34, -0.84, -0.54, -0.23, 0.22, 0.89, 1.73 and 2.84 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 34 (Figure D34), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 35 (“the way they use their arms?”), the Rasch half-point thresholds ranged from -1.44 to 2.08 (-1.44, -0.86, -0.54, -0.25, 0.08, 0.53, 1.14 and 2.08 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 35 (Figure D35), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 36 (“the way they use their legs?”), the Rasch half-point thresholds ranged from -1.49 to 2.44 (-1.49, -0.76, -0.32, 0.04, 0.4, 0.82, 1.42 and 2.44 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6,

Step 7 and Step 8 respectively), and according to the item characteristic curve of item 36 (Figure D36), the expected scoring range was from 1 to 9 and the expected mean score was 4.5; in item 37 ("the way they use their hands?"), the Rasch half-point thresholds ranged from -1.53 to 2.38 (-1.53, -0.92, -0.57, -0.27, 0.06, 0.52, 1.22 and 2.38 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 37 (Figure D37), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 38 ("their ability to dress themselves?"), the Rasch half-point thresholds ranged from -1.13 to 2.3 (-1.13, -0.51, -0.18, 0.11, 0.41, 0.81, 1.36 and 2.3 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 38 (Figure D38), the expected scoring range was from 1 to 9 and the expected mean score was 4.5; in item 39 ("their ability to eat or drink independently?"), the Rasch half-point thresholds ranged from -1.22 to 1.57 (-1.22, -0.73, -0.49, -0.28, -0.07, 0.23, 0.7 and 1.57 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 39 (Figure D39), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 40 ("their ability to use the toilet by themselves?"), the Rasch half-point thresholds ranged from -1.15 to 2.54 (-1.15, -0.57, -0.27, 0, 0.3, 0.71, 1.37 and 2.54 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 40 (Figure D40), the expected scoring range was from 1 to 9 and

the expected mean score was 4.5; in item 41 (“the special equipment they have at home?”), the Rasch half-point thresholds ranged from -2.09 to 2.05 (-2.09, -1.39, -0.91, -0.44, 0.11, 0.69, 1.26 and 2.05 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 41 (Figure D41), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 42 (“the special equipment they have at their school?”), the Rasch half-point thresholds ranged from -1.47 to 2.26 (-1.47, -0.99, -0.71, -0.44, -0.08, 0.51, 1.29 and 2.26 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 42 (Figure D42), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 43 (“the special equipment that is available in the community?”), the Rasch half-point thresholds ranged from -1.35 to 2.4 (-1.35, -0.78, -0.46, -0.17, 0.17, 0.65, 1.34 and 2.4 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 43 (Figure D43), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 44 (“is your child bothered by hospital visits?”), the Rasch half-point thresholds ranged from -1.23 to 1.3 (-1.23, -0.67, -0.37, -0.12, 0.12, 0.37, 0.69 and 1.3 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 44 (Figure D44), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 45 (“is your child bothered when they miss school for health reasons?”), the Rasch half-point thresholds ranged from -0.93

to 1.22 (-0.93, -0.47, -0.24, -0.06, 0.12, 0.33, 0.62 and 1.22 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 45 (Figure D45), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 46 ("is your child bothered by being handled by other people?"), the Rasch half-point thresholds ranged from -1.63 to 1.3 (-1.63, -0.91, -0.56, -0.26, 0.02, 0.31, 0.66 and 1.3 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 46 (Figure D46), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 47 ("does your child worry about who will take care of them in the future?"), the Rasch half-point thresholds ranged from -1.21 to 1.25 (-1.21, -0.33, 0.3, 1.25, , , and in Step 1, Step 2, Step 3, Step 4, , , and respectively), and according to the item characteristic curve of item 47 (Figure D47), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 48 ("is your child concerned about having cerebral palsy?"), the Rasch half-point thresholds ranged from -0.92 to 1.31 (-0.92, -0.43, -0.17, 0.05, 0.25, 0.46, 0.74 and 1.31 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 48 (Figure D48), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 49 ("how much pain does your child have?"), the Rasch half-point thresholds ranged from -1.45 to 1.36 (-1.45, -0.76, -0.39, -0.11, 0.15, 0.41, 0.73 and 1.36 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively),

and according to the item characteristic curve of item 49 (Figure D49), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 50 (“how does your child feel about the amount of pain they have?”), the Rasch half-point thresholds ranged from -1.2 to 1.37 (-1.2, -0.61, -0.28, -0.01, 0.23, 0.48, 0.77 and 1.37 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 50 (Figure D50), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 51 (“how much discomfort does your child experience?”), the Rasch half-point thresholds ranged from -1.24 to 1.39 (-1.24, -0.64, -0.33, -0.09, 0.15, 0.4, 0.72 and 1.39 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 51 (Figure D51), the expected scoring range was from 1 to 9 and the expected mean score was 3.5; in item 52 (“how happy is your child?”), the Rasch half-point thresholds ranged from -1.84 to 1.84 (-1.84, -1.35, -1.08, -0.81, -0.49, 0.02, 0.79 and 1.84 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 52 (Figure D52), the expected scoring range was from 1 to 9 and the expected mean score was 6.5; in item 53 (“your child's access to treatment?”), the Rasch half-point thresholds ranged from -1.98 to 2.11 (-1.98, -1.4, -1.04, -0.69, -0.27, 0.31, 1.08 and 2.11 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 53 (Figure D53), the expected scoring range was from 1 to 9 and the expected mean score

was 6; in item 54 (“your child's access to therapy?”), the Rasch half-point thresholds ranged from -1.74 to 1.96 (-1.74, -1.18, -0.87, -0.58, -0.25, 0.23, 0.95 and 1.96 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 54 (Figure D54), the expected scoring range was from 1 to 9 and the expected mean score was 6; in item 55 (“your child's access to specialised medical or surgical care?”), the Rasch half-point thresholds ranged from -1.66 to 2.18 (-1.66, -0.98, -0.59, -0.26, 0.08, 0.51, 1.12 and 2.18 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 55 (Figure D55), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 56 (“your ability to get advice from a paediatrician?”), the Rasch half-point thresholds ranged from -1.76 to 1.92 (-1.76, -1.04, -0.66, -0.33, 0.02, 0.46, 1.04 and 1.92 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 56 (Figure D56), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 57 (“your access to respite care?”), the Rasch half-point thresholds ranged from -1.85 to 2.41 (, -1.85, -0.52, -0.04, 0.32, 0.69, 1.19 and 2.41 in , Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 57 (Figure D57), the expected scoring range was from 2 to 9 and the expected mean score was 4.5; in item 58 (“the amount of respite care you receive?”), the Rasch half-point thresholds ranged from -1.63 to 2.26 (-1.63, -0.65, -0.21, 0.11, 0.42, 0.78,

1.25 and 2.26 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 58 (Figure D58), the expected scoring range was from 1 to 9 and the expected mean score was 4.5; in item 59 ("how easy it is to get respite?"), the Rasch half-point thresholds ranged from -1.25 to 2.23 (-1.25, -0.51, -0.13, 0.17, 0.46, 0.82, 1.34 and 2.23 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 59 (Figure D59), the expected scoring range was from 1 to 9 and the expected mean score was 4; in item 60 ("your child's access to community services and facilities?"), the Rasch half-point thresholds ranged from -0.93 to 2.58 (-0.93, -0.37, -0.07, 0.19, 0.49, 0.89, 1.48 and 2.58 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 60 (Figure D60), the expected scoring range was from 1 to 9 and the expected mean score was 4; in item 61 ("your child's access to extra help with learning at preschool or school?"), the Rasch half-point thresholds ranged from -1.39 to 2.36 (-1.39, -0.74, -0.39, -0.09, 0.25, 0.72, 1.38 and 2.36 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 61 (Figure D61), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 62 ("your physical health?"), the Rasch half-point thresholds ranged from -1.95 to 2.4 (-1.95, -1.22, -0.77, -0.38, 0.04, 0.63, 1.44 and 2.4 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 62

(Figure D62), the expected scoring range was from 1 to 9 and the expected mean score was 5.5; in item 63 (“your work situation?”), the Rasch half-point thresholds ranged from -1.55 to 2.46 (-1.55, -0.9, -0.54, -0.19, 0.24, 0.8, 1.49 and 2.46 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 63 (Figure D63), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 64 (“your family's financial situation?”), the Rasch half-point thresholds ranged from -1.37 to 2.82 (-1.37, -0.74, -0.4, -0.1, 0.28, 0.88, 1.83 and 2.82 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 64 (Figure D64), the expected scoring range was from 1 to 9 and the expected mean score was 5; in item 65 (“how happy are you?”), the Rasch half-point thresholds ranged from -1.4 to 2.04 (-1.4, -0.8, -0.47, -0.17, 0.17, 0.61, 1.18 and 2.04 in Step 1, Step 2, Step 3, Step 4, Step 5, Step 6, Step 7 and Step 8 respectively), and according to the item characteristic curve of item 65 (Figure D65), the expected scoring range was from 1 to 9 and the expected mean score was 5.5 (Table 4.2b, Page 139-140).

Table 4.2b

Rasch half-point thresholds of C CP-QOL Child: Parent Scale

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
1			-2.03	-1.3	-0.75	-0.07	0.78	1.81
2					-1.95	-0.92	-0.03	1.08
3					-1.49	-0.58	0.13	1.17
4			-2.29	-1.46	-0.73	0.07	0.93	1.99
5	-2.04	-1.38	-1	-0.67	-0.31	0.24	1.08	2.12
6	-1.64	-1.33	-1.15	-0.96	-0.7	-0.07	0.92	2.07
7			-2.35	-1.42	-0.85	-0.27	0.51	1.68
8	-1.7	-1.14	-0.85	-0.6	-0.32	0.12	0.84	1.91
9		-2.04	-1.06	-0.65	-0.3	0.13	0.73	1.67
10	-1.69	-1.21	-1.01	-0.85	-0.69	-0.46	-0.02	0.98
11	-1.68	-1.3	-1.09	-0.9	-0.7	-0.41	0.09	1.04
12	-1.59	-1.22	-0.99	-0.76	-0.44	0.15	1.05	2.36
13	-2.08	-1.42	-1	-0.61	-0.15	0.49	1.37	2.48
14	-1.79	-1.33	-1.08	-0.85	-0.57	-0.06	0.91	2.09
15	-2.09	-1.37	-0.97	-0.6	-0.12	0.59	1.54	2.76
16	-1.12	-0.7	-0.49	-0.32	-0.12	0.14	0.6	1.51
17	-1.43	-0.98	-0.73	-0.5	-0.21	0.28	1.08	2.15
18	-1.34	-0.82	-0.56	-0.34	-0.09	0.29	0.9	1.81
19	-1.21	-0.73	-0.46	-0.23	0.06	0.49	1.12	2.03
20	-1.32	-0.83	-0.57	-0.33	-0.06	0.37	0.99	1.91
21	-1.14	-0.68	-0.44	-0.21	0.07	0.5	1.1	2.02
22	-1.31	-0.95	-0.76	-0.59	-0.4	-0.09	0.61	1.81
23	-1.26	-0.74	-0.43	-0.13	0.3	1.1	3.18	
24	-1.34	-0.88	-0.6	-0.32	0.1	0.93	3.07	
25	-2.22	-1.43	-0.9	-0.49	-0.11	0.37	1.15	2.35
26	-1.51	-0.86	-0.55	-0.3	-0.05	0.28	0.83	1.82
27		-2.13	-1.24	-0.88	-0.58	-0.19	0.44	1.55
28	-1.76	-1.33	-1.09	-0.87	-0.6	-0.17	0.63	1.81
29	-1.39	-0.82	-0.49	-0.16	0.24	0.76	1.44	2.71
30	-1.4	-0.7	-0.3	0.06	0.49	1.04	1.71	2.72
31	-1.96	-1.36	-1.03	-0.73	-0.39	0.15	1.15	2.53
32	-1.35	-0.99	-0.77	-0.57	-0.3	0.2	1.07	2.22
33	-1.36	-0.85	-0.53	-0.19	0.3	0.96	1.69	2.8
34	-1.34	-0.84	-0.54	-0.23	0.22	0.89	1.73	2.84
35	-1.44	-0.86	-0.54	-0.25	0.08	0.53	1.14	2.08

Table 4.2b *(Continued)*

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
36	-1.49	-0.76	-0.32	0.04	0.4	0.82	1.42	2.44
37	-1.53	-0.92	-0.57	-0.27	0.06	0.52	1.22	2.38
38	-1.13	-0.51	-0.18	0.11	0.41	0.81	1.36	2.3
39	-1.22	-0.73	-0.49	-0.28	-0.07	0.23	0.7	1.57
40	-1.15	-0.57	-0.27	0	0.3	0.71	1.37	2.54
41	-2.09	-1.39	-0.91	-0.44	0.11	0.69	1.26	2.05
42	-1.47	-0.99	-0.71	-0.44	-0.08	0.51	1.29	2.26
43	-1.35	-0.78	-0.46	-0.17	0.17	0.65	1.34	2.4
44	-1.23	-0.67	-0.37	-0.12	0.12	0.37	0.69	1.3
45	-0.93	-0.47	-0.24	-0.06	0.12	0.33	0.62	1.22
46	-1.63	-0.91	-0.56	-0.26	0.02	0.31	0.66	1.3
47	-1.21	-0.33	0.3	1.25				
48	-0.92	-0.43	-0.17	0.05	0.25	0.46	0.74	1.31
49	-1.45	-0.76	-0.39	-0.11	0.15	0.41	0.73	1.36
50	-1.2	-0.61	-0.28	-0.01	0.23	0.48	0.77	1.37
51	-1.24	-0.64	-0.33	-0.09	0.15	0.4	0.72	1.39
52	-1.84	-1.35	-1.08	-0.81	-0.49	0.02	0.79	1.84
53	-1.98	-1.4	-1.04	-0.69	-0.27	0.31	1.08	2.11
54	-1.74	-1.18	-0.87	-0.58	-0.25	0.23	0.95	1.96
55	-1.66	-0.98	-0.59	-0.26	0.08	0.51	1.12	2.18
56	-1.76	-1.04	-0.66	-0.33	0.02	0.46	1.04	1.92
57		-1.85	-0.52	-0.04	0.32	0.69	1.19	2.41
58	-1.63	-0.65	-0.21	0.11	0.42	0.78	1.25	2.26
59	-1.25	-0.51	-0.13	0.17	0.46	0.82	1.34	2.23
60	-0.93	-0.37	-0.07	0.19	0.49	0.89	1.48	2.58
61	-1.39	-0.74	-0.39	-0.09	0.25	0.72	1.38	2.36
62	-1.95	-1.22	-0.77	-0.38	0.04	0.63	1.44	2.4
63	-1.55	-0.9	-0.54	-0.19	0.24	0.8	1.49	2.46
64	-1.37	-0.74	-0.4	-0.1	0.28	0.88	1.83	2.82
65	-1.4	-0.8	-0.47	-0.17	0.17	0.61	1.18	2.04

The item content, categorized by different domains, was shown as blow (each question begins with “How do you think your child feels about…”):

Table 4.2c

CP-QOL Child: Parent Scale grouped by domain

Social Wellbeing and Acceptance
1 the way they get along with people, generally?
3 the way they get along with their brothers and sisters?
4 the way they get along with other children at preschool or school?
5 the way they get along with other children outside of preschool or school?
6 the way they get along with adults?
7 the way they get along with their teachers and/or careers?
10 going out on trips with the family?
11 how they are accepted by their family?
12 how they are accepted by other children at preschool or school?
13 how they are accepted by other children outside of preschool or school?
14 how they are accepted by adults?
15 how they are accepted by people in general?
Functioning
8 their ability to play on their own?
22 the way they communicate with people they know well?
23 the way they communicate with people they don't know well?
24 the way other people communicate with them?
27 how they sleep?
29 their ability to keep up academically with their peers?
34 their opportunities in life?
35 the way they use their arms?
37 the way they use their hands?
38 their ability to dress themselves?
39 their ability to eat or drink independently?
40 their ability to use the toilet by themselves?
Participation and Physical Health
9 their ability to play with friends?
16 being able to do the things they want to do?
17 their ability to participate at preschool or school?
18 their ability to participate in recreational activities?
19 their ability to participate in sporting activities?
20 their ability to participate in social events outside of preschool or school?
21 their ability to participate in their community?
25 their physical health?
26 the way they get around?

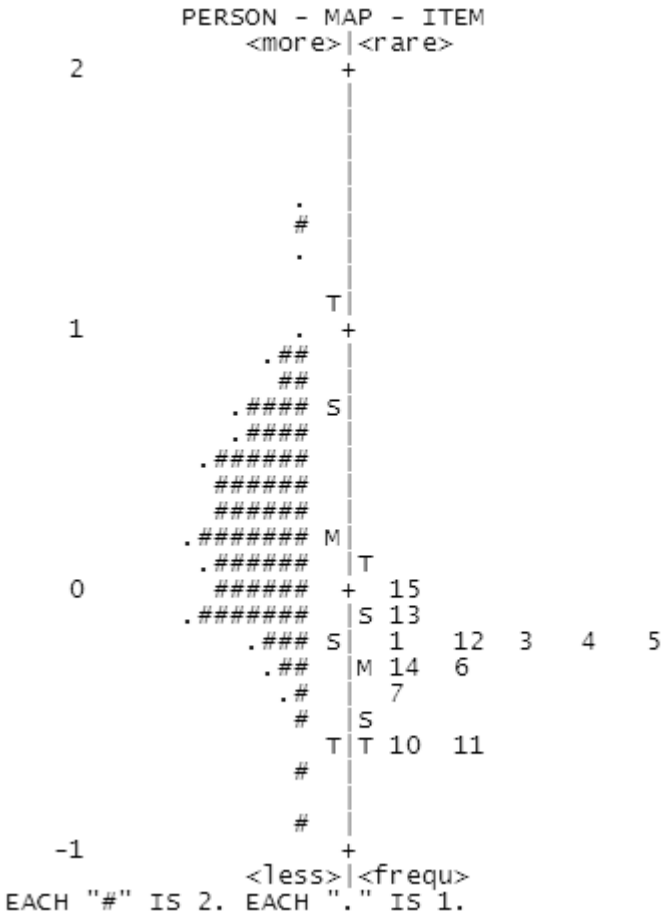
Table 4.2c *(Continued)*

Participation and Physical Health
30 their ability to keep up physically with their peers?
36 the way they use their legs?
Emotional Wellbeing
2 the way they get along with you?
28 the way they look?
31 their life in general?
32 themselves?
33 their future?
52 how happy is your child?
Access to Services
41 the special equipment they have at home?
42 the special equipment they have at their school?
43 the special equipment that is available in the community?
53 your child's access to treatment?
54 your child's access to therapy?
55 your child's access to specialised medical or surgical care?
56 your ability to get advice from a paediatrician?
57 your access to respite care?
58 the amount of respite care you receive?
59 how easy it is to get respite?
60 your child's access to community services and facilities?
61 your child's access to extra help with learning at preschool or school?
Pain and Impact of Disability
44 is your child bothered by hospital visits?
45 is your child bothered when they miss school for health reasons?
46 is your child bothered by being handled by other people?
47 does your child worry about who will take care of them in the future?
48 is your child concerned about having cerebral palsy?
49 how much pain does your child have?
50 how does your child feel about the amount of pain they have?
51 how much discomfort does your child experience?
Family Health
62 your physical health?
63 your work situation?
64 your family's financial situation?
65 how happy are you?

4.2.1. Subscales of parent proxy report

Figure 4.2.1a

Social Wellbeing and Acceptance Variable Map



Note: 1 the way they get along with people, generally?

3 the way they get along with their brothers and sisters?

4 the way they get along with other children at preschool or school?

5 the way they get along with other children outside of preschool or school?

6 the way they get along with adults?

7 the way they get along with their teachers and/or careers?

10 going out on trips with the family?

11 how they are accepted by their family?

12 how they are accepted by other children at preschool or school?

13 how they are accepted by other children outside of preschool or school?

14 how they are accepted by adults?

15 how they are accepted by people in general?

The person-item map (Figure 4.2.1a, Page 143) shows that the Social Wellbeing and Acceptance scale was relatively "easy" for this sample. The mean for the sample is two standard deviations above the mean for the items. The item difficulties cluster together.

The easiest items (10, "going out on trips with the family?"; 11, "how they are accepted by their family?") are approximately two standard deviations below the item mean. The most difficult item (15, "how they are accepted by people in general?") is approximately one and a half standard deviations above the item mean.

In the preceding table, item 14 ("how they are accepted by adults?") and 6 ("the way they get along with adults?") were just fallen into the mean range of the scale. The item mean, however, was far lower than the person mean, and the wide range of person distribution revealed that there are only limited effects on distinguishing this group of people within the scale.

Moreover, for the alignment of person-item distributions, around 65% person was above the item limit. These information have two meanings, the first one is that there are so many (65%) high "ability" people (or high satisfactory level in terms of social wellbeing and acceptance) within this specific group; the second one is that this subscale itself cannot adequately

measure those (65%) high “ability” people, possibly, design and insert some high “difficulty” items or amend the content of the redundant items is recommended.

Table 4.2.1a

Social Wellbeing and Acceptance Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
11	-0.6	0.07	1.36	2.2	1.32	2.1	0.23
3	-0.18	0.08	1.04	0.4	1.04	0.4	0.36
14	-0.32	0.07	1.04	0.3	0.98	-0.1	0.43
7	-0.45	0.07	0.98	-0.1	1.02	0.2	0.43
12	-0.16	0.08	0.99	0	0.99	0	0.43
10	-0.62	0.07	0.98	-0.1	0.94	-0.4	0.47
4	-0.23	0.08	0.96	-0.3	0.97	-0.2	0.43
1	-0.25	0.08	0.95	-0.3	0.97	-0.2	0.46
5	-0.23	0.07	0.91	-0.6	0.9	-0.7	0.53
13	-0.1	0.07	0.88	-0.9	0.86	-1.1	0.56
6	-0.33	0.08	0.88	-0.7	0.87	-0.9	0.51
15	-0.03	0.07	0.82	-1.4	0.78	-1.8	0.61

The item difficulty logits (Item Score in Table 4.2.1a) ranged from -0.62 to -0.03 on calibrations of the Social Wellbeing and Acceptance scale. The easiest item was item 10 (“going out on trips with the family?”) [-0.62, SE=0.07]. The most difficult item was item 15 (“how they are accepted by people in general?”) [-0.03, SE=0.07].

On the examination of the fit statistics of the Social Wellbeing and Acceptance scale, the infit and outfit statistics for items ranged from 0.78 to 1.36. The lowest MNSQ was the outfit for

item 15 (“how they are accepted by people in general?”) [0.78, SD=-1.8]. The highest MNSQ was the infit for item 11 (“how they are accepted by their family?”) [1.36, SD=2.2]

Within this scale, item 3 (“the way they get along with their brothers and sisters?”), 14 (“how they are accepted by adults?”), 7 (“the way they get along with their teachers and/or careers?”), 12 (“how they are accepted by other children at preschool or school?”), 10 (“going out on trips with the family?”), 4 (“the way they get along with other children at preschool or school?”), 1 (“the way they get along with people, generally?”), and 5 (“the way they get along with other children outside of preschool or school?”) were fallen into an optimal fit level (infit/outfit, 1±0.1).

Item 11 had a slightly low score to total correlation (r=0.23). In order to properly fit the criteria of unidimensionality, item 11 may recommended to be separated from other items in the scale and interpreted accordingly.

Figure 4.2.1b

Social Wellbeing and Acceptance Item Characteristic Curves

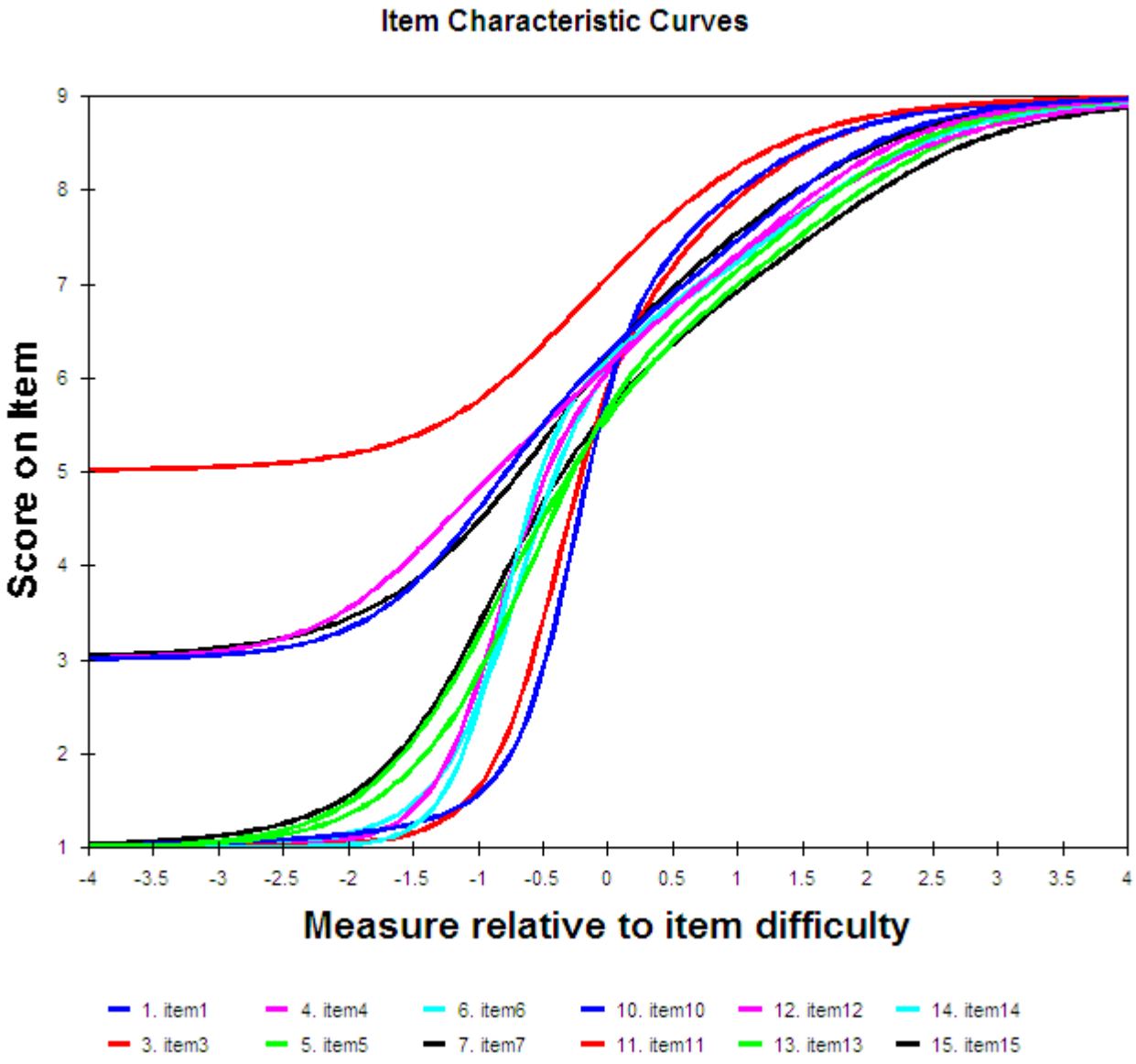


Table 4.2.1b

Social Wellbeing and Acceptance Scale: Rasch half-point thresholds

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
11	-1.68	-1.3	-1.09	-0.9	-0.7	-0.41	0.09	1.04
3					-1.49	-0.58	0.13	1.17
14	-1.79	-1.33	-1.08	-0.85	-0.57	-0.06	0.91	2.09
7			-2.35	-1.42	-0.85	-0.27	0.51	1.68
12	-1.59	-1.22	-0.99	-0.76	-0.44	0.15	1.05	2.36
10	-1.69	-1.21	-1.01	-0.85	-0.69	-0.46	-0.02	0.98
4			-2.29	-1.46	-0.73	0.07	0.93	1.99
1			-2.03	-1.3	-0.75	-0.07	0.78	1.81
5	-2.04	-1.38	-1	-0.67	-0.31	0.24	1.08	2.12
13	-2.08	-1.42	-1	-0.61	-0.15	0.49	1.37	2.48
6	-1.64	-1.33	-1.15	-0.96	-0.7	-0.07	0.92	2.07
15	-2.09	-1.37	-0.97	-0.6	-0.12	0.59	1.54	2.76

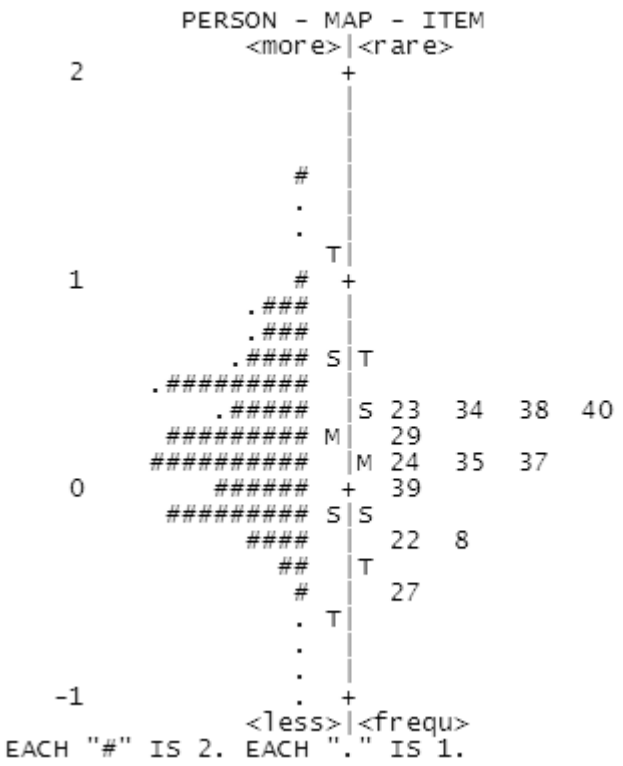
On Rasch half-point step calibrations of the Social Wellbeing and Acceptance scale (Figure 4.2.1b; Table 4.2.1b), the Rasch half-point thresholds for items ranged from -2.09 to 2.76. The lowest threshold was the step 1 for item 15 (“how they are accepted by people in general?”). The highest threshold was the step 8 for item 15 (“how they are accepted by people in general?”).

According to Table 4.2.1b, item 7, 4 and 1 had missed values in step 1 and step 2, it was due to the category 1 and 2 for these three items received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 7, 4 and 1 may recommended to be assigned a 7-point Likert option instead.

Item 3 had missed values in step 1, 2, 3 and 4; it was due to the category 1, 2, 3 and 4 for item 3 received zero respondent rates (no endorsement). In order to simplify the scale and shorten the completion time, item 3 may recommended to be assigned a 5-point Likert option instead.

Figure 4.2.2a

Functioning Variable Map



Note: 8 their ability to play on their own?

22 the way they communicate with people they know well?

23 the way they communicate with people they don't know well?

24 the way other people communicate with them?

27 how they sleep?

29 their ability to keep up academically with their peers?

34 their opportunities in life?

35 the way they use their arms?

37 the way they use their hands?

38 their ability to dress themselves?

39 their ability to eat or drink independently?

40 their ability to use the toilet by themselves?

The person-item map (Figure 4.2.2a, Page 150) shows that the Functioning scale was relatively “easy” for this sample. The mean for the sample is a half standard deviation above the mean for the items. The item difficulties cluster together.

The easiest item (27, “how they sleep?”) is approximately two standard deviations below the item mean. The most difficult items (23, “the way they communicate with people they don't know well?”; 34, “their opportunities in life?”; 38, “their ability to dress themselves?”; 40, “their ability to use the toilet by themselves?”) are approximately one standard deviation above the item mean.

In the preceding table, item 24 (“the way other people communicate with them?”), 35 (“the way they use their arms?”), and 37 (“the way they use their hands?”) were just fallen into the mean range of the scale. The item mean, however, was lower than the person mean, and the wide range of person distribution limited the item effects on distinguishing this group of people within the scale.

More specifically, for the alignment of person-item distributions, around 35% person was above the item limit. These information have two meanings, the first one is that there are so

many (35%) high “ability” people (or high satisfactory level in terms of functioning) within this group of samples; the second one is that this subscale itself cannot adequately measure those (35%) high “ability” people, possibly, design and insert some high “difficulty” items or amend the content of the redundant items is recommended.

Table 4.2.2a

Functioning Wellbeing Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
29	0.31	0.06	1.06	0.6	1.13	1.1	0.48
23	0.31	0.06	0.97	-0.2	1.05	0.4	0.51
8	-0.21	0.06	0.95	-0.3	1	0	0.53
27	-0.45	0.07	0.97	-0.1	0.98	-0.1	0.48
22	-0.19	0.06	0.93	-0.4	0.96	-0.2	0.52
40	0.38	0.05	0.92	-0.7	0.95	-0.4	0.6
38	0.4	0.05	0.93	-0.6	0.94	-0.5	0.61
34	0.36	0.06	0.9	-0.7	0.92	-0.5	0.55
35	0.1	0.05	0.88	-1	0.9	-0.8	0.6
39	-0.03	0.05	0.84	-1.3	0.85	-1.1	0.65
37	0.13	0.05	0.8	-1.7	0.8	-1.7	0.65
24	0.16	0.06	0.78	-1.6	0.8	-1.4	0.61

The item difficulty logits (Item Score in Table 4.2.2a) ranged from -0.45 to 0.4 on calibrations of the Functioning scale. The easiest item was item 27 (“how they sleep?”) [-0.45, SE=0.07]. The most difficult item was item 38 (“their ability to dress themselves?”) [0.4, SE=0.05].

On the examination of the fit statistics of the Functioning scale, the infit and outfit statistics for items ranged from 0.78 to 1.13. The lowest MNSQ was the infit for item 24 (“the way other people communicate with them?”) [0.78, SD=-1.6]. The highest MNSQ was the outfit for item 29 (“their ability to keep up academically with their peers?”) [1.13, SD=1.1]

Within this scale, item 23 (“the way they communicate with people they don't know well?”), 8 (“their ability to play on their own?”), 27 (“how they sleep?”), 22 (“the way they communicate with people they know well?”), 40 (“their ability to use the toilet by themselves?”), 38 (“their ability to dress themselves?”), 34 (“their opportunities in life?”), and 35 (“the way they use their arms?”) were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.2.2b

Functioning Item Characteristic Curves

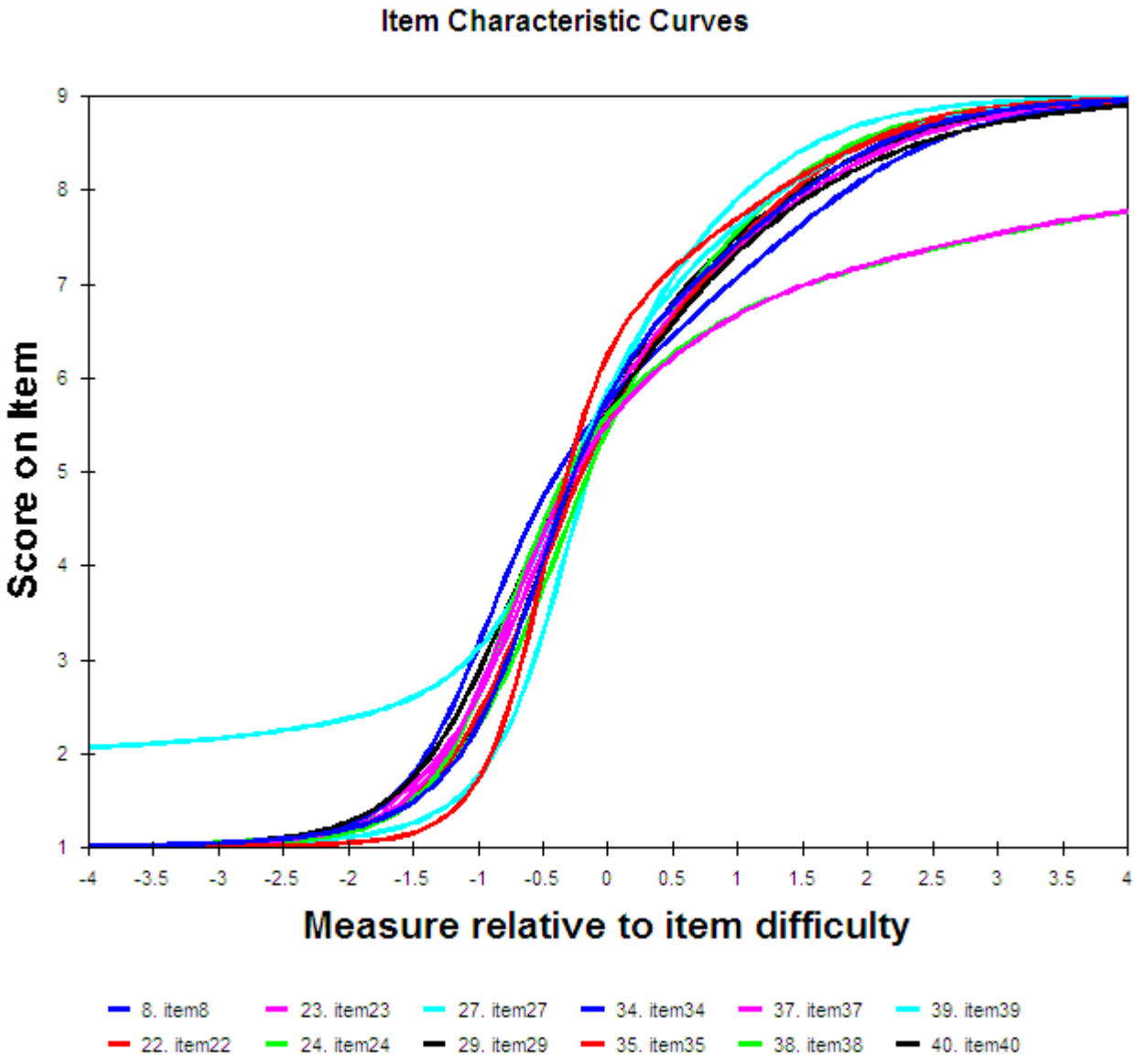


Table 4.2.2b

Functioning Wellbeing Scale: Rasch half-point thresholds

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
29	-1.39	-0.82	-0.49	-0.16	0.24	0.76	1.44	2.71
23	-1.26	-0.74	-0.43	-0.13	0.3	1.1	3.18	
8	-1.7	-1.14	-0.85	-0.6	-0.32	0.12	0.84	1.91
27		-2.13	-1.24	-0.88	-0.58	-0.19	0.44	1.55
22	-1.31	-0.95	-0.76	-0.59	-0.4	-0.09	0.61	1.81
40	-1.15	-0.57	-0.27	0	0.3	0.71	1.37	2.54
38	-1.13	-0.51	-0.18	0.11	0.41	0.81	1.36	2.3
34	-1.34	-0.84	-0.54	-0.23	0.22	0.89	1.73	2.84
35	-1.44	-0.86	-0.54	-0.25	0.08	0.53	1.14	2.08
39	-1.22	-0.73	-0.49	-0.28	-0.07	0.23	0.7	1.57
37	-1.53	-0.92	-0.57	-0.27	0.06	0.52	1.22	2.38
24	-1.34	-0.88	-0.6	-0.32	0.1	0.93	3.07	

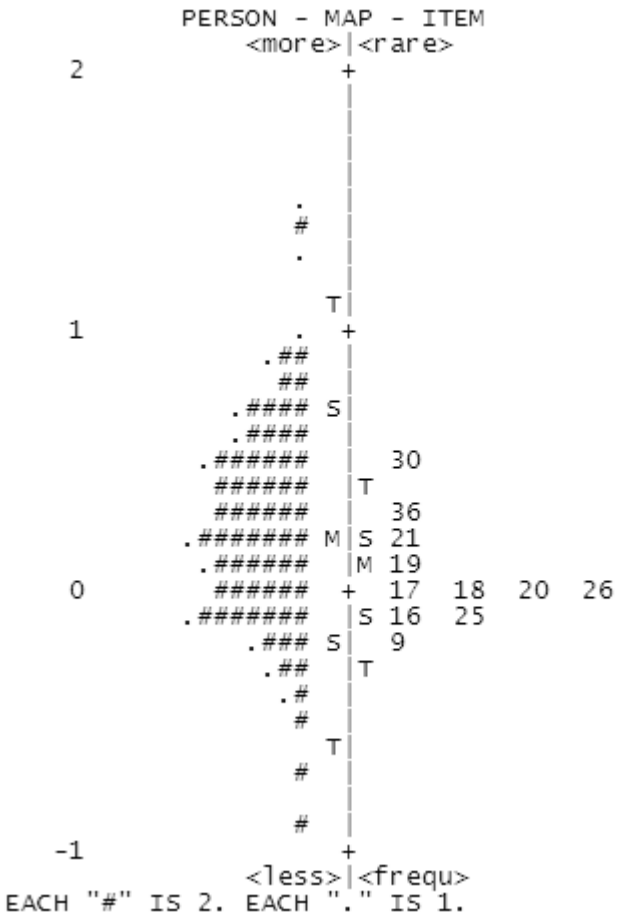
On Rasch half-point step calibrations of the Functioning scale (Figure 4.2.2b; Table 4.2.2b), the Rasch half-point thresholds for items ranged from -1.53 to 2.84. The lowest threshold was the step 1 for item 37 (“the way they use their hands?”). The highest threshold was the step 8 for item 34 (“their opportunities in life?”).

According to Table 4.2.2b, item 23 and item 24 had missed values in step 8, it was due to the category 9 for these two items received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 23 and item 24 may recommended to be assigned a 8-point Likert option instead.

Item 27 had missed values in step 1; it was due to the category 1 for item 27 received zero respondent rates (no endorsement). In order to simplify the scale and shorten the completion time, item 27 may recommended to be assigned an 8-point Likert option instead.

Figure 4.2.3a

Participation and Physical Health Variable Map



Note: 9 their ability to play with friends?

16 being able to do the things they want to do?

17 their ability to participate at preschool or school?

18 their ability to participate in recreational activities?

19 their ability to participate in sporting activities?

20 their ability to participate in social events outside of preschool or school?

21 their ability to participate in their community?

25 their physical health?

26 the way they get around?

30 their ability to keep up physically with their peers?

36 the way they use their legs?

The person-item map (Figure 4.2.3a, Page 157) shows that the Participation and Physical Health scale was relatively “easy” for this sample. The mean for the sample is one standard deviation above the mean for the items. The item difficulties cluster together.

The easiest item (9, “their ability to play with friends?”) is approximately one and a half standard deviations below the item mean. The most difficult item (30, “their ability to keep up physically with their peers?”) is approximately two standard deviations above the item mean.

In the preceding table, only item 19 (“their ability to play with friends?”) was just fallen into the mean range of the scale.

In details, for the alignment of person-item distributions, around 20% person was above the item upper limit, and around 15% person was below the item lower limit. These information have two meanings, the first one is that there are some (20%) high “ability” people (or high satisfactory level in terms of participation and physical health) and some (15%) low “ability” people (or low satisfactory level in terms of participation and physical health) within this specific group; the second one is that this subscale itself cannot adequately measure those high (20%) “ability” and low (15%) “ability” people, possibly, design and insert some items to the upper and lower ends or amend the content of the redundant items is recommended.

Table 4.2.3a

Participation and Physical Health Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
30	0.46	0.05	0.97	-0.3	0.98	-0.1	0.56
16	-0.05	0.05	0.94	-0.4	0.88	-0.8	0.6
19	0.14	0.05	0.91	-0.7	0.9	-0.8	0.57
21	0.16	0.05	0.86	-1.1	0.88	-0.9	0.61
9	-0.24	0.06	0.88	-1	0.88	-1	0.58
36	0.33	0.05	0.84	-1.6	0.85	-1.4	0.65
25	-0.15	0.06	0.83	-1.5	0.85	-1.3	0.61
20	0.02	0.05	0.78	-1.8	0.79	-1.6	0.66
26	-0.04	0.05	0.78	-1.9	0.74	-2.2	0.68
18	-0.02	0.05	0.74	-2.1	0.72	-2.2	0.69
17	-0.04	0.06	0.7	-2.1	0.71	-2.1	0.67

The item difficulty logits (Item Score in Table 4.2.3a) ranged from -0.24 to 0.46 on calibrations of the Participation and Physical Health scale. The easiest item was item 9 (“their ability to play with friends?”) [-0.24, SE=0.06]. The most difficult item was item 30 (“their ability to keep up physically with their peers?”) [0.46, SE=0.05].

On the examination of the fit statistics of the Participation and Physical Health scale, the infit and outfit statistics for items ranged from 0.7 to 0.98. The lowest MNSQ was the infit for item 17 (“their ability to participate at preschool or school?”) [0.7, SD=-2.1]. The highest MNSQ was the outfit for item 30 (“their ability to keep up physically with their peers?”) [0.98, SD=-0.1]

Within this scale, item 30 (“their ability to keep up physically with their peers?”), 16 (“being able to do the things they want to do?”), and 19 (“their ability to participate in sporting activities?”) were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.2.3b

Participation and Physical Health Item Characteristic Curves

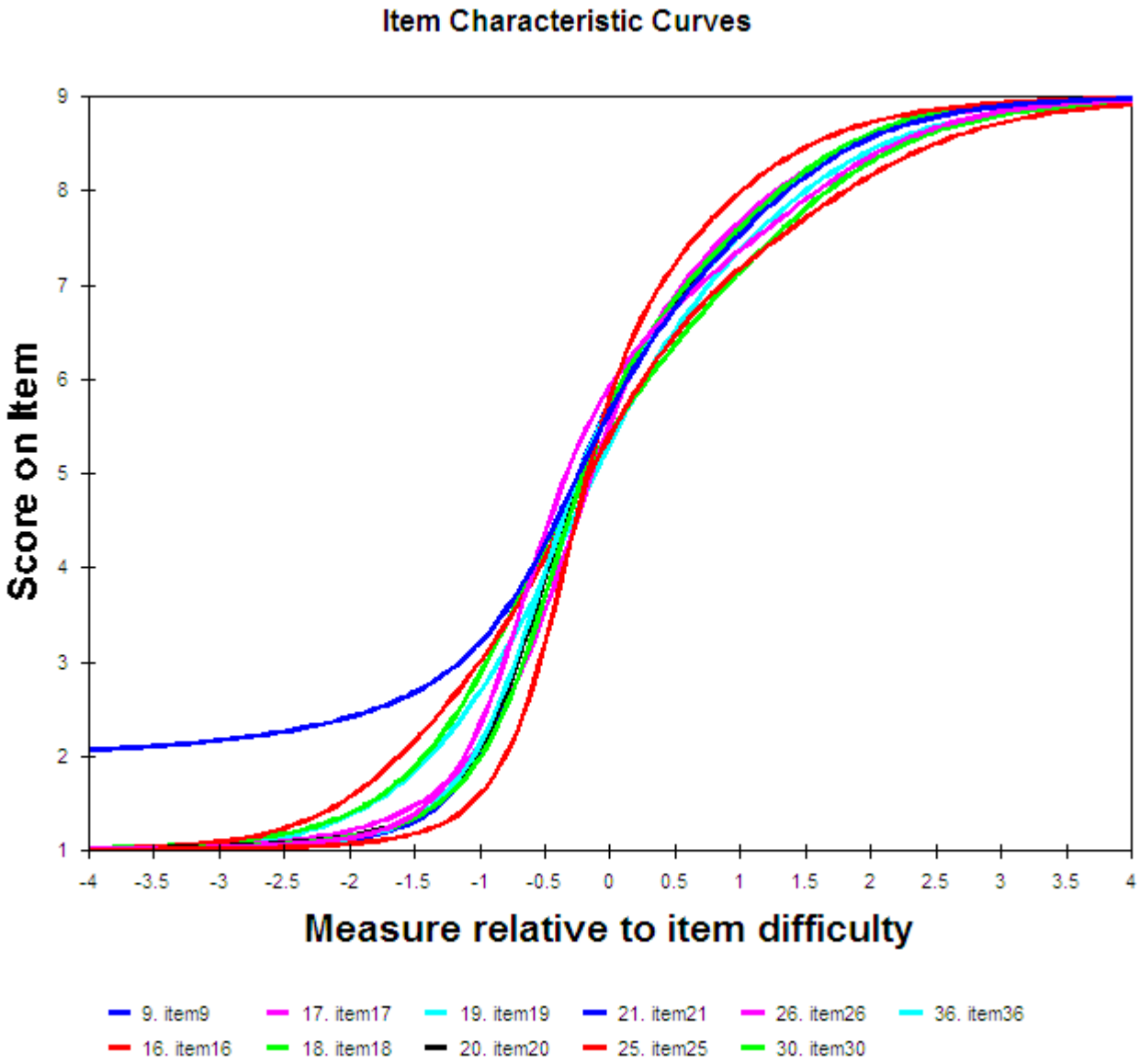


Table 4.2.3b

Participation and Physical Health Scale: Rasch half-point thresholds

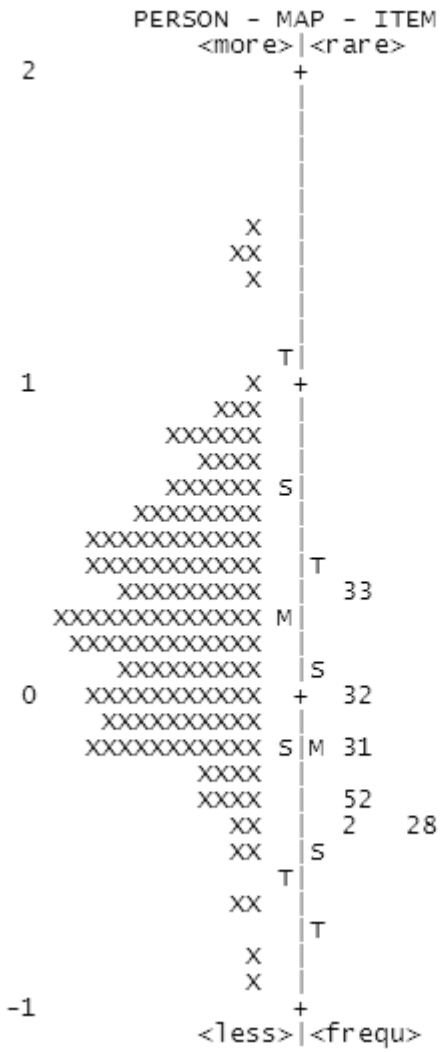
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
30	-1.4	-0.7	-0.3	0.06	0.49	1.04	1.71	2.72
16	-1.12	-0.7	-0.49	-0.32	-0.12	0.14	0.6	1.51
19	-1.21	-0.73	-0.46	-0.23	0.06	0.49	1.12	2.03
21	-1.14	-0.68	-0.44	-0.21	0.07	0.5	1.1	2.02
9		-2.04	-1.06	-0.65	-0.3	0.13	0.73	1.67
36	-1.49	-0.76	-0.32	0.04	0.4	0.82	1.42	2.44
25	-2.22	-1.43	-0.9	-0.49	-0.11	0.37	1.15	2.35
20	-1.32	-0.83	-0.57	-0.33	-0.06	0.37	0.99	1.91
26	-1.51	-0.86	-0.55	-0.3	-0.05	0.28	0.83	1.82
18	-1.34	-0.82	-0.56	-0.34	-0.09	0.29	0.9	1.81
17	-1.43	-0.98	-0.73	-0.5	-0.21	0.28	1.08	2.15

On Rasch half-point step calibrations of the Participation and Physical Health scale (Figure 4.2.3b; Table 4.2.3b), the Rasch half-point thresholds for items ranged from -2.22 to 2.72. The lowest threshold was the step 1 for item 25 ("their physical health?"). The highest threshold was the step 8 for item 30 ("their ability to keep up physically with their peers?").

According to Table 4.2.3b, item 9 had missed values in step 1, it was due to the category 1 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 9 may recommended to be assigned an 8-point Likert option instead.

Figure 4.2.4a

Emotional Wellbeing Variable Map



Note: 2 the way they get along with you?

28 the way they look?

31 their life in general?

32 themselves?

33 their future?

52 how happy is your child?

The person-item map (Figure 4.2.4a, Page 162) shows that the Emotional Wellbeing scale was relatively “easy” for this sample. The mean for the sample is one and a half standard deviations above the mean for the items. The item difficulties cluster together.

The easiest items (2, “the way they get along with you?”; 28, “the way they look?”) are approximately one standard deviation below the item mean. The most difficult item (33, “their future?”) is approximately two standard deviations above the item mean.

In the preceding table, only item 31 (“your ability to participate in your community?”) was just fallen into the mean range of the scale. The item mean, however, was lower than the person mean, and the wide range of person distribution limited the item effects on distinguishing this group of people within the scale.

Furthermore, for the alignment of person-item distributions, around 35% person was above the item upper limit, and around 10% person was below the item lower limit. These information have two meanings, the first one is that there are some (35%) high “ability” people (or high satisfactory level in terms of emotional wellbeing) and some (10%) low “ability” people (or low satisfactory level in terms of emotional wellbeing) within this group of samples; the second one is that this subscale itself cannot adequately measure those high

(35%) “ability” and low (10%) “ability” people, possibly, design and insert some items to the upper and lower ends or amend the content of the redundant items is recommended.

Table 4.2.4a

Emotional Wellbeing Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
2	-0.45	0.08	1.01	0.1	1.03	0.3	0.37
33	0.37	0.06	0.92	-0.6	0.92	-0.5	0.55
52	-0.35	0.07	0.85	-1	0.87	-0.9	0.59
32	-0.04	0.06	0.75	-1.7	0.78	-1.5	0.63
31	-0.19	0.07	0.74	-1.9	0.76	-1.7	0.65
28	-0.4	0.07	0.75	-1.7	0.75	-1.9	0.6

The item difficulty logits (Item Score in Table 4.2.4a) ranged from -0.45 to 0.37 on calibrations of the Emotional Wellbeing scale. The easiest item was item 2 (“the way they get along with you?”) [-0.45, SE=0.08]. The most difficult item was item 33 (“their future?”) [0.37, SE=0.06].

On the examination of the fit statistics of the Emotional Wellbeing scale, the infit and outfit statistics for items ranged from 0.75 to 1.03. The lowest MNSQ was the outfit for item 28 (“the way they look?”) [0.75, SD=-1.9]. The highest MNSQ was the outfit for item 2 (“the way they get along with you?”) [1.03, SD=0.3]

Within this scale, item 2 (“the way they get along with you?”) and 33 (“their future?”) were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.2.4b

Emotional Wellbeing Item Characteristic Curves

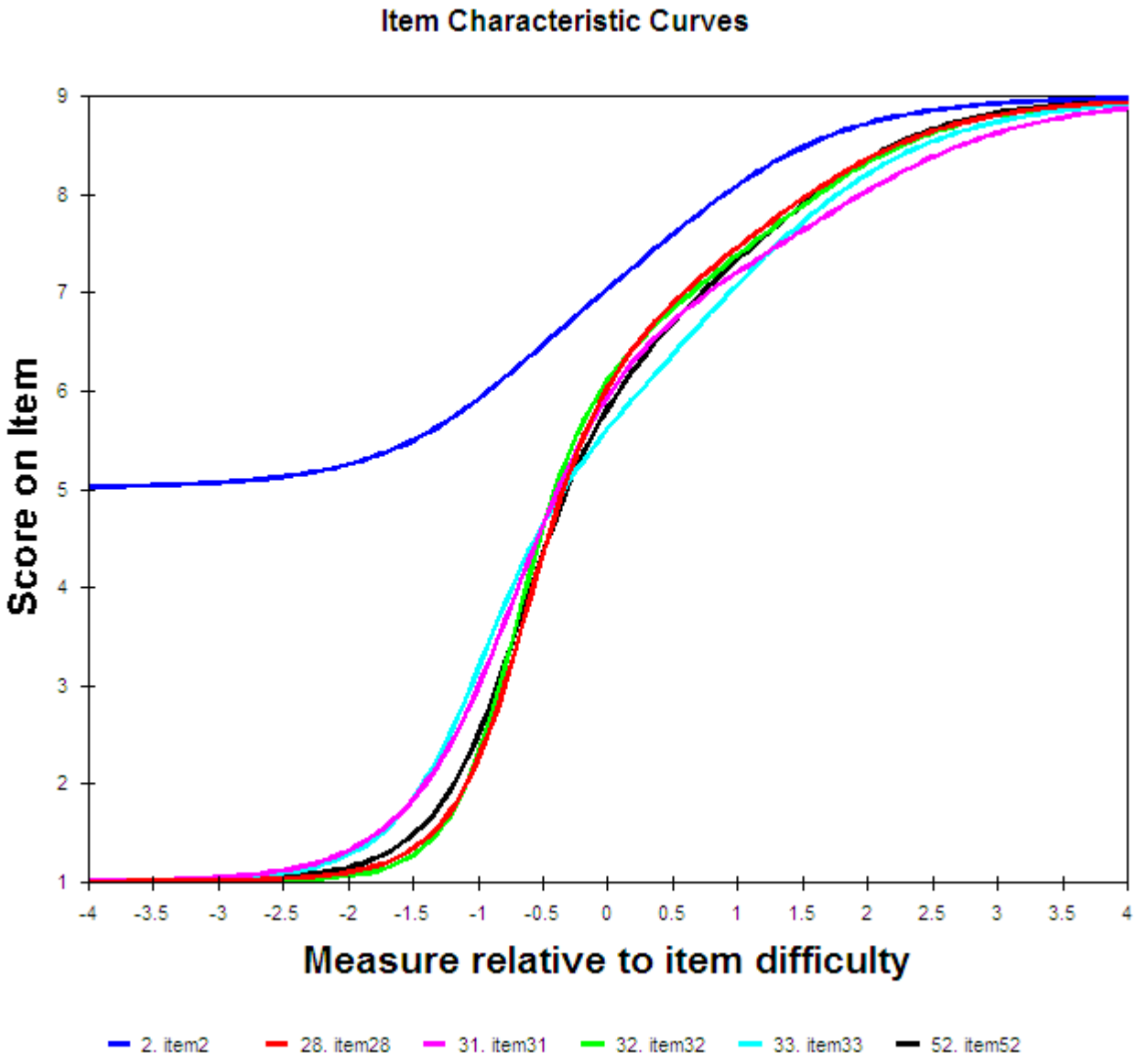


Table 4.2.4b

Emotional Wellbeing Scale: Rasch half-point thresholds

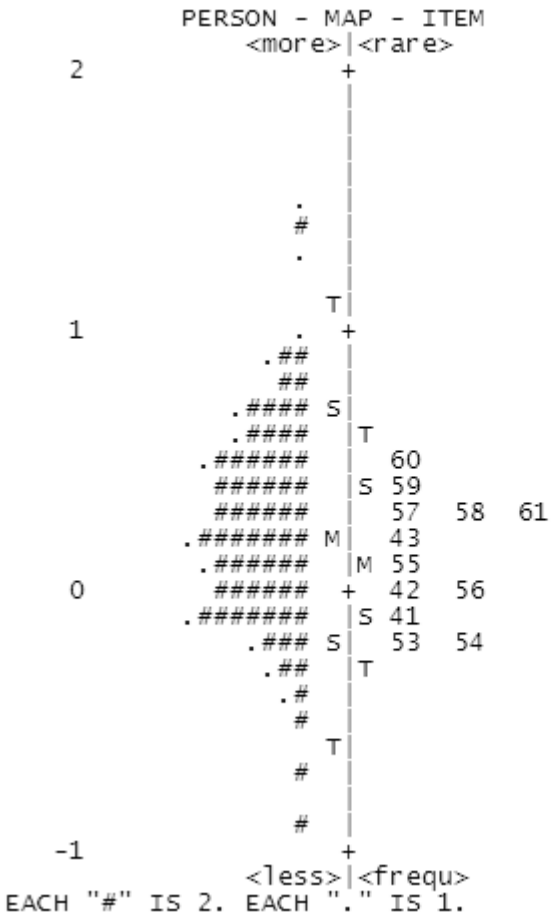
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
2					-1.95	-0.92	-0.03	1.08
33	-1.36	-0.85	-0.53	-0.19	0.3	0.96	1.69	2.8
52	-1.84	-1.35	-1.08	-0.81	-0.49	0.02	0.79	1.84
32	-1.35	-0.99	-0.77	-0.57	-0.3	0.2	1.07	2.22
31	-1.96	-1.36	-1.03	-0.73	-0.39	0.15	1.15	2.53
28	-1.76	-1.33	-1.09	-0.87	-0.6	-0.17	0.63	1.81

On Rasch half-point step calibrations of the Emotional Wellbeing scale (Figure 4.2.4b; Table 4.2.4b), the Rasch half-point thresholds for items ranged from -1.96 to 2.8. The lowest threshold was the step 1 for item 31 ("their life in general?"). The highest threshold was the step 8 for item 33 ("their future?").

According to Table 4.2.4b, item 2 had missed values in step 1, 2, 3 and 4, it was due to the category 1, 2, 3 and 4 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 2 may recommended to be assigned a 5-point Likert option instead.

Figure 4.2.5a

Access to Services Variable Map



Note: 41 the special equipment they have at home?

42 the special equipment they have at their school?

43 the special equipment that is available in the community?

53 your child's access to treatment?

54 your child's access to therapy?

55 your child's access to specialised medical or surgical care?

56 your ability to get advice from a paediatrician?

57 your access to respite care?

58 the amount of respite care you receive?

59 how easy it is to get respite?

60 your child's access to community services and facilities?

61 your child's access to extra help with learning at preschool or school?

The person-item map (Figure 4.2.5a, Page 167) shows that the Access to Services scale was relatively “easy” for this sample. The mean for the sample is one fourth standard deviation above the mean for the items. The item difficulties cluster together.

The easiest items (53, “your child's access to treatment?”; 54, “your child's access to therapy?”) are approximately one and a half standard deviations below the item mean. The most difficult item (60, “your child's access to community services and facilities?”) is approximately one and a half standard deviations above the item mean.

In the preceding table, only item 55 (“your child's access to specialised medical or surgical care?”) was just fallen into the mean range of the scale.

Moreover, for the alignment of person-item distributions, around 20% person was above the item upper limit, and around 15% person was below the item lower limit. These information have two meanings, the first one is that there are some (20%) high “ability” people (or high satisfactory level in terms of access to services) and some (15%) low “ability” people (or low satisfactory level in terms of access to services) within this group of samples; the second one is that this subscale itself cannot adequately measure those high (20%) “ability” and low

(15%) “ability” people, possibly, design and insert some items to the upper and lower ends or amend the content of the redundant items is recommended.

Table 4.2.5a

Access to Services Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
59	0.39	0.1	1.08	0.4	1.11	0.5	0.42
43	0.24	0.06	1.01	0.1	1.08	0.6	0.52
56	-0.04	0.06	1.04	0.4	1.05	0.5	0.52
55	0.06	0.05	1	0	0.97	-0.2	0.55
61	0.27	0.05	1	0	1	0	0.57
42	0.05	0.07	0.96	-0.2	0.99	0	0.51
54	-0.17	0.06	0.98	-0.1	0.9	-0.7	0.54
53	-0.23	0.07	0.97	-0.2	0.91	-0.7	0.51
58	0.3	0.1	0.95	-0.1	0.97	-0.1	0.51
57	0.31	0.11	0.95	-0.2	0.95	-0.1	0.49
41	-0.08	0.08	0.92	-0.5	0.92	-0.5	0.53
60	0.55	0.05	0.9	-1	0.9	-0.9	0.65

The item difficulty logits (Item Score in Table 4.2.5a) ranged from -0.23 to 0.55 on calibrations of the Access to Services scale. The easiest item was item 53 (“your child's access to treatment?”) [-0.23, SE=0.07]. The most difficult item was item 60 (“your child's access to community services and facilities?”) [0.55, SE=0.05].

On the examination of the fit statistics of the Access to Services scale, the infit and outfit statistics for items ranged from 0.9 to 1.11. The lowest MNSQ was the infit for item 60

(“your child's access to community services and facilities?”) [0.9, SD=-1]. The highest MNSQ was the outfit for item 59 (“how easy it is to get respite?”) [1.11, SD=0.5]

Within this scale, except item 59 (“how easy it is to get respite?”), all the items were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.2.5b

Access to Services Health Item Characteristic Curves

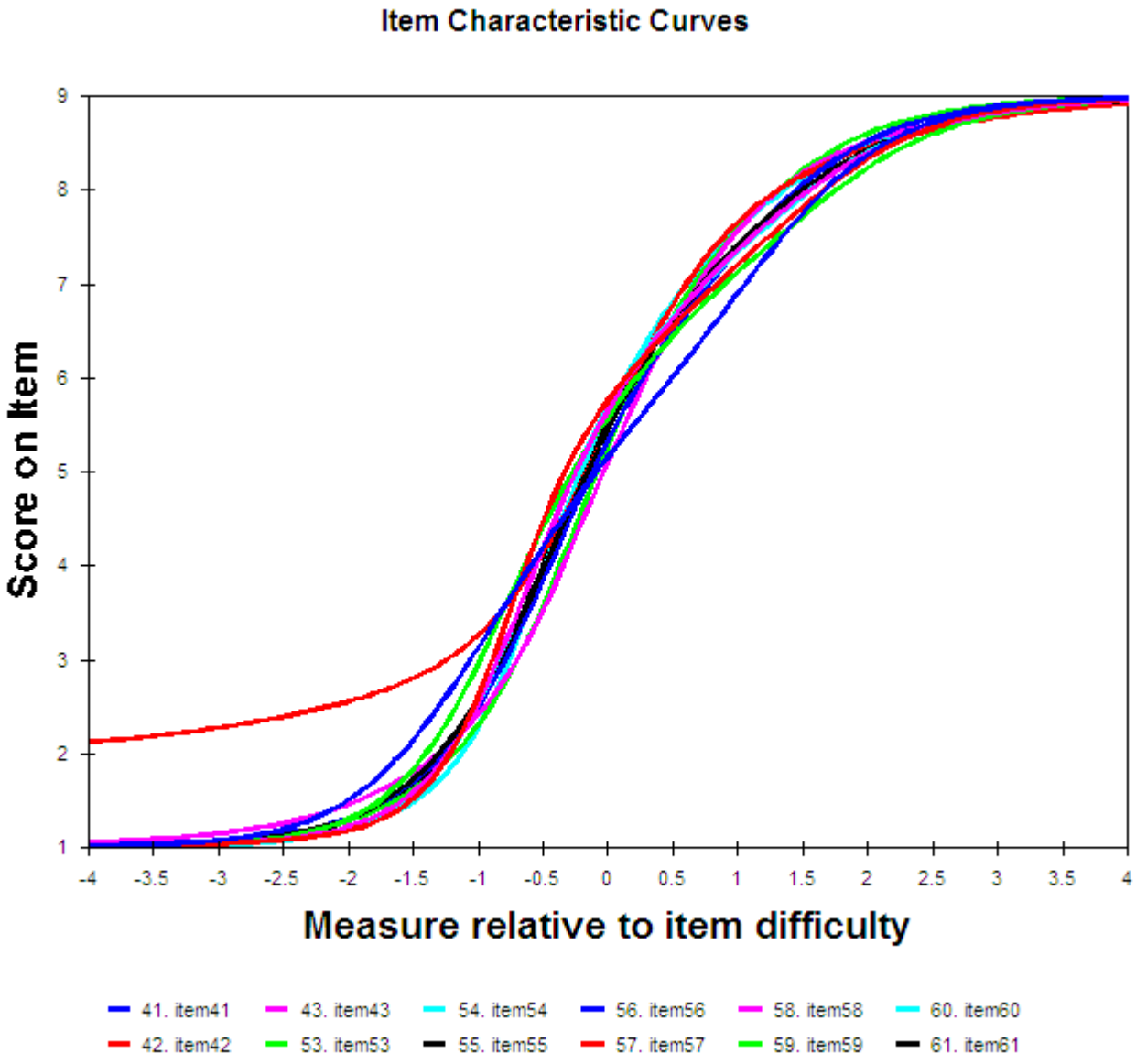


Table 4.2.5b

Access to Services Scale: Rasch half-point thresholds

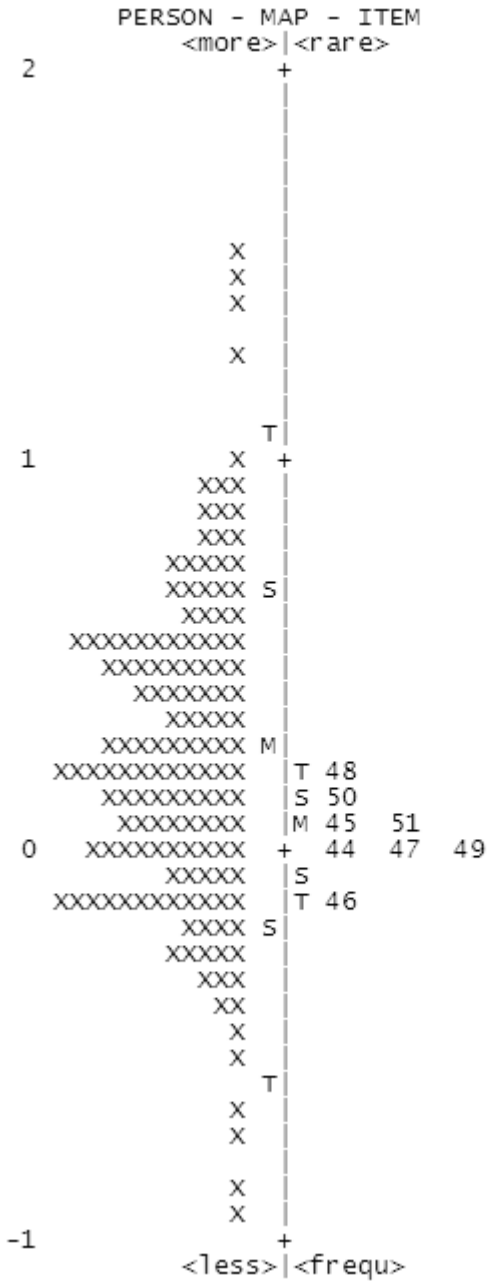
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
59	-1.25	-0.51	-0.13	0.17	0.46	0.82	1.34	2.23
43	-1.35	-0.78	-0.46	-0.17	0.17	0.65	1.34	2.4
56	-1.76	-1.04	-0.66	-0.33	0.02	0.46	1.04	1.92
55	-1.66	-0.98	-0.59	-0.26	0.08	0.51	1.12	2.18
61	-1.39	-0.74	-0.39	-0.09	0.25	0.72	1.38	2.36
42	-1.47	-0.99	-0.71	-0.44	-0.08	0.51	1.29	2.26
54	-1.74	-1.18	-0.87	-0.58	-0.25	0.23	0.95	1.96
53	-1.98	-1.4	-1.04	-0.69	-0.27	0.31	1.08	2.11
58	-1.63	-0.65	-0.21	0.11	0.42	0.78	1.25	2.26
57		-1.85	-0.52	-0.04	0.32	0.69	1.19	2.41
41	-2.09	-1.39	-0.91	-0.44	0.11	0.69	1.26	2.05
60	-0.93	-0.37	-0.07	0.19	0.49	0.89	1.48	2.58

On Rasch half-point step calibrations of the Access to Services scale (Figure 4.2.5b; Table 4.2.5b), the Rasch half-point thresholds for items ranged from -2.09 to 2.58. The lowest threshold was the step 1 for item 41 (“the special equipment they have at home?”). The highest threshold was the step 8 for item 60 (“your child's access to community services and facilities?”).

According to Table 4.2.5b, item 57 had missed values in step 1, it was due to the category 1 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 57 may recommended to be assigned an 8-point Likert option instead.

Figure 4.2.6a

Pain and Impact of Disability Variable Map



Note: 44 is your child bothered by hospital visits?

45 is your child bothered when they miss school for health reasons?

46 is your child bothered by being handled by other people?

47 does your child worry about who will take care of them in the future?

48 is your child concerned about having cerebral palsy?

49 how much pain does your child have?

50 how does your child feel about the amount of pain they have?

51 how much discomfort does your child experience?

The person-item map (Figure 4.2.6a, Page 173) shows that the Pain and Impact of Disability scale was relatively “easy” for this sample. The mean for the sample is two standard deviations above the mean for the items. The item difficulties cluster together.

The easiest item (46, “is your child bothered by being handled by other people?”) is approximately two standard deviations below the item mean. The most difficult item (48, “is your child concerned about having cerebral palsy?”) is approximately two standard deviations above the item mean.

In the preceding table, item 45 (“is your child bothered when they miss school for health reasons?”) and 51 (“how much discomfort does your child experience?”) were just fallen into the mean range of the scale. The item mean, however, was lower than the person mean, and the wide range of person distribution revealed that there are only limited effects on distinguishing this group of people within the scale.

More specifically, for the alignment of person-item distributions, around 50% person was above the item upper limit, and around 15% person was below the item lower limit. These information have two meanings, the first one is that there are so many (50%) high “ability” people (or high satisfactory level in terms of pain and impact of disability) and some (15%)

low “ability” people (or low satisfactory level in terms of pain and impact of disability) within this group of samples; the second one is that this subscale itself cannot adequately measure those high (50%) “ability” and low (15%) “ability” people, possibly, design and insert some items to the upper and lower ends or amend the content of the redundant items is recommended.

Table 4.2.6a

Pain and Impact of Disability Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
45	0.08	0.04	1.73	5.5	2.1	6.9	0.32
44	0.02	0.04	1.5	4	1.95	6.5	0.34
48	0.17	0.04	1.8	6.1	1.94	6.4	0.29
49	-0.01	0.05	1.47	3.9	1.8	5.8	0.34
51	0.05	0.04	1.44	3.6	1.72	5.2	0.38
50	0.1	0.04	1.42	3.5	1.63	4.8	0.4
47	0.01	0.07	1.27	2.6	1.43	3.7	0.2
46	-0.14	0.05	1.34	2.8	1.41	3.2	0.4

The item difficulty logits (Item Score in Table 4.2.6a) ranged from -0.14 to 0.17 on calibrations of the Pain and Impact of Disability scale. The easiest item was item 46 (“is your child bothered by being handled by other people?”) [-0.14, SE=0.05]. The most difficult item was item 48 (“is your child concerned about having cerebral palsy?”) [0.17, SE=0.04].

On the examination of the fit statistics of the Pain and Impact of Disability scale, the infit and outfit statistics for items ranged from 1.34 to 2.1. The lowest MNSQ was the infit for item 46

(“is your child bothered by being handled by other people?”) [1.34, SD=2.8]. The highest MNSQ was the outfit for item 45 (“is your child bothered when they miss school for health reasons?”) [2.1, SD=6.9]

Item 45 (“is your child bothered when they miss school for health reasons?”) [infit, 1.73, outfit, 2.1], 44 (“is your child bothered by hospital visits?”) [outfit, 1.95], 48 (“is your child concerned about having cerebral palsy?”) [infit, 1.8, outfit, 1.94], 49 (“how much pain does your child have?”) [outfit, 1.8], 51 (“how much discomfort does your child experience?”) [outfit, 1.72], 50 (“how does your child feel about the amount of pain they have?”) [outfit, 1.63] were out of the Likert scale critical range (i.e., 0.5-1.5). Item 48 and 47, had a slightly low score to total correlation (r=0.29, 0.2). In order to properly fit the criteria of unidimensionality, item 47 may recommended to be separated from other items in the scale and interpreted accordingly.

Figure 4.2.6b

Pain and Impact of Disability Item Characteristic Curves

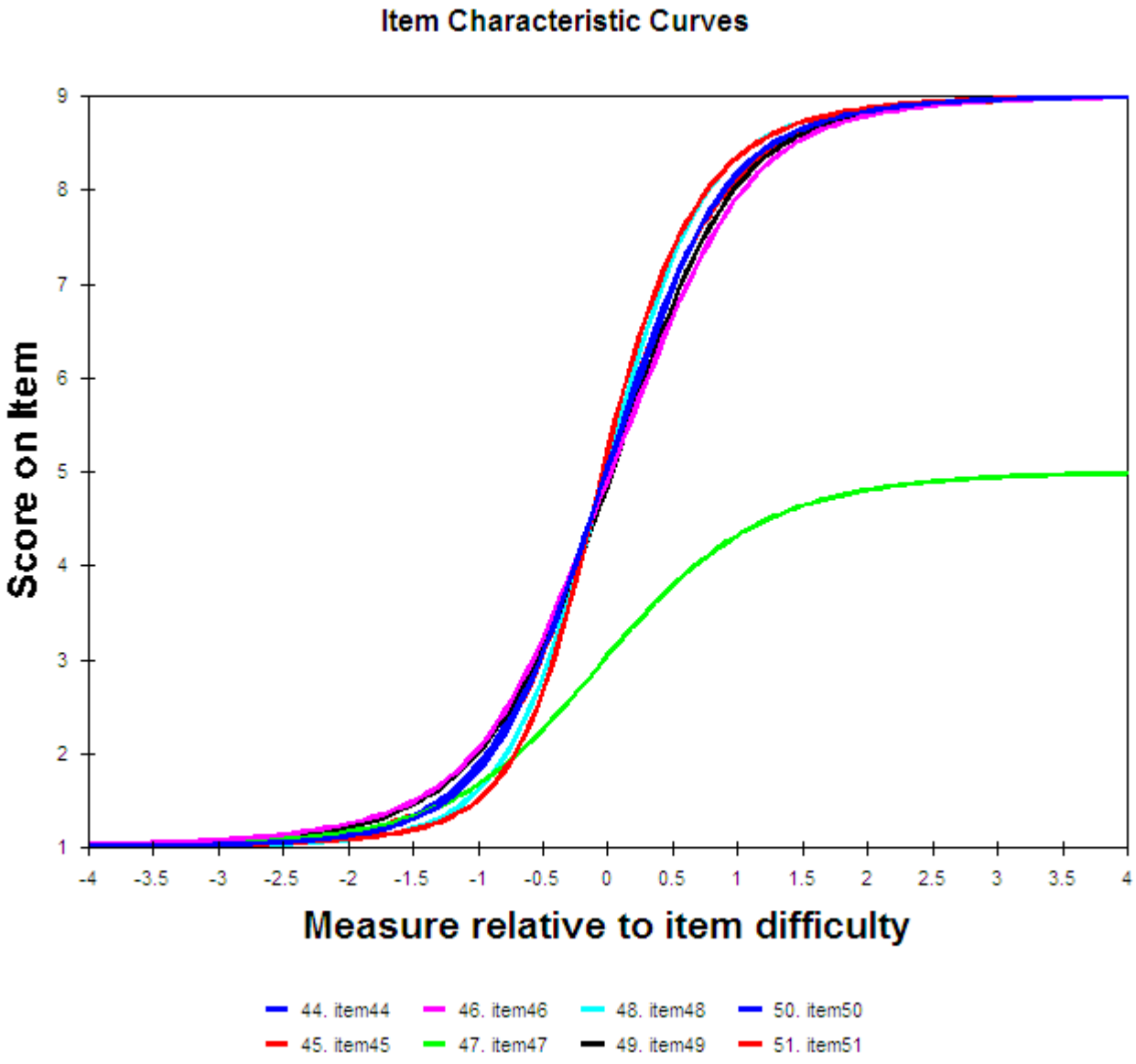


Table 4.2.6b

Pain and Impact of Disability Scale: Rasch half-point thresholds

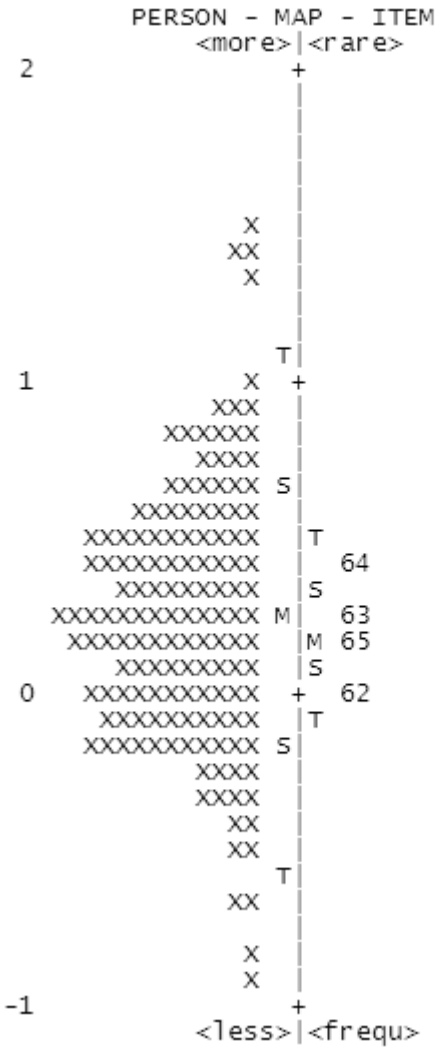
Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
45	-0.93	-0.47	-0.24	-0.06	0.12	0.33	0.62	1.22
44	-1.23	-0.67	-0.37	-0.12	0.12	0.37	0.69	1.3
48	-0.92	-0.43	-0.17	0.05	0.25	0.46	0.74	1.31
49	-1.45	-0.76	-0.39	-0.11	0.15	0.41	0.73	1.36
51	-1.24	-0.64	-0.33	-0.09	0.15	0.4	0.72	1.39
50	-1.2	-0.61	-0.28	-0.01	0.23	0.48	0.77	1.37
47	-1.21	-0.33	0.3	1.25				
46	-1.63	-0.91	-0.56	-0.26	0.02	0.31	0.66	1.3

On Rasch half-point step calibrations of the Pain and Impact of Disability scale (Figure 4.2.6b; Table 4.2.6b), the Rasch half-point thresholds for items ranged from -1.63 to 1.39. The lowest threshold was the step 1 for item 46 ("is your child bothered by being handled by other people?"). The highest threshold was the step 8 for item 51 ("how much discomfort does your child experience?").

According to Table 4.2.6b, item 47 had missed values in step 5, 6, 7 and 8, it was due to the category 6, 7, 8 and 9 for this item received zero respondent rate (no endorsement). In order to simplify the scale and shorten the completion time, item 47 may recommended to be assigned a 5-point Likert option instead.

Figure 4.2.7a

Family Health Variable Map



Note: 62 your physical health?

63 your work situation?

64 your family's financial situation?

65 how happy are you?

The person-item map (Figure 4.2.7a, Page 179) shows that the Family Health scale was relatively “easy” for this sample. The mean for the sample is a half standard deviation above the mean for the items. The item difficulties cluster together.

The easiest item (62, “your physical health?”) is approximately one and a half standard deviations below the item mean. The most difficult item (64, “your family's financial situation?”) is approximately one and a half standard deviations above the item mean.

In the preceding table, only item 65 (“how happy are you?”) was just fallen into the mean range of the scale. The item mean, however, was lower than the person mean, and the wide range of person distribution revealed that there are only limited effects on distinguishing this group of people within the scale.

In details, for the alignment of person-item distributions, around 30% person was above the item upper limit, and around 20% person was below the item lower limit. These information have two meanings, the first one is that there are some (30%) high “ability” people (or high satisfactory level in terms of family health) and some (20%) low “ability” people (or low satisfactory level in terms of family health) within this specific group; the second one is that this subscale itself cannot adequately measure those high (30%) “ability” and low (20%)

“ability” people, possibly, design and insert some items to the upper and lower ends or amend the content of the redundant items is recommended.

Table 4.2.7a

Family Health Scale Scores, Fit Statistics and Item Correlations

Item	Score	(SE)	Infit MNSQ	(SD)	Outfit MNSQ	(SD)	Correlation
64	0.4	0.05	1.2	1.6	1.22	1.7	0.42
63	0.23	0.06	1.04	0.4	1.03	0.3	0.5
62	0.03	0.06	1.04	0.4	1.01	0.1	0.48
65	0.15	0.05	0.95	-0.3	0.94	-0.5	0.59

The item difficulty logits (Item Score in Table 4.2.7a) ranged from 0.03 to 0.4 on calibrations of the Family Health scale. The easiest item was item 62 (“your physical health?”) [0.03, SE=0.06]. The most difficult item was item 64 (“your family's financial situation?”) [0.4, SE=0.05].

On the examination of the fit statistics of the Family Health scale, the infit and outfit statistics for items ranged from 0.94 to 1.22. The lowest MNSQ was the outfit for item 65 (“how happy are you?”) [0.94, SD=-0.5]. The highest MNSQ was the outfit for item 64 (“your family's financial situation?”) [1.22, SD=1.7]

Within this scale, except item 64 (“your family's financial situation?”), all the items were fallen into an optimal fit level (infit/outfit, 1±0.1).

Figure 4.2.7b
Family Health Item Characteristic Curves

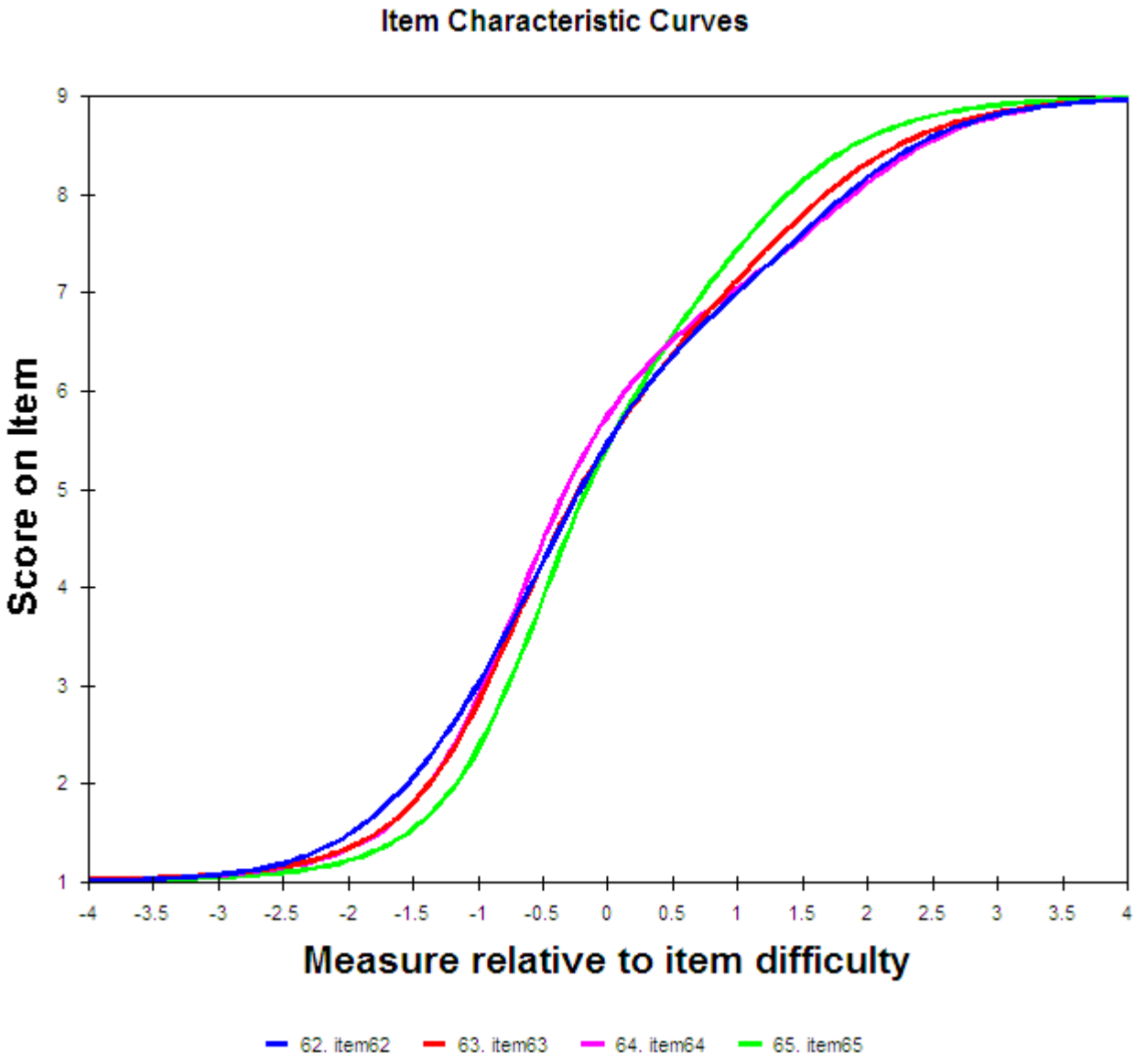


Table 4.2.7b

Family Health Scale: Rasch half-point thresholds

Item	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
64	-1.37	-0.74	-0.4	-0.1	0.28	0.88	1.83	2.82
63	-1.55	-0.9	-0.54	-0.19	0.24	0.8	1.49	2.46
62	-1.95	-1.22	-0.77	-0.38	0.04	0.63	1.44	2.4
65	-1.4	-0.8	-0.47	-0.17	0.17	0.61	1.18	2.04

On Rasch half-point step calibrations of the Family Health scale (Figure 4.2.7b; Table 4.2.7b), the Rasch half-point thresholds for items ranged from -1.95 to 2.82. The lowest threshold was the step 1 for item 62 (“your physical health?”). The highest threshold was the step 8 for item 64 (“your family's financial situation?”).

4.3. Comparison between child and parent reports

Table 4.3.1

Rasch half-point thresholds, logits, and correlations between child and parent versions of Social Wellbeing and Acceptance Scales

	Logit	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Child report									
Item									
1	-0.08			-1.65	-0.95	-0.49	0.03	0.79	1.76
3	-0.14	-1.08	-0.75	-0.56	-0.4	-0.23	0.02	0.45	1.31
4	-0.14	-1.5	-0.91	-0.63	-0.4	-0.15	0.2	0.75	1.62
5	0	-1.31	-0.86	-0.61	-0.37	-0.08	0.36	0.97	1.86
6	0			-1.59	-0.88	-0.37	0.19	0.85	1.74
7	-0.34	-1.39	-1.07	-0.88	-0.7	-0.48	-0.12	0.45	1.32
10	-0.59				-2.88	-0.91	-0.27	0.23	0.99
11	-0.42	-1.39	-1.04	-0.84	-0.68	-0.49	-0.24	0.2	1.02
12	-0.18	-1.53	-0.88	-0.6	-0.38	-0.16	0.14	0.61	1.46
13	-0.06	-1.62	-0.98	-0.63	-0.32	0	0.39	0.91	1.77
14	-0.17	-1.54	-1.1	-0.85	-0.63	-0.36	0.09	0.89	2.04
15	-0.04	-1.64	-0.97	-0.61	-0.3	0.06	0.46	0.94	1.75
Parent report									
Item									
1	-0.25			-2.03	-1.3	-0.75	-0.07	0.78	1.81
3	-0.18					-1.49	-0.58	0.13	1.17
4	-0.23			-2.29	-1.46	-0.73	0.07	0.93	1.99
5	-0.23	-2.04	-1.38	-1	-0.67	-0.31	0.24	1.08	2.12
6	-0.33	-1.64	-1.33	-1.15	-0.96	-0.7	-0.07	0.92	2.07
7	-0.45			-2.35	-1.42	-0.85	-0.27	0.51	1.68
10	-0.62	-1.69	-1.21	-1.01	-0.85	-0.69	-0.46	-0.02	0.98
11	-0.6	-1.68	-1.3	-1.09	-0.9	-0.7	-0.41	0.09	1.04
12	-0.16	-1.59	-1.22	-0.99	-0.76	-0.44	0.15	1.05	2.36
13	-0.1	-2.08	-1.42	-1	-0.61	-0.15	0.49	1.37	2.48
14	-0.32	-1.79	-1.33	-1.08	-0.85	-0.57	-0.06	0.91	2.09
15	-0.03	-2.09	-1.37	-0.97	-0.6	-0.12	0.59	1.54	2.76
Number of paired item	12	6	6	10	11	12	12	12	12
Correlation between child and parent reports	0.836**	0.229	0.097	0.261	0.067	0.429	0.893**	0.898**	0.782**

*p <0.05; **p <0.01

Pearson correlations were applied to compare the child self-report and parent proxy report for item difficulty (logit) and step calibration estimates. Correlations measure the relationship or similarity between two scales of estimates.

Correlations between child and parent version Social Wellbeing and Acceptance Scales were highly significant in the item difficulty ($r=0.836$, $p<0.01$), Step 6 ($r=0.893$, $p<0.01$), Step 7 ($r=0.898$, $p<0.01$), and Step 8 ($r=0.782$, $p<0.01$).

Thus, in Social Wellbeing and Acceptance Scale, collapsing the original 9 categories into 4 categories is recommended.

Table 4.3.2

Rasch half-point thresholds, logits, and correlations between child and parent versions of Functioning Scales

		Logit	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Child report										
	Item									
	8	0	-1.45	-0.89	-0.57	-0.29	0.01	0.37	0.89	1.81
	22	-0.04			-1.51	-0.88	-0.45	0.05	0.75	1.75
	23	0.39	-1.27	-0.71	-0.38	-0.04	0.41	0.97	1.61	2.51
	24	0.06	-1.6	-1.02	-0.67	-0.3	0.18	0.69	1.2	1.97
	27	-0.31	-1.58	-1.11	-0.84	-0.6	-0.33	0.03	0.52	1.34
	29	0.23	-0.98	-0.48	-0.23	-0.01	0.23	0.54	0.97	1.76
	34	0.07	-1.44	-0.78	-0.43	-0.14	0.17	0.52	0.95	1.73
	35	0.04	-1.55	-0.88	-0.49	-0.15	0.18	0.53	0.96	1.73
	37	0.1	-1.35	-0.72	-0.37	-0.1	0.17	0.48	0.91	1.72
	38	0.22	-1.17	-0.51	-0.19	0.05	0.29	0.57	0.97	1.72
	39	-0.09	-1.28	-0.69	-0.44	-0.25	-0.06	0.18	0.54	1.3
	40	0.18	-1.03	-0.5	-0.23	-0.01	0.22	0.49	0.87	1.58
Parent report										
	Item									
	8	-0.21	-1.7	-1.14	-0.85	-0.6	-0.32	0.12	0.84	1.91
	22	-0.19	-1.31	-0.95	-0.76	-0.59	-0.4	-0.09	0.61	1.81
	23	0.31	-1.26	-0.74	-0.43	-0.13	0.3	1.1	3.18	
	24	0.16	-1.34	-0.88	-0.6	-0.32	0.1	0.93	3.07	
	27	-0.45		-2.13	-1.24	-0.88	-0.58	-0.19	0.44	1.55
	29	0.31	-1.39	-0.82	-0.49	-0.16	0.24	0.76	1.44	2.71
	34	0.36	-1.34	-0.84	-0.54	-0.23	0.22	0.89	1.73	2.84
	35	0.1	-1.44	-0.86	-0.54	-0.25	0.08	0.53	1.14	2.08
	37	0.13	-1.53	-0.92	-0.57	-0.27	0.06	0.52	1.22	2.38
	38	0.4	-1.13	-0.51	-0.18	0.11	0.41	0.81	1.36	2.3
	39	-0.03	-1.22	-0.73	-0.49	-0.28	-0.07	0.23	0.7	1.57
	40	0.38	-1.15	-0.57	-0.27	0	0.3	0.71	1.37	2.54
Number of paired item		12	10	11	12	12	12	12	12	10
Correlation between child and parent reports		0.853**	0.47	0.743**	0.626*	0.831**	0.831**	0.907**	0.929**	0.575

*p <0.05; **p <0.01

Correlations between child and parent version Functioning Scales was significant in Step 3 ($r=0.626$, $p<0.05$), and were highly significant in the item difficulty ($r=0.853$, $p<0.01$), Step 2 ($r=0.743$, $p<0.01$), Step 3 ($r=0.626$, $p<0.01$), Step 4 ($r=0.831$, $p<0.01$), Step 6 ($r=0.907$, $p<0.01$), and Step 7 ($r=0.929$, $p<0.01$).

Thus, in Functioning Scale, collapsing the original 9 categories into 7 categories is recommended.

Table 4.3.3

Rasch half-point thresholds, logits, and correlations between child and parent versions of Participation and Physical Health Scales

		Logit	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Child report										
	Item									
	9	-0.26	-1.96	-1.09	-0.69	-0.39	-0.1	0.22	0.64	1.37
	16	-0.14	-1.22	-0.78	-0.54	-0.33	-0.12	0.14	0.5	1.18
	17	-0.13	-1.51	-0.96	-0.66	-0.39	-0.09	0.27	0.74	1.53
	18	-0.2	-1.82	-1.09	-0.75	-0.45	-0.12	0.29	0.82	1.64
	19	0.09	-1.17	-0.72	-0.45	-0.18	0.13	0.49	0.9	1.61
	20	-0.08	-1.67	-0.96	-0.59	-0.29	0.03	0.4	0.86	1.64
	21	0.02	-1.73	-1	-0.59	-0.2	0.21	0.62	1.07	1.82
	25	-0.13		-2.99	-0.8	-0.28	0.09	0.48	0.96	1.73
	26	-0.02	-1.59	-0.86	-0.48	-0.19	0.08	0.39	0.83	1.65
	30	0.4	-0.94	-0.38	-0.09	0.17	0.44	0.77	1.22	1.98
	36	0.33	-1.22	-0.53	-0.15	0.14	0.42	0.74	1.18	2
Parent report										
	Item									
	9	-0.24		-2.04	-1.06	-0.65	-0.3	0.13	0.73	1.67
	16	-0.05	-1.12	-0.7	-0.49	-0.32	-0.12	0.14	0.6	1.51
	17	-0.04	-1.43	-0.98	-0.73	-0.5	-0.21	0.28	1.08	2.15
	18	-0.02	-1.34	-0.82	-0.56	-0.34	-0.09	0.29	0.9	1.81
	19	0.14	-1.21	-0.73	-0.46	-0.23	0.06	0.49	1.12	2.03
	20	0.02	-1.32	-0.83	-0.57	-0.33	-0.06	0.37	0.99	1.91
	21	0.16	-1.14	-0.68	-0.44	-0.21	0.07	0.5	1.1	2.02
	25	-0.15	-2.22	-1.43	-0.9	-0.49	-0.11	0.37	1.15	2.35
	26	-0.04	-1.51	-0.86	-0.55	-0.3	-0.05	0.28	0.83	1.82
	30	0.46	-1.4	-0.7	-0.3	0.06	0.49	1.04	1.71	2.72
	36	0.33	-1.49	-0.76	-0.32	0.04	0.4	0.82	1.42	2.44
Number of paired item		11	9	11	11	11	11	11	11	11
Correlation between child and parent reports		0.952**	-0.032	0.494	0.756**	0.882**	0.938**	0.936**	0.910**	0.849**

*p <0.05; **p <0.01

Correlations between child and parent version Participation and Physical Health Scales were highly significant in the item difficulty ($r=0.952$, $p<0.01$), Step 3 ($r=0.756$, $p<0.01$), Step 4 ($r=0.882$, $p<0.01$), Step 5 ($r=0.938$, $p<0.01$), Step 6 ($r=0.936$, $p<0.01$), Step 7 ($r=0.910$, $p<0.01$), and Step 8 ($r=0.849$, $p<0.01$).

Thus, in Participation and Physical Health Scale, collapsing the original 9 categories into 7 categories is recommended.

Table 4.3.4

Rasch half-point thresholds, logits, and correlations between child and parent versions of Emotional Wellbeing Scales

		Logit	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Child report										
	Item									
	2	-0.32	-1.51	-1.05	-0.82	-0.63	-0.42	-0.13	0.44	1.44
	28	-0.23	-1.53	-1.11	-0.87	-0.64	-0.35	0.1	0.76	1.7
	31	-0.23	-1.66	-1.09	-0.8	-0.55	-0.26	0.15	0.74	1.63
	32	0.1			-1.39	-0.66	-0.18	0.29	0.83	1.68
	33	0.1	-1.3	-0.74	-0.43	-0.16	0.14	0.5	0.96	1.79
	52	-0.28	-1.59	-1.1	-0.82	-0.58	-0.29	0.11	0.62	1.39
Parent report										
	Item									
	2	-0.45					-1.95	-0.92	-0.03	1.08
	28	-0.4	-1.76	-1.33	-1.09	-0.87	-0.6	-0.17	0.63	1.81
	31	-0.19	-1.96	-1.36	-1.03	-0.73	-0.39	0.15	1.15	2.53
	32	-0.04	-1.35	-0.99	-0.77	-0.57	-0.3	0.2	1.07	2.22
	33	0.37	-1.36	-0.85	-0.53	-0.19	0.3	0.96	1.69	2.8
	52	-0.35	-1.84	-1.35	-1.08	-0.81	-0.49	0.02	0.79	1.84
Number of paired item		6	4	4	5	5	6	6	6	6
Correlation between child and parent reports		0.879*	0.999**	0.995**	0.23	0.862	0.805	0.974**	0.920**	0.722

*p <0.05; **p <0.01

Correlations between child and parent version Emotional Wellbeing Scales was significant in item difficulty ($r=0.879$, $p<0.05$), and were highly significant in the Step 1 ($r=0.999$, $p<0.01$), Step 2 ($r=0.995$, $p<0.01$), Step 6 ($r=0.974$, $p<0.01$), and Step 7 ($r=0.920$, $p<0.01$).

Thus, in Emotional Wellbeing Scale, collapsing the original 9 categories into 5 categories is recommended.

Table 4.3.5

Rasch half-point thresholds, logits, and correlations between child and parent versions of Pain and Impact of Disability Scales

		Logit	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Child report										
	Item									
	44	0.12	-1.12	-0.53	-0.21	0.05	0.28	0.5	0.77	1.27
	45	0.19	-0.88	-0.35	-0.08	0.12	0.31	0.51	0.75	1.2
	46	0.12	-1.04	-0.48	-0.2	0.03	0.24	0.47	0.74	1.27
	47	1.08	-1.24	-0.22	0.48	1.29	1.76	2	2.22	2.55
	48	0.29	-0.63	-0.19	0.03	0.21	0.38	0.55	0.78	1.23
	49	-0.07	-1.22	-0.71	-0.42	-0.19	0.04	0.27	0.56	1.09
	50	0.01	-1.12	-0.62	-0.34	-0.1	0.14	0.37	0.64	1.12
	51	-0.04	-1.28	-0.72	-0.4	-0.13	0.11	0.35	0.63	1.12
Parent report										
	Item									
	44	0.02	-1.23	-0.67	-0.37	-0.12	0.12	0.37	0.69	1.3
	45	0.08	-0.93	-0.47	-0.24	-0.06	0.12	0.33	0.62	1.22
	46	-0.14	-1.63	-0.91	-0.56	-0.26	0.02	0.31	0.66	1.3
	47	0.01	-1.21	-0.33	0.3	1.25				
	48	0.17	-0.92	-0.43	-0.17	0.05	0.25	0.46	0.74	1.31
	49	-0.01	-1.45	-0.76	-0.39	-0.11	0.15	0.41	0.73	1.36
	50	0.1	-1.2	-0.61	-0.28	-0.01	0.23	0.48	0.77	1.37
	51	0.05	-1.24	-0.64	-0.33	-0.09	0.15	0.4	0.72	1.39
Number of paired item		8	8	8	8	8	7	7	7	7
Correlation between child and parent reports		0.005	0.584	0.688	0.858**	0.957**	0.044	-0.27	-0.488	-0.651

*p <0.05; **p <0.01

Correlations between child and parent version Pain and Impact of Disability Scales were highly significant in the Step 3 ($r=0.858$, $p<0.01$) and Step 4 ($r=0.957$, $p<0.01$).

Thus, in Pain and Impact of Disability Scale, collapsing the original 9 categories into 3 categories is recommended.

4.4. Summary of the key findings

The table below summarized the statistics of item fit calibrations. For child scale, 1 item in the subscale of social well-being and acceptance and 3 items in the subscale of pain and impact of disability were recommended to exclude from the analysis and interpret separately.

Table 4.4.1

Misfit and low utility item screening

Subscale		Swb	Fun	Part	Emb	Acc	Pain	Fam
Misfit[a] item	Child	3	-	-	-	N/A	44, 45, 46, 48	N/A
	Parent	-	-	-	-	-	44, 45, 48, 50, 51	-
Low utility[b] item	Child	3	-	-	-	N/A	45, 46, 47, 48	N/A
	Parent	11	-	-	-	-	47, 48	-
Item Recommended to be separated from the scale	Child	3	-	-	-	N/A	45, 46, 48	N/A
	Parent	-	-	-	-	-	-	-

Swb : social well-being and acceptance
Fun : functioning
Part : participation and physical health
Ewb : emotional well-being
Acc : access to services
Pain : pain and impact of disability
Fam : family health

a Infit/Outfit MNSQ outside the critical range 0.5-1.5
b Point-measure correlation below 0.30

Since the parent and child questionnaires we collected were not paired-up data, we could not simply calculate the correlations between parents and children on each item; then we used the step calibration results of the two forms, and calculate the correlations between parents and children on each step, therefore the so called “item” was step now, and the so called “person” was item now, the results turned out became comparable.

At the very beginning, we assumed that the child form and parent form was somehow coherent, the recommendation of collapsing categories for each subscale was based on the step correlations between child form and parent form. If the correlation, r, was higher than 0.6 and the p-value was smaller than 0.05, we could conclude that this step can effectively differentiate persons with different abilities.

For those items or subscales with low correlation between child and parent form, it implies a different point of view between children and parents in the certain content or the item/subscale itself was poorly defined and might somehow confuse the participants (need to further investigate on the individual fit statistics in order to make a precise estimate).

Table 4.4.2

Quality of subscale coherence on measure and steps between Child scale and Parent scale

Subscale	Logit	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Swb	Good	Poor	Poor	Poor	Poor	Poor	Good	Good	Good
Fun	Good	Poor	Good	Good	Good	Good	Good	Good	Poor
Part	Good	Poor	Poor	Good	Good	Good	Good	Good	Good
Emb	Good	Good	Good	Poor	Poor	Poor	Good	Good	Poor
Pain	Poor	Poor	Poor	Good	Good	Poor	Poor	Poor	Poor

Swb : social well-being and acceptance

Fun : functioning

Part : participation and physical health

Ewb : emotional well-being

Acc : access to services

Pain : pain and impact of disability

Fam : family health

Good: $r > 0.6$ and $p < 0.05$

Poor: $r = 0.6$ or below and $p = 0.05$ or above

CHAPTER 5

DISCUSSION

5.1. Do items separate persons appropriately and consequently provide an acceptable measure quality?

In answer to this question, we can directly observe the person-item map, and check the distribution of persons and items. When the quality of the scale is good, items will have different difficulties, as well as their allocation can separated persons with different “ability” level sufficiently.

(“Ability” with the characteristics of this questionnaire accordingly, refers to the Quality of Life [QOL])

According to the person-item map of Social Wellbeing and Acceptance Scale of Child self report (see Figure 4.1.1, Page 81), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -0.4 to 1.2 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute on near the bottom of the vertical midline, and their distribution range is approximately half of the persons, ranges from -0.59 to 0 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons thus cannot be sufficiently distinguished.

To improve the quality of this scale, designing a series of “high difficulty” items is crucial, so that persons in a relatively high ability level will be sufficiently distinguished.
(“High difficulty” refers to a person difficult to select a high score in this item; it may due to the nature of the item, i.e., may involve higher than normal quality of life, or reach to the luxury standard.)

According to the person-item map of Social Wellbeing and Acceptance Scale of Parent proxy report (see Figure 4.2.1a, Page 140), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -0.9 to 1.5 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute on near the bottom of the vertical midline, and their distribution range is approximately less than half of the persons, ranges from -0.62 to -0.03 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons and a few low ability persons thus cannot be sufficiently distinguished. This situation is similar to Social Wellbeing and Acceptance Scale of child self report.

To improve the quality of this scale, designing a series of “high difficulty” items, 1-2 less difficulty (easy) items, and test it properly, are crucial so that persons in a relatively high ability level and a few one in the relatively low ability level will be sufficiently distinguished.

According to the person-item map of Functioning Scale of Child self report (see Figure 4.1.2a, Page 88), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -0.4 to 1.3 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute from the middle to near the bottom of the vertical midline, and their distribution range is

approximately half of the persons, ranges from -0.31 to 0.39 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons thus cannot be sufficiently distinguished.

To improve the quality of this scale, designing a series of “high difficulty” items is crucial, so that persons in a relatively high ability level will be sufficiently distinguished.

According to the person-item map of Functioning Scale of Parent proxy report (see Figure 4.2.2a, Page 147), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -1 to 1.6 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute from the middle to near the bottom of the vertical midline, and their distribution range is approximately less than half of the persons, ranges from -0.45 to 0.4 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons and a few

low ability persons thus cannot be sufficiently distinguished. This situation is similar to Functioning Scale of child self report.

To improve the quality of this scale, designing a series of “high difficulty” items, 1-3 less difficulty (easy) items, and test it properly, are crucial so that persons in a relatively high ability level and a few one in the relatively low ability level will be sufficiently distinguished.

According to the person-item map of Participation and Physical Health Scale of Child self report (see Figure 4.1.3a, Page 93), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -0.4 to 1.3 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute from the middle to near the bottom of the vertical midline, and their distribution range is approximately half of the persons, ranges from -0.26 to 0.4 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons thus cannot be sufficiently distinguished.

To improve the quality of this scale, designing a series of “high difficulty” items is crucial, so that persons in a relatively high ability level will be sufficiently distinguished.

According to the person-item map of Participation and Physical Health Scale of Parent proxy report (see Figure 4.2.3a, Page 154), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -0.9 to 1.5 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute in the middle of the vertical midline, and their distribution range is approximately half of the persons, ranges from -0.24 to 0.46 logits. In other words, this group of items can only be distinguished half of the persons according to their ability, and only work in the middle zone, most of the high ability persons and a few low ability persons thus cannot be sufficiently distinguished. This situation is similar to Participation and Physical Health Scale of child self report.

To improve the quality of this scale, designing a series of “high difficulty” and “low difficulty” items are crucial so that persons in a relatively high ability level and a relatively low ability level will be sufficiently distinguished.

According to the person-item map of Emotional Wellbeing Scale of Child self report (see Figure 4.1.4a, Page 98), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -0.5 to 1.3 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute on near the bottom of the vertical midline, and their distribution range is approximately half of the persons, ranges from -0.32 to 0.1 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons thus cannot be sufficiently distinguished.

To improve the quality of this scale, designing a series of “high difficulty” items is crucial, so that persons in a relatively high ability level will be sufficiently distinguished.

According to the person-item map of Emotional Wellbeing Scale of Parent proxy report (see Figure 4.2.4a, Page 159), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -1 to 1.5 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute from the middle to near the bottom of the vertical midline, and their distribution range is approximately half of the persons, ranges from -0.45 to 0.37 logits. In other words, this group of items can only be distinguished half of the persons according to their ability, and only work in the lower middle zone, most of the high ability persons and a few low ability persons thus cannot be sufficiently distinguished. This situation is similar to Emotional Wellbeing Scale of child self report.

To improve the quality of this scale, designing a series of “high difficulty” items, 5-6 less difficulty (easy) items, and test it properly, are crucial so that persons in a relatively high ability level and a few one in the relatively low ability level will be sufficiently distinguished.

According to the person-item map of Pain and Impact of Disability Scale of Child self report (see Figure 4.1.5a, Page 103), persons locate on the left side of the diagram, exhibit a wide

range of distribution, and range from -0.5 to 1.3 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute from the middle to near the bottom (except item 47) of the vertical midline, and their distribution range is approximately half of the persons, ranges from -0.07 to 1.08 logits. In other words, this group of items can only be distinguished approximately half of the persons according to their ability, and only work in the middle to low zone, most of the high ability persons thus cannot be sufficiently distinguished.

To improve the quality of this scale, designing a series of “high difficulty” items is crucial, so that persons in a relatively high ability level will be sufficiently distinguished.

According to the person-item map of Pain and Impact of Disability Scale of Parent proxy report (see Figure 4.2.5a, Page 164), persons locate on the left side of the diagram, exhibit a wide range of distribution, and range from -1 to 1.6 logits. Fewer persons locate in the highest and the lowest ends, many of persons converge to the central, and gradually form a bell shape.

Compare to the person, items locate on the right side of the diagram, they distribute in the lower middle of the vertical midline, and their distribution range is approximately less than half of the persons, ranges from -0.14 to 0.17 logits. In other words, this group of items can only be distinguished less than half of the persons according to their ability, and only work in the lower middle zone, most of the high ability persons and a few low ability persons thus cannot be sufficiently distinguished. This situation is similar to Pain and Impact of Disability Scale of child self report.

To improve the quality of this scale, designing a series of “high difficulty” and “low difficulty” items are crucial so that persons in a relatively high ability level and a relatively low ability level will be sufficiently distinguished.

5.2. Do items exhibit an appropriate goodness of fit within each subscale?

In answer to this question, we can investigate the MNSQ value through testing infit and outfit statistics of each item. In this study, the value of the MNSQ in Likert scale was hence estimated, and generally the acceptable range is from 0.5 to 1.5, as a standard boundary. In Chapter 2 we have introduced these concepts: the MNSQ of infit is sensitive to unexpected response close to the person's trait level, and the MNSQ of outfit is more sensitive to unexpected responses on items far from the person's trait level, while the ideal value for the MNSQ is 1.0.

If the MNSQ is far from 1.0, we use ± 0.5 as boundaries, which indicate that there is something other than the trait influencing the responses on that specific item. i.e., it could be that the wording of the item itself is not understood in the same way by all respondents, or that another trait is influencing the responses on a given item. This item is named as a misfit item, because this item is contributing poorly to the measurement of the defined trait.

In the Child self report (see Table A1), it has a total of 52 items, of which, 47 items with infit and outfit MNSQ inside of the acceptable range, 0.5-1.5, exhibited a proper item-model fit, it could be seemed as fit items, which accounted for 90.4% of the whole item set.

On the other hand, there are 5 items exceed the maximum MNSQ tolerable limit of 1.5, they are item 3 (outfit, 1.85), item 44 (infit, 1.56; outfit, 1.61), item 45 (infit, 1.99; outfit, 2.5), item 46 (infit, 1.83; outfit, 2.19), and item 48 (infit, 2.01; outfit, 4.84). Item 3 is in Social Wellbeing and Acceptance subscale, item 44, 45, 46 and 48 are Pain and Impact of Disability subscale, and those items can be regarded as misfit items, which accounted for 9.6% of the whole item set.

Within the whole scale, fit items accounted for the number of over 90%; therefore, we can claim that the items in this scale exhibit an appropriate goodness of fit.

In the parent proxy report (see Table B1), a total of 65 items, including 59 items with the infit and outfit MNSQ within the acceptable range, 0.5-1.5. These items exhibit a proper item-model fit, and can be regarded as fit items, which accounted for 90.8% of the whole item set.

On the other hand, there are 6 items exceed the maximum MNSQ tolerable limit of 1.5, they are item 44 (outfit, 1.95), item 45 (infit, 1.73; outfit, 2.1), item 48 (infit, 1.8; outfit, 1.94), item 49 (outfit, 1.8), item 50 (outfit, 1.63), and item 51 (outfit, 1.72). Surprisingly, they all

belong to the same subscale - Pain and Impact of Disability, these misfit items accounted for 9.2% of the whole item set.

Within the whole scale, fit items accounted for the number of over 90%; therefore, we can claim that the items in this scale exhibit an appropriate goodness of fit. This situation is similar to the child self report by comparing its fit statistics.

However, it is noticeable that, whatever in child self report as well as parent proxy report, most of the items in the Pain and Impact of Disability subscale appeared misfit in certain cases; these values revealed that Pain and Impact of Disability subscale seems somehow problematic, and worth to be further investigated.

5.3. Do items contribute to a unidimensional construct within each subscale as well as for the full scale?

In answer to this question, we can test item MNSQ as references, and more specifically, we test each item's correlation to the entire scale, generally, the value of correlation coefficient (r) greater than 0.3 can be regarded as correlation, if the r value ranged from 0.4 to 0.6, implicates a moderate level of correlation, if the r value greater than 0.7, implicates a high level of correlation.

In addition, through investigating item characteristic curve (ICC) the Rasch half-point threshold information was obtained, comparing each subscale of the ICC in pattern, to see whether they are consistent or not, and if consistency is high, it provides evidence for unidimensional construct. Furthermore, ICC provides some useful category information, we can make appropriate adjustments regarding to those information, in order to improve the quality of an individual subscale as well as the entire scale.

In order to further investigate the issue, whether or not the items contribute to a unidimensional construct, the following discussion will be divided into two sections:

(1) examine the item-scale correlation coefficient;

(2) compare the item characteristic curves of each subscale.

5.3.1. Section I: Examine the item-scale correlation coefficient

In the child self report (see Table A1), only 5 items of item-scale correlation coefficient is less than 0.3, which includes item 3 (r = 0.26), item 45 (r = 0.22), item 46 (r = 0.26) , item 47 (r = 0.16), and item 48 (r = 0.23). Of which, item 3 is in Social Wellbeing and Acceptance subscale, and item 45, 46, 47 and 48 is in Pain and Impact of Disability subscale.

According to the item fit statistics and variable map of Pain and Impact of Disability subscale (Figure 4.1.5a; Table 4.1.5a), although item 47's item-scale correlation coefficient is the lowest one within the scale, but its MNSQ value is still within the acceptable range, 0.5-1.5, while the rest of them are out of the critical range. On the other hand, item 47 is the most difficult item (1.08 logit in item difficulty) within the Pain and Impact of Disability subscale, making it in the person-item map shows the highest relative position. If this item was

removed, the quality of Pain and Impact of Disability subscale will be reduced. It is proposed to be retained.

As an affiliate, the location of item 48 and item 45 in the person-item map shows that they can effectively distinguish persons with different ability level, and no other alternative items in the same position can take their role, so it is proposed to be retained.

To simplify this subscale in order to improve its quality, only item 46 can be considered to be removed from the scale, because within the scale, item 44 has the same item difficulty (0.12), as well as the same relative location in the person-item map, and compare with the item-scale coefficient, item 44 also has a higher value (r = 0.36) than item 46 (r = 0.26), and its value is even greater than 0.3. Although item 44's MNSQ value (infit, 1.56; outfit, 1.61) is also out of the critical range, but the deviation is far less than item 46 (infit, 1.83; outfit, 2.19).

According to the item fit statistics and the variable map of Social Wellbeing and Acceptance subscale (Figure 4.1.1a; Table 4.1.1a), to simplify this subscale in order to improve its quality, item 3 can be considered to be removed from the scale, because within the scale, item 4 has the same item difficulty (-0.14), as well as the same relative location in the person-item map, and compare with the item-scale coefficient, item 4 has a higher value (r = 0.53) than item 3

(r = 0.26), and its value is even greater than 0.3, the critical point of correlation. Moreover, item 4's MNSQ values (infit, 0.95; outfit, 0.99) are within the acceptable range of 0.5-1.5, whereas item 3 is not (outfit, 1.85).

For the full scale of child self report, most of the items' item-scale correlation coefficients are ranging from 0.4 to 0.7, which indicates that the correlations of most items within the scale are moderate or above.

In the parent proxy report (see Table B1), only 3 items of item-scale correlation coefficient is less than 0.3, which includes item 11 (r = 0.23), item 47 (r = 0.2), and item 48 (r = 0.29). Of which, item 11 is in Social Wellbeing and Acceptance subscale, and item 47 and 48 are in Pain and Impact of Disability subscale.

According to the item fit statistics and variable map of Social Wellbeing and Acceptance subscale (Figure 4.2.1a; Table 4.2.1a), to simplify this subscale in order to improve its quality, item 11 can be considered to be removed, because within the scale, item 10 has a very close item difficult (-0.62) to item 11 (-0.6), as well as the similar relative location in the person-item map, and compare with the item-scale coefficient, item 10 also has a higher

value (r = 0.47) than item 11 (r = 0.23), and it value is even greater than 0.3, the critical point of correlation.

According to the item fit statistics and variable map of Pain and Impact of Disability subscale (Figure 4.2.6a; Table 4.2.6a), although item 48 and item 49's item-scale correlation coefficients are less than 0.3, and their MNSQ values are both out of the critical range, however, their item difficulty within the Pain and Impact of Disability subscale can effectively differentiate persons with different ability levels, and no other alternative items in the same position can take their role, so these 2 items are still recommended to be retained.

For the full scale of parent proxy report, most of the items' item-scale correlation coefficients are ranging from 0.4 to 0.6, which indicates that the correlations of most items within the scale are in the level of moderate.

5.3.2. Section II: Compare the Item Characteristic Curves of each subscale

According to Social Wellbeing and Acceptance's ICC (Figure 4.1.1b; Figure 4.2.1b; Table 4.1.1b; Table 4.2.1b), in the child self report, ICC thresholds range from -1.64 to 2.04, while in the parent proxy report, ICC thresholds range from -2.09 to 2.76. Therefore, parent proxy report has a wider range of step calibration estimates.

On the other hand, in the child self report, item 6 and item 1's category 1 and 2, and item 10's category 1, 2 and 3 in this set data are out of utility. In the parent proxy report, item 7, 4 and 1's category 1 and 2, and item 3's category 1, 2, 3 and 4 in this set of data are also out of utility.

In comparison with the category utilization of the child self report and the parent proxy report, there are some inconsistencies in item 3, item 4, item 6, item 7 and item 10.

According to Functioning's ICC (Figure 4.1.2b; Figure 4.2.2b; Table 4.1.1b; Table 4.2.1b), in the child self report, ICC thresholds range from -1.58 to 2.51, while in the parent proxy report, ICC thresholds range from -1.53 to 2.84. Therefore, parent proxy report has a wider range of step calibration estimates.

On the other hand, in the child self report, item 22's category 1 and 2 in this set data are out of utility. In the parent proxy report, item 23 and item 24's category 9, and item 27's category 1 in this set of data are also out of utility.

In comparison with the category utilization of the child self report and the parent proxy report, there are some inconsistencies in item 22, item 23, item 24 and item 27.

According to Participation and Physical Health's ICC (Figure 4.1.3b; Figure 4.2.3b; Table 4.1.1b; Table 4.2.1b), in the child self report, ICC thresholds range from -1.96 to 2, while in the parent proxy report, ICC thresholds range from -2.22 to 2.72. Therefore, parent proxy report has a wider range of step calibration estimates.

On the other hand, in the child self report, item 25's category 1 in this set data is out of utility. In the parent proxy report, item 9's category 1 in this set of data is also out of utility.

In comparison with the category utilization of the child self report and the parent proxy report, there are some inconsistencies in item 9 and item 25.

According to Emotional Wellbeing's ICC (Figure 4.1.4b; Figure 4.2.4b; Table 4.1.1b; Table 4.2.1b), in the child self report, ICC thresholds range from -1.66 to 1.79, while in the parent proxy report, ICC thresholds range from -1.96 to 2.80. Therefore, parent proxy report has a wider range of step calibration estimates.

On the other hand, in the child self report, item 32's category 1 and 2 in this set data are out of utility. In the parent proxy report, item 2's category 1, 2, 3 and 4 in this set of data are also out of utility.

In comparison with the category utilization of the child self report and the parent proxy report, there are some inconsistencies in item 2 and item 32.

According to Pain and Impact of Disability's ICC (Figure 4.1.5b; Figure 4.2.6b; Table 4.1.1b; Table 4.2.1b), in the child self report, ICC thresholds range from -1.28 to 2.55, while in the parent proxy report, ICC thresholds range from -1.63 to 1.39. Therefore, parent proxy report has a wider range of step calibration estimates.

On the other hand, in the parent proxy report, item 47's category 6, 7, 8 and 9 in this set of data are out of utility, whereas no items' category is out of utility in the child self report.

In comparison with the category utilization of the child self report and the parent proxy report, there are some category inconsistencies in item 47.

The item fit statistics are summarized in Table 5.1a. The average infit and outfit were 1.00 (SD=0.29) and 1.10 (SD=0.63), respectively in the child scale, and 0.99 (SD=0.22) and 1.03 (SD=0.30), respectively in the parent scale. The fit statistics of item 3 (outfit, 0.26), item 44 (infit, 1.56; outfit, 1.61), item 45 (infit, 1.99; outfit, 2.50), item 46 (infit, 1.83; outfit, 2.19) and item 48 (infit, 2.01; outfit, 4.84) in child scale, item 44 (infit, 1.50; outfit, 1.95), item 45 (infit, 1.73; outfit, 2.10), item 48 (infit, 1.80; outfit, 1.94), item 50 (outfit, 1.63) and item 51 (outfit, 1.72) in parent scale were outside the critical range, 0.5 to 1.5; those items were identified as misfit items.

The point-measure correlation is also shown in Table 5.1a. The correlate of item 3 (0.26), item 45 (0.22), item 46 (0.23), item 47 (0.16) and item 48 (0.23) in child scale, item 11 (0.23), item 47 (0.20) and item 48 (0.29) in parent scale were below the cut point, 0.30; those items were identified as low utility items.

Table 5.1a

Item fit statistics and point-measure correlations of C CP QOL Child scale

Child self-report scale					Parent proxy scale				
Subscale	Item	Measure(SE)	Infit(Outfit)	Correlate	Subscale	Item	Measure(SE)	Infit(Outfit)	Correlate
Swb	1	-0.08(0.07)	0.85(0.85)	0.56	Swb	1	-0.25(0.08)	0.95(0.97)	0.46
	3	-0.14(0.06)	1.32(1.85)	0.26		3	-0.18(0.08)	1.04(1.04)	0.36
	4	-0.14(0.06)	0.95(0.99)	0.53		4	-0.23(0.08)	0.96(0.97)	0.43
	5	0.00(0.06)	0.89(0.93)	0.57		5	-0.23(0.07)	0.91(0.90)	0.53
	6	0.00(0.07)	0.97(0.97)	0.47		6	-0.33(0.08)	0.88(0.87)	0.51
	7	-0.34(0.07)	0.98(1.04)	0.46		7	-0.45(0.07)	0.98(1.02)	0.43
	10	-0.59(0.07)	0.94(0.94)	0.48		10	-0.62(0.07)	0.98(0.94)	0.47
	11	-0.42(0.07)	0.92(0.84)	0.50		11	-0.6090.07)	1.36(1.32)	0.23
	12	-0.18(0.06)	0.90(0.96)	0.54		12	-0.16(0.08)	0.99(0.99)	0.43
	13	-0.06(0.06)	0.79(0.80)	0.64		13	-0.10(0.07)	0.88(0.86)	0.56
	14	-0.17(0.07)	0.84(0.85)	0.56		14	-0.32(0.07)	1.04(0.98)	0.43
	15	-0.04(0.06)	1.04(1.06)	0.50		15	-0.03(0.07)	0.82(0.78)	0.61
Fun	8	0.00(0.06)	1.27(1.46)	0.34	Fun	8	-0.21(0.06)	0.95(1.00)	0.53
	22	-0.04(0.07)	0.84(0.85)	0.57		22	-0.19(0.06)	0.93(0.96)	0.52
	23	0.39(0.06)	1.20(1.28)	0.34		23	0.31(0.06)	0.97(1.05)	0.51
	24	0.06(0.06)	0.94(0.94)	0.55		24	0.16(0.06)	0.78(0.80)	0.61
	27	-0.31(0.06)	0.85(0.81)	0.57		27	-0.45(0.07)	0.97(0.98)	0.48
	29	0.23(0.05)	1.11(1.36)	0.49		29	0.31(0.06)	1.06(1.13)	0.48
	34	0.07(0.05)	0.74(0.73)	0.70		34	0.36(0.06)	0.90(0.92)	0.55
	35	0.04(0.05)	0.89(0.90)	0.61		35	0.10(0.05)	0.88(0.90)	0.60
	37	0.10(0.05)	0.94(0.97)	0.59		37	0.13(0.05)	0.80(0.80)	0.65
	38	0.22(0.05)	0.89(0.91)	0.63		38	0.40(0.05)	0.93(0.94)	0.61
	39	-0.09(0.05)	0.83(0.82)	0.61		39	-0.03(0.05)	0.84(0.85)	0.65
	40	0.18(0.05)	1.02(1.09)	0.56		40	0.38(0.05)	0.92(0.95)	0.60
Part	9	-0.26(0.05)	0.95(1.00)	0.53	Part	9	-0.24(0.06)	0.88(0.88)	0.58
	16	-0.14(0.05)	0.92(0.89)	0.58		16	-0.05(0.05)	0.94(0.88)	0.60
	17	-0.13(0.06)	0.76(0.75)	0.67		17	-0.04(0.06)	0.70(0.71)	0.67
	18	-0.20(0.06)	0.80(0.79)	0.63		18	-0.02(0.05)	0.74(0.72)	0.69
	19	0.09(0.05)	0.90(0.90)	0.61		19	0.14(0.05)	0.91(0.90)	0.57
	20	-0.08(0.05)	0.85(0.85)	0.61		20	0.02(0.05)	0.78(0.79)	0.66
	21	0.02(0.06)	0.92(0.91)	0.59		21	0.16(0.05)	0.86(0.88)	0.61
	25	-0.13(0.06)	0.81(0.81)	0.65		25	-0.15(0.06)	0.83(0.85)	0.61

Table 3 *(Continued)*

Child self-report scale					Parent proxy scale				
Subscale	Item	Measure(SE)	Infit(Outfit)	Correlate	Subscale	Item	Measure(SE)	Infit(Outfit)	Correlate
Part	26	-0.02(0.05)	0.74(0.71)	0.69	Part	26	-0.04(0.05)	0.78(0.74)	0.68
	30	0.40(0.05)	0.94(1.02)	0.61		30	0.46(0.05)	0.97(0.98)	0.56
	36	0.33(0.05)	0.88(0.88)	0.65		36	0.33(0.05)	0.84(0.85)	0.65
Emb	2	-0.32(0.07)	0.89(0.86)	0.49	Emb	2	-0.45(0.08)	1.01(1.03)	0.37
	28	-0.23(0.07)	0.83(0.83)	0.58		28	-0.40(0.07)	0.75(0.75)	0.60
	31	-0.23(0.06)	0.66(0.64)	0.72		31	-0.19(0.07)	0.74(0.76)	0.65
	32	0.10(0.06)	0.67(0.68)	0.73		32	-0.04(0.06)	0.75(0.78)	0.63
	33	0.10(0.05)	0.77(0.77)	0.68		33	0.37(0.06)	0.92(0.92)	0.55
	52	-0.28(0.06)	0.72(0.69)	0.68		52	-0.35(0.07)	0.85(0.87)	0.59
					Acc	41	-0.08(0.08)	0.92(0.92)	0.53
						42	0.05(0.07)	0.96(0.99)	0.51
						43	0.24(0.06)	1.01(1.08)	0.52
						53	-0.23(0.07)	0.97(0.91)	0.51
						54	-0.17(0.06)	0.98(0.90)	0.54
						55	0.06(0.05)	1.00(0.97)	0.55
						56	-0.04(0.06)	1.04(1.05)	0.52
						57	0.31(0.11)	0.95(0.95)	0.49
						58	0.30(0.10)	0.95(0.97)	0.51
						59	0.39(0.10)	1.08(1.11)	0.42
						60	0.55(0.05)	0.90(0.90)	0.65
						61	0.27(0.05)	1.00(1.00)	0.57
Pain	44	0.12(0.05)	1.56(1.61)	0.36	Pain	44	0.02(0.04)	1.50(1.95)	0.34
	45	0.19(0.04)	1.99(2.50)	0.22		45	0.08(0.04)	1.73(2.10)	0.32
	46	0.12(0.04)	1.83(2.19)	0.23		46	-0.14(0.05)	1.34(1.41)	0.40
	47	1.08(0.07)	1.42(1.39)	0.16		47	0.01(0.07)	1.27(1.43)	0.20
	48	0.29(0.04)	2.01(4.84)	0.23		48	0.17(0.04)	1.80(1.94)	0.29
	49	-0.07(0.05)	1.15(1.12)	0.50		49	-0.01(0.05)	1.47(1.80)	0.34
	50	0.01(0.05)	1.23(1.22)	0.48		50	0.10(0.04)	1.42(1.63)	0.40
	51	-0.04(0.05)	1.18(1.22)	0.48		51	0.05(0.04)	1.44(1.72)	0.38
					Fam	62	0.03(0.06)	1.04(1.01)	0.48
						63	0.23(0.06)	1.04(1.03)	0.50
						64	0.40(0.05)	1.20(1.22)	0.42
						65	0.15(0.05)	0.95(0.94)	0.59

Swb : social well-being and acceptance

Fun : functioning

Part : participation and physical health
Ewb : emotional well-being
Acc : access to services
Pain : pain and impact of disability
Fam : family health

Table 5.1b

Do items contribute to a unidimensional construct of each subscale?

Scale		Do items contribute to a unidimensional construct?	Misfit[a] item	Low utility[b] item	Item recommended to be separated from the scale
Child	Swb	No	3	3	3
	Fun	Yes	-	-	-
	Part	Yes	-	-	-
	Emb	Yes	-	-	-
	Pain	No	44, 45, 46, 48	45, 46, 47, 48	45, 46, 48
Parent	Swb	No	-	11	-
	Fun	Yes	-	-	-
	Part	Yes	-	-	-
	Emb	Yes	-	-	-
	Acc	Yes	-	-	-
	Pain	No	44, 45, 48, 50, 51	47, 48	48
	Fam	Yes	-	-	-

Swb : social well-being and acceptance
Fun : functioning
Part : participation and physical health
Ewb : emotional well-being
Acc : access to services
Pain : pain and impact of disability
Fam : family health

[a] Infit/Outfit MNSQ outside the critical range 0.5-1.5
[b] Point-measure correlation below 0.30

According to the results, in child form, within the “Swb” subscale, itme 3 (“the way you get along with your brothers and sisters?”), on one hand it was a misfit item, on the other hand it was a low utility item, to develop a unidimensional construct, this kind of item would recommend to be separated from the scale since it shows the characteristics of both inconsistency and “irrelevance” for that particular scale. Besides, with the “Pain” subscale, item 45 (“are you bothered when you miss school for health reasons?”), item 46 (“are you bothered by being handled by other people?”) and item 48 (“are you concerned about having cerebral palsy?”) share the same characteristics as item 3; in parent form, the similar case appears in item 48 (“is your child concerned about having cerebral palsy?”) (Table 5.1b).

Regarding to the test of unidimensionality, the summary of each subscale was shown in the above table (Table 5.1b). For the full scale construct, we can see that there are five misfit items (item 3, item 44, item 45, item 46 and item 48) and five low utility items (item 3, item 45, item 46 item 47 and item 48) in the child scale; and five misfit items (item 44, item 45, item 48, item 50 and item 51) and three low utility items (item 11, item 47 and item 48) in the parent scale. Thus the unidimensional construct of these two scales could not be justified.

5.4. Question 4: does child self report correlate to parent proxy report?

In answering this question, we perform the Pearson product correlation test for the results of child self report and parent proxy report, but since we have separated the data collection stage into two phases, and also during the process, most of the questionnaires were answered only by individual children or parents solely, the responses from child hence cannot be paired up with his/her parent, then, the correlation analysis cannot be performed, therefore, we conduct the analysis at the item level, with each corresponding step calibration results for comparison, thus, in this analysis, the number of paired items is the sample size (n), and the items' category itself is the new "item", and the currently new correlation analysis results will be discussed below.

In the following part, we compare the Pearson correlation of item difficulty and step calibration estimates of each subscale, as well as the full scale, one by one.

According to the results of Pearson correlation analysis in the step calibrations between child and parent reports in Social Wellbeing and Acceptance subscale (Table 4.3.1b), we can see that the correlation coefficient value greater than 0.3 includes step 5 (r=0.429), step 6 (r=0.893), step 7 (r=0.898) and step 8 (r=0.782). However, not all the correlations are significant ($p < 0.05$), according to the results (Table 5.1), only step 6 ($r = 0.893$, $p = 0.000$), step 7 ($r = 0.898$, $p = 0.000$) and step 8 ($r = 0.782$, $p = 0.003$) are with significant correlation.

In addition, both child version and parent version scales display a high level of correlation ($r > 0.8$) in item difficulty (r=0.836), and this correlation is highly significant ($p < 0.01$) (see Table 5.1).

Table 5.1

Correlation and p-values in item difficulty and step calibrations between child and parent versions of Social Wellbeing and Acceptance subscale

	Item difficulty	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Pearson correlation	0.836 (n=12)	0.229 (n=6)	0.097 (n=6)	0.261 (n=10)	0.067 (n=11)	0.429 (n=12)	0.893 (n=12)	0.898 (n=12)	0.782 (n=12)
p-value	0.001	0.662	0.855	0.466	0.845	0.164	0.000	0.000	0.003

According to the results of Pearson correlation analysis in the step calibrations between child and parent reports in Functioning subscale (Table 4.3.2b), we can see that the correlation coefficient value greater than 0.3 includes step 1 (r=0.470), step 2 (r=0.743), step 3 (r=0.626), step 4 (r=0.831), step 5 (r=0.907), step 6 (r=0.929), step 7 (r=0.903) and step 8 (r=0.575). However, not all the correlations are significant ($p < 0.05$), according to the results (Table 5.2), only step 2 ($r = 0.743$, $p = 0.009$), step 3 ($r = 0.626$, $p = 0.029$), step 4 ($r = 0.831$, $p = 0.001$), step 5 ($r = 0.907$, $p = 0.000$), step 6 ($r = 0.929$, $p = 0.000$) and step 7 ($r = 0.903$, $p = 0.000$) are with significant correlation.

In addition, both child version and parent version scales display a high level of correlation ($r > 0.8$) in item difficulty (r=0.853), and this correlation is highly significant ($p < 0.01$) (see Table 5.2).

Table 5.2

Correlation and p-values in item difficulty and step calibrations between child and parent versions of Functioning subscale

	Item difficulty	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Pearson correlation	0.853 (n=12)	0.470 (n=10)	0.743 (n=11)	0.626 (n=12)	0.831 (n=12)	0.907 (n=12)	0.929 (n=12)	0.903 (n=12)	0.575 (n=10)
p-value	0.000	0.0170	0.009	0.029	0.001	0.000	0.000	0.000	0.082

According to the results of Pearson correlation analysis in the step calibrations between child and parent reports in Participation and Physical Health subscale (Table 4.3.3b), we can see that the correlation coefficient value greater than 0.3 includes step 2 (r=0.494), step 3 (r=0.756), step 4 (r=0.882), step 5 (r=0.938), step 6 (r=0.936), step 7 (r=0.910) and step 8 (r=0.849). However, not all the correlations are significant ($p < 0.05$), according to the results (Table 5.3), only step 3 ($r = 0.756$, $p = 0.007$), step 4 ($r = 0.882$, $p = 0.000$), step 5 ($r = 0.938$, $p = 0.000$), step 6 ($r = 0.936$, $p = 0.000$), step 7 ($r = 0.910$, $p = 0.000$) and step 8 ($r = 0.849$, $p = 0.001$) are with significant correlation.

In addition, both child version and parent version scales display a high level of correlation ($r > 0.8$) in item difficulty (r=0.952), and this correlation is highly significant ($p < 0.01$) (see Table 5.3).

Table 5.3

Correlation and p-values in item difficulty and step calibrations between child and parent versions of Participation and Physical Health subscale

	Item difficulty	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Pearson correlation	0.952 (n=11)	-0.032 (n=9)	0.494 (n=11)	0.756 (n=11)	0.882 (n=11)	0.938 (n=11)	0.936 (n=11)	0.910 (n=11)	0.849 (n=11)
p-value	0.000	0.934	0.122	0.007	0.000	0.000	0.000	0.000	0.001

According to the results of Pearson correlation analysis in the step calibrations between child and parent reports in Emotional Wellbeing subscale (Table 4.3.4b), we can see that the correlation coefficient value greater than 0.3 includes step 1 (r=0.999), step 2 (r=0.995), step 4 (r=0.862), step 5 (r=0.805), step 6 (r=0.974), step 7 (r=0.920) and step 8 (r=0.722). However, not all the correlations are significant ($p < 0.05$), according to the results (Table 5.4), only step 1 ($r = 0.999$, $p = 0.001$), step 2 ($r = 0.995$, $p = 0.005$), step 6 ($r = 0.974$, $p = 0.001$) and step 7 ($r = 0.920$, $p = 0.009$) are with significant correlation.

In addition, both child version and parent version scales display a high level of correlation ($r > 0.8$) in item difficulty (r=0.879), and this correlation is highly significant ($p < 0.01$) (see Table 5.4).

Table 5.4

Correlation and p-values in item difficulty and step calibrations between child and parent versions of Emotional Wellbeing subscale

	Item difficulty	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Pearson correlation	0.879 (n=6)	0.999 (n=4)	0.995 (n=4)	0.230 (n=5)	0.862 (n=5)	0.805 (n=6)	0.974 (n=6)	0.920 (n=6)	0.722 (n=6)
p-value	0.021	0.001	0.005	0.710	0.060	0.053	0.001	0.009	0.105

According to the results of Pearson correlation analysis in the step calibrations between child and parent reports in Pain and Impact of Disability subscale (Table 4.3.5b), we can see that the correlation coefficient value greater than 0.3 includes step 1 (r=0.584), step 2 (r=0.688), step 3 (r=0.858), step 4 (r=0.957), step 7 (r=-0.488) and step 8 (r=-0.651). However, not all the correlations are significant ($p < 0.05$), according to the results (Table 5.5), only step 3 ($r = 0.858$, $p = 0.000$) and step 4 ($r = 0.957$, $p = 0.000$) are with significant correlation.

Table 5.5

Correlation and p-values in item difficulty and step calibrations between child and parent versions of Pain and Impact of Disability subscale

	Item difficulty	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Pearson correlation	0.005 (n=52)	0.584 (n=40)	0.688 (n=43)	0.858 (n=49)	0.957 (n=50)	0.044 (n=51)	-0.270 (n=51)	-0.488 (n=51)	-0.651 (n=49)
p-value	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000

5.5. Correlations between child self report and parent proxy report (full scale)

Finally, we compare the correlations between child self report and parent proxy report's item difficulty and step calibration estimates (see Table 5.6), and discuss as below.

Table 5.6

Correlation and p-values in item difficulty and step calibrations between child and parent reports (full scale)

	Item difficulty	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Pearson correlation	0.709 (n=52)	0.501 (n=40)	0.572 (n=43)	0.665 (n=49)	0.710 (n=50)	0.805 (n=51)	0.915 (n=51)	0.859 (n=51)	0.779 (n=49)
p-value	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000

From the above table (Table 5.6), we can see that both the item difficulty and step calibration estimates of the two scales are shown a high level of correlations (r ranged from 0.501 to 0.915), and these correlations are highly significant ($p < 0.01$).

To conclude, the complete scales of child and parent reports do correlate each other.

Actually, the quality of step coherence for each subscale is summarized in Table 4.4.2. For step coherence, 2 steps in “Pain”, 4 steps in “Emb”, 3 steps in “Swb”, and 6 steps in “Fun” and “Part” were identified with “good” quality ($r>0.6$; $p<0.05$); whereas the remaining steps were with “poor” quality ($r \leqq 0.6$ and $p \geqq 0.05$). The measure coherence for those subscales were all with “good” quality ($r>0.8$; $p<0.05$) except “Pain” ($r=0.05$; $p=0.001$).

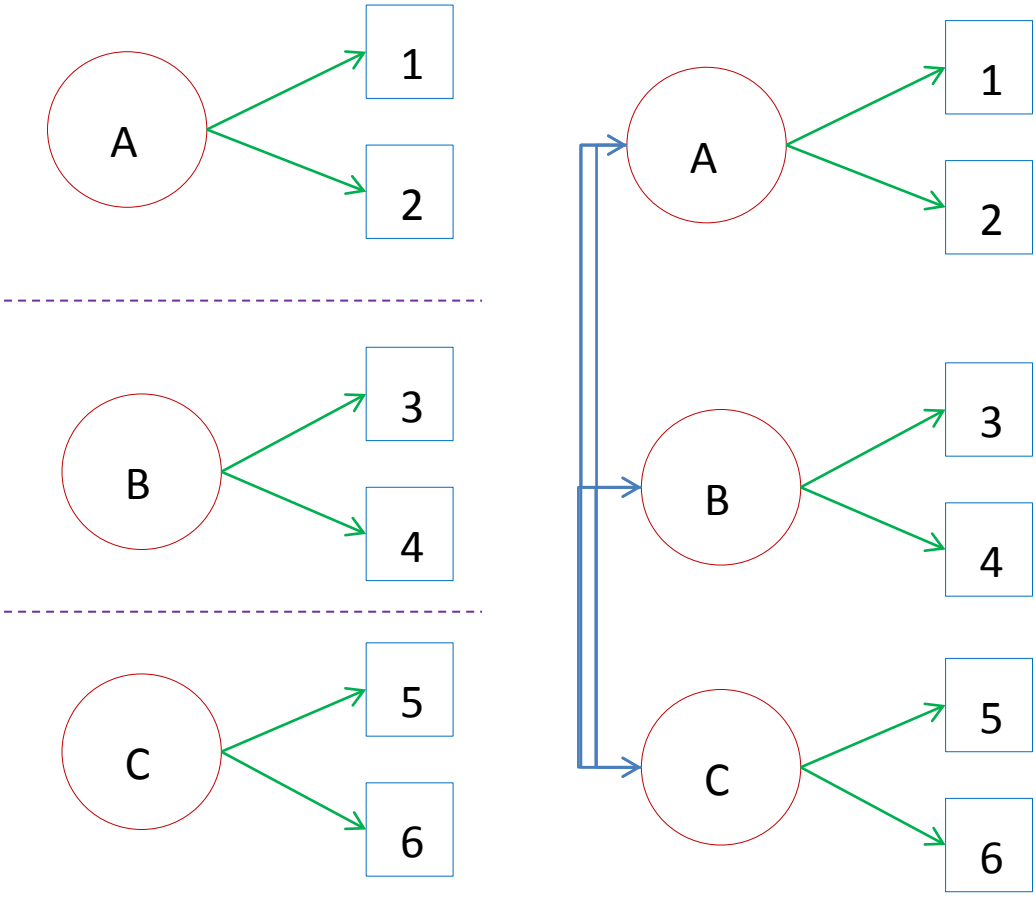
Since the C CP QOL-Child is originally designed as a 9-point Likert scale, each item has 9 categories for respondents to opt. After assessing the subscale coherence (Table 4.4.2), the well-functioning category ranged from 3 to 7 (2 to 6 steps), “Fun” and “Part” were at the top end, with 7 coherent categories; whereas “Pain” was at the low end, with 3 coherent categories, but “Pain” was the only subscale identified with poor overall measure coherence, it is problematic, possibly it may due to different perceptions between parent and child, large variation among different levels of motor disability, or etc. It is better to separate “Pain” from the scale and interpret accordingly. If so, the well-functioning category should be ranged from 4 to 7, around 5 in average. For shortening the access time of C CP QOL-Child, 5 categories are hence recommended to utilize.

5.6. Limitations of the study

In this study, the data analysis part is conducted adopting an undimensional approach, however, in accordance with the original questionnaire design, it contains several subscales, these subscales can actually be regarded as different dimensions, then using multidimensional IRT models to analysis the data set may be more appropriate, because undimensional procedures may be inefficient when data are truly multidimensional, because the unidimensional models may not adequately describe the data (Cheng, et al., 2009; Wang, et al., 2004). With undimensional approach, zero correlations between latent traits were assumed, but in reality, there are always nonzero correlations between latent traits (see Figure 5.1) (Cheng, et al., 2009; Wang, et al., 2004; Yao & Boughton, 2007).

Figure 5.1

Comparison between unidimensional and multidimensional models



Note. unidimensional model (left), multidimensional model (right); A, B and C: dimensions; 1-6: items

However, due to the small sample size, if multidimensional IRT models were adopt, the item calibration estimate may not be that accurate, and may lead to misinterpretation in the analysis results.

On the other hand, participants' living area (urban or rural), their economic status, and the time gap between two data collection phrases may be one of the latent factors affecting the

QOL perception which cannot be detected on this kind of studies, and may influence the accuracy of person-item calibrations.

5.7. Future Research

This study provided a detailed look at item level of the Chinese CP-QOL Child, but since the sample for item calibrations was recruited in Taiwan, if the Chinese CP-QOL Child is applied to another Chinese communities, i.e., mainland China, Hong Kong, or Macau, the usefulness of the Chinese CP-QOL Child in these places would need to be further investigated, as the culture between those places are not entirely the same.

Part II

Item Hierarchy of the Chinese version of Cerebral Palsy Quality of Life for Children (C-CP QOL-Child)

CHAPTER 1

INTRODUCTION

Cerebral palsy (CP) is the name used to describe a group of clinical syndromes characterized by motor deficits due to non-progressive brain damage early in life. Its incidence has been estimated as 2-2.5 per thousand live births (Levitt, 2010; Odding, Roebroeck, & Stam, 2006; Stanley & Blair, 2000). Treatment of disorders associated with these motor deficits may involve a range of interventions such as physiotherapy, orthopaedic surgery and medications for spasticity (Amichai, Harries, Dvir, Patish, & Copeliovitch, 2009; Hoving et al., 2009; Trost, Schwartz, Krach, Dunn, & Novacheck, 2008). Researchers have been using empirical-based outcome measures to guide the understanding of health condition and effectiveness of interventions for children with CP (Bagley et al., 2007; Boyd & Hays, 2001). In recent years, a crucial health model, the International Classification of Function, Disability and Health (Langerak et al., 2009; Schiariti, Fayed, Cieza, Klassen, & O'Donnell, 2011), has been used to describe the impact of health condition on a number of life domains including quality of life (QOL). Hence QOL has gained increasing attention as an important health or intervention outcome measure (Dieruf et al., 2009; Vargus-Adams & Martin, 2011). QOL has been defined by the World Health Organization as individuals' perception of their position in

life, in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns (World Health Organization, 1993).

While measuring QOL is a vital part of assessing the health condition of children with CP, scales that have been used in the past (Houlihan, O'Donnell, Conaway, & Stevenson, 2004; Wake, Ba, & Reddihough, 2003) have been found to have substantial limitations (Wake, et al., 2003). Many of the scales measure the absence of health difficulties or limitations rather than well-being; and these scales focus on physical health and functioning of children with CP without paying attention to other potentially important domains such as supportive physical environment, family health, and acceptance in the family and community. In recognizing such limitations, the Cerebral Palsy Quality of Life for Children (CP QOL-Child) was developed. It is a condition-specific QOL measure for children with CP which can be used to evaluate self perceived well-being of several life domains of such children (Waters, Maher, Salmon, Reddihough, & Boyd, 2005). It has two parallel forms, the primary caregiver proxy form and the child self-report form. The proxy form was reported to have very good reliability and validity. The child self-report form also has acceptable psychometric properties (Waters et al., 2007). Subsequently, the Chinese version of the scale (C CP QOL-Child) has been constructed (Wang et al., 2010).

Analyses based on Classical Test Theory (CTT) have been used to validate C CP QOL-Child (Wang, et al., 2010). Although the results were generally positive, two major conceptual limitations using CTT have been pointed out: the lack of an explicit ordered continuum of the test items that represent a unidimensional construct, and the lack of justification of rating scale data (Bond & Fox, 2007). In contrast to the CTT approach, the item response theory (IRT)-based Rasch analysis has gained as a potent application of psychometric testing for the development or refinement of measuring instruments (Bond & Fox, 2007). Rasch analysis provides a scaling methodology that enables the examination of the construction of true interval-scale, construct validation and also to evaluate whether the responses conformed to what would be expected, by looking at patterns of item responses and drawing probabilistic inferences (Bond & Fox, 2007). The aim of this study was to use Rasch analysis to assess further the psychometric properties of C CP QOL-Child. Specific objectives were to: (1) to establish the goodness of fit for each item; (2) to examine the extent to which C CP QOL-Child contributed to a unidimensional construct; (3) to investigate item quality by mapping item-person responses. We were particularly interested to explore, from a clinical perspective, how intervention strategies could be designed and tailor-made for individuals based on the item scores.

CHAPTER 2

MATERIALS AND METHODS

2.1. Participants

A convenience sample of primary caregivers of children with CP was recruited. Inclusion criteria were: (1) primary caregiver of a child aged 4 to 12 years, and the child was diagnosed with CP by a pediatric neurologist; (2) capable of completing the questionnaire without any assistance. Caregivers of children suffering from neurodegenerative diseases or psychiatric illness were excluded.

2.2. Measures

The primary caregiver proxy form (65 items) of C CP QOL-Child was administrated to the caregivers. The instrument measures seven domains of QOL for a child with CP: (1) social well-being and acceptance (Swb); (2) functioning (Fun); (3) participation and physical health (Part); (4) emotional well-being (Ewb); (5) access to services (Acc); (6) pain and impact of disability (Pain) and (7) family health (Fam).When using C CP QOL-Child, the stem of the test item is “How do you think your child feels about …?” or “How do you feel about …?”

This type of item stem is used because it does not measure the child's condition or her/his functioning; it assesses how the primary caregivers feel about their child's condition. The instrument uses a Likert scale. All the items except one are rated on a nine-point scale ranging from 1 to 9. One item in "Pain" is rated on a five-point scale. In addition, two conditions of motor deficits of the child were recorded. The neuromotor pattern of movements and severity of motor disability were assessed by the treating physical therapist using the neuromotor classification of CP (Tecklin, 2008), and the Gross Motor Function Classification System (GMFCS) (Palisano et al., 1997), respectively. The GMFCS levels range from Level I to Level V, with greater numerical level indicating more severe motor disability.

2.3. Procedure

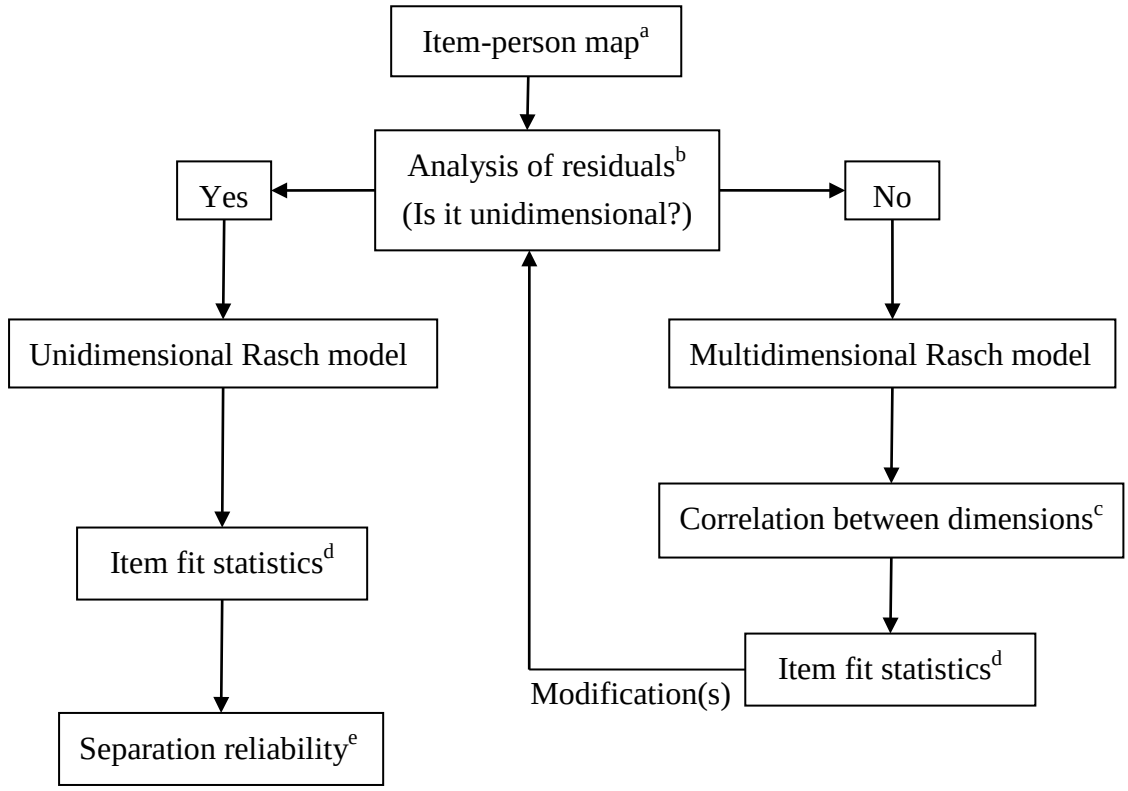
We contacted primary caregivers of children aged 4–12 with CP in the rehabilitation department of five hospitals, three rehabilitation clinics, three early intervention centers and two special education schools in Southern Taiwan and invited them to participate in the study. Ethics approval was given by Kaohsiung Medical University; informed consent was obtained prior to data collection.

2.4. Data analysis

Data were analyzed using Winsteps 3.70.1.1 (MSEA Press, Chicago, USA) and ConQuest 2.0 (ACER Press, Camberwell, Australia). Before conducting any analysis, eight items in “Pain” originally designed in a negative direction were reverse-scored in order to give all items a score in a positive direction; hence, higher scores indicating happier status or better well-being. Figure 6.1 shows the framework of the analysis; detailed explanations of the analysis are given below.

Figure 6.1.

Flow chart of the analysis



[a] give an overall picture of item-person distributions

[b] test the assumption of undimensionality of the scale

[c] test quality of each dimension

[d] test quality of each item

[e] test quality of the whole-scale

A variable map was constructed to investigate the item-person relationship and the relative person location (Bond & Fox, 2007). Statistically, the variable map shows the item distribution based on the degree of difficulty of each item. Practically, items on the upper end of the map are experienced and by children with better QOL; whereas items on lower end are experienced by more children, regardless of the QOL status.

Item fit statistics were also computed. The sufficiency of the fit of each item is assessed by the goodness of fit statistics of the residuals, mean-square of measures, including information-weighted fit (Infit) and outlier usual unweighted fit (Outfit) (Wright & Masters, 1982). These statistics provide endorsement related information by examining how participants with the same degree of QOL rate each of the items in the scale. Practically, for Likert rating scales, the range of infit and outfit values that was generally considered appropriate is between 0.6 and 1.4 (Bond & Fox, 2007; Wright, Linacre, Gustafson, & Martin-Lof, 1994). An infit and outfit value of less than 0.6 implies that the item does not provide adequate information beyond that provided by the rest of the items in the scale. This can occur when two or more items are similar or highly correlated. By contrast, an infit and outfit value greater than 1.4 implies that the item does not define the same construct as defined by the rest of the items; it is either a poorly constructed or understood item, or is ambiguously defined (Tang, Wong, Chiu, & Ungvari, 2007).

Residual analysis was conducted to assess whether the scale exhibit a unidimensional construct. Unidimensionality is the basic assumption for the total score, usually computed by summing the score of every single item in the scale, to be valid. It is essential for an unambiguous interpretation of the total score of the instrument on what it is purported to measure. If the total score of a measuring instrument does not represent a common line of

inquiry among the items, then it is unclear if two individuals with the same score can be considered comparable. If the assumption of unidimensionality is not satisfied, then to adequately describe the data, multidimensional approach might need to be adopted in order to squeeze as much information as possible from the data, and to provide a more accurate estimation of the number of constructs contained in the measurement (Cheng, Wang, & Ho, 2009; Tang, et al., 2007; Wang, Chen, & Cheng, 2004; Yao & Boughton, 2007).

Separation reliability analysis was also performed. Generally speaking, person or item reliability can be considered equivalent to KR-20, Cronbach alpha, and generalizability coefficient in CTT models. Statistically, the separation index is the ratio of "true" variance to error variance (signal-to-noise ratio, which is analogous to the Fisher discriminant ratio). For practical applications in this study, the person separation index indicates the ability of C CP QOL-Child to differentiate children with CP based on their quality of life; whereas the item separation index represents the ability to define a distinct hierarchy of order of items. The greater the separation, the more likely that groups of items will be better separated and the differences between respondents will be better distinguished (Tang, et al., 2007).

CHAPTER 3

RESULTS

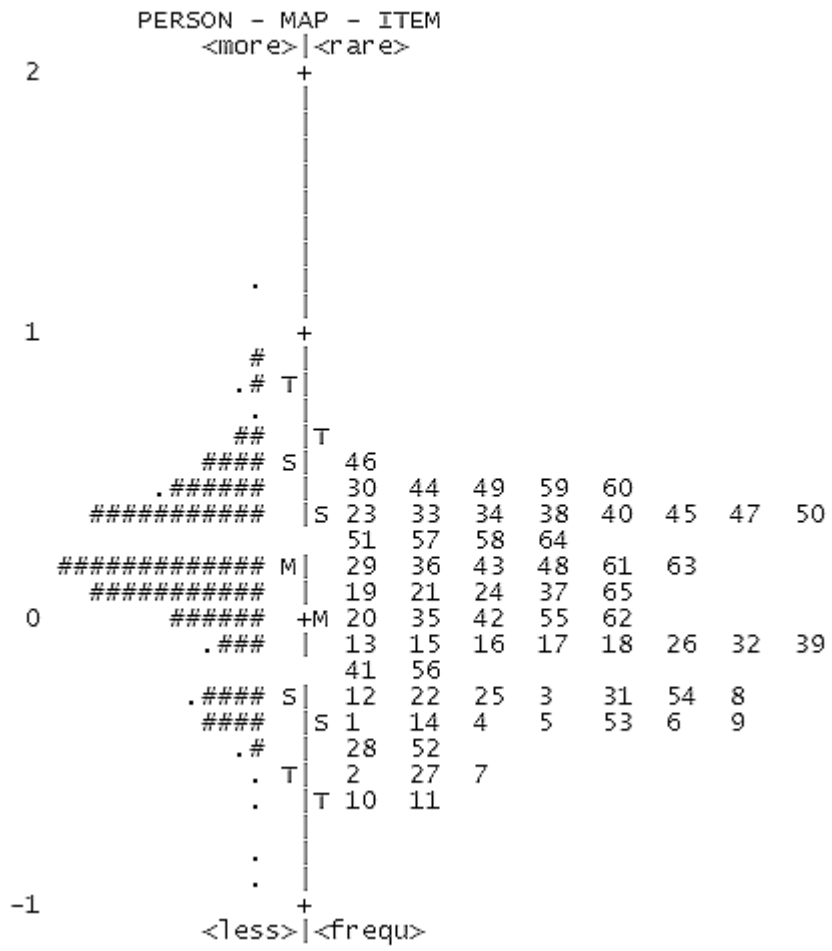
A total of 145 primary caregivers completed the proxy form of C CP QOL-Child. Table 3.1 (see Page 45) displays the demographic characteristics of the participants and their child. The mean age of the caregivers was 39.2 years old. Most of them (77.2%) were the mother, followed by father (15.9%), babysitter (4.1%), and grandmother (2.8%). Sixty-two primary caregivers (42.7%) had tertiary education level, whereas 42.1% had senior high school level. The marital status of most of the caregivers (90.3%) was married. A majority (73.1%) reported that the household income and expenditure were balanced; 30 (20.7%) reported that the household income was insufficient.

3.1. Variable map

As shown in the person-item map (Figure 6.2), generally the location of the item set was aligned with the person distribution; in the other words, the item distribution demonstrated an appropriate depth and width of difficulty for measuring person's ability. More specifically, only a few people with extreme scores (the highest 5% or the lowest 2%) could not be adequately measured by this set of items; and the items as a whole were slightly "easy" for a small proportion of people in the target population.

Figure 6.2

Variable Map of the C-CP QOL-Child Scale. Person (left side) versus Item (right side).



Each “#” is 2; each “.” is 1.

The numbers at the left margin are logits of person ability/item difficulty; on left side, each “#” is 2, each “.” is 1, number of person; on right side, item number 1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14 and 15 are in the domain of social well-being and acceptance (SWB), item number 8, 22, 23, 24, 27, 29, 34, 35, 37, 38, 39 and 40 are in the domain of functioning (FUN), item number 9, 16, 17, 18, 19, 20, 21, 25, 26, 30 and 36 are in the domain of participation and physical health (PART), item number 2, 28, 31, 32, 33 and 52 are in the domain of emotional well-being (EWB), item number 41, 42, 43, 53, 54, 55, 56, 57, 58, 59, 60 and 61 are in the domain of access to services (ACC), item number 44, 45, 46, 47, 48, 49, 50 and 51 are in the domain of pain and impact of disability (PAIN), item number 62, 63, 64 and 65 are in the domain of family health (FAM).

3.2. Analysis of residuals

Residual analysis indicated that 31.2% of the variance was explained by the measure, which was below the more ideal value of 50% or higher. The largest secondary dimension explained 8.5% of the variance, which was bigger than the maximum tolerant percentage, 5% or below, of unidimensionality (eigenvalue=8.1). The assumption of undimensionality hence could not be justified for this 65-item scale.

3.3. Correlation of dimensions

A 7-dimension Rasch model was carried out; the correlation matrix of the 7 dimensions is shown in Table 6.2. It was noted that Dimension 6 (Pain) had a small negative correlation with all other dimensions; while medium to high positive correlations were observed between the other dimensions. The reliability for dimension 1 to 7 were 0.90 (Swb), 0.97 (Fun), 0.95 (Part), 0.93 (Ewb), 0.98 (Acc), 0.79 (Pain) and 0.89 (Fam), respectively. In order to fit the criteria of unidimensionality, the eight items in dimension 6 (Pain) were thus recommended to be separated from other items and interpreted accordingly, resulting in a revised 57-item scale.

Table 6.2

Correlations between Dimensions of C-CP QOL-Child Scale

	Dimension					
Dimension	1 (Swb)	2 (Fun)	3 (Part)	4 (Ewb)	5 (Acc)	6 (Pain)
2 (Fun)	0.76					
3 (Part)	0.87	0.89				
4 (Ewb)	0.92	0.90	0.91			
5 (Acc)	0.64	0.82	0.71	0.74		
6 (Pain)	-0.23	-0.36	-0.36	-0.27	-0.16	
7 (Fam)	0.54	0.61	0.64	0.60	0.63	-0.32

Swb : social well-being and acceptance

Fun : functioning

Part : participation and physical health

Ewb : emotional well-being

Acc : access to services

Pain : pain and impact of disability

Fam : family health

3.4. Item fit statistics

The item fit statistics of the original 65-item scale and the modified 57-item scale are shown in Table 6.3. First of all, it was noted that in the 65-item scale, the infit and outfit statistics of items 44 to 51 (i.e., the "Pain" items) were outside the critical range, 0.6 to 1.4; whereas the remaining items were within the range. Deletion of these "Pain" items to form a new 57-item version of C CP QOL-Child is further justified. The average infit and outfit statistics for items were 0.98 (SD=0.33) and 1.03 (SD=0.47), respectively in the original 65-item scale; and 1.00 (SD=0.14) and 1.01 (SD=0.14), respectively in the modified 57-item scale.

In the modified 57-item scale, although item 11 ("how they are accepted by their family?") [infit, 1.53; outfit, 1.49] and item 64 ("family's financial situation") [outfit, 1.42] were out of the Likert scale critical range (i.e., 0.6-1.4), they were still fallen in the range of "clinical observation" (0.5-1.7) (Bond & Fox, 2007; Wright, et al., 1994) which is another range more commonly used in clinical settings; hence these two items are kept in the scale.

Table 6.3

Item Fit Statistics of Individual Items in Different Versions of C-CP QOL-Child Scale

	Version			
	Original 65-item scale		Modified 57-item scale	
Item	Measure(SE)	Infit(Outfit)	Measure(SE)	Infit(Outfit)
1	-0.29(0.07)	0.89(0.89)	-0.26(0.08)	1.01(1.02)
2	-0.49(0.08)	0.96(0.96)	-0.46(0.09)	1.08(1.11)
3	-0.22(0.08)	1.00(1.03)	-0.17(0.09)	1.15(1.18)
4	-0.29(0.08)	0.88(0.88)	-0.22(0.08)	0.98(0.99)
5	-0.26(0.06)	0.85(0.83)	-0.26(0.07)	0.97(0.96)
6	-0.34(0.08)	0.84(0.83)	-0.40(0.08)	0.91(0.92)
7	-0.47(0.07)	0.89(0.88)	-0.47(0.08)	1.03(1.05)
8	-0.23(0.06)	0.84(0.89)	-0.24(0.07)	1.00(1.07)
9	-0.27(0.06)	0.81(0.81)	-0.25(0.06)	0.93(0.93)
10	-0.59(0.07)	0.88(0.87)	-0.72(0.07)	1.05(1.05)
11	-0.58(0.07)	1.20(1.23)	-0.70(0.07)	1.53(1.49)
12	-0.18(0.07)	0.91(0.92)	-0.17(0.08)	1.02(1.03)
13	-0.14(0.06)	0.83(0.81)	-0.11(0.07)	0.93(0.90)
14	-0.33(0.07)	0.97(0.96)	-0.38(0.08)	1.13(1.08)
15	-0.07(0.07)	0.80(0.77)	-0.03(0.07)	0.86(0.82)
16	-0.08(0.05)	0.91(0.86)	-0.06(0.05)	1.06(1.03)
17	-0.07(0.06)	0.75(0.73)	-0.04(0.07)	0.75(0.77)
18	-0.05(0.05)	0.72(0.71)	-0.02(0.06)	0.78(0.79)
19	0.09(0.05)	0.88(0.87)	0.16(0.05)	1.03(1.05)
20	-0.02(0.05)	0.75(0.76)	0.02(0.06)	0.83(0.87)
21	0.10(0.05)	0.80(0.83)	0.18(0.05)	0.94(1.00)
22	-0.20(0.06)	0.81(0.83)	-0.23(0.07)	0.98(1.03)
23	0.26(0.06)	0.88(0.91)	0.33(0.06)	1.05(1.15)
24	0.13(0.06)	0.74(0.72)	0.16(0.07)	0.81(0.85)
25	-0.19(0.06)	0.79(0.79)	-0.16(0.06)	0.88(0.91)
26	-0.07(0.05)	0.81(0.8)	-0.04(0.05)	0.87(0.85)
27	-0.46(0.06)	0.95(0.95)	-0.49(0.07)	1.09(1.10)
28	-0.40(0.07)	0.74(0.72)	-0.47(0.07)	0.78(0.79)
29	0.23(0.05)	0.89(0.90)	0.34(0.06)	1.13(1.19)
30	0.36(0.05)	0.86(0.87)	0.53(0.06)	1.04(1.06)
31	-0.21(0.07)	0.73(0.73)	-0.22(0.07)	0.75(0.78)
32	-0.07(0.06)	0.72(0.71)	-0.06(0.07)	0.78(0.81)
33	0.29(0.06)	0.84(0.85)	0.42(0.06)	0.99(1.00)

Table 6.3 *(Continued)*

	Version			
	Original 65-item scale		Modified 57-item scale	
Item	Measure(SE)	Infit(Outfit)	Measure(SE)	Infit(Outfit)
34	0.28(0.06)	0.83(0.86)	0.40(0.06)	0.97(1.02)
35	0.05(0.05)	0.83(0.84)	0.12(0.06)	0.96(1.00)
36	0.24(0.05)	0.85(0.86)	0.38(0.05)	0.96(1.00)
37	0.07(0.05)	0.77(0.78)	0.14(0.06)	0.87(0.87)
38	0.31(0.04)	0.81(0.80)	0.46(0.05)	1.01(1.00)
39	-0.06(0.05)	0.73(0.75)	-0.04(0.05)	0.87(0.91)
40	0.30(0.05)	0.79(0.79)	0.42(0.05)	0.97(0.99)
41	-0.12(0.07)	0.86(0.86)	-0.09(0.08)	0.98(0.97)
42	0.02(0.07)	0.87(0.90)	0.05(0.08)	1.03(1.06)
43	0.20(0.06)	0.96(1.02)	0.27(0.07)	1.16(1.27)
44	0.35(0.04)	1.94(2.30)	-	-
45	0.29(0.04)	1.90(2.42)	-	-
46	0.49(0.04)	1.82(2.04)	-	-
47	0.30(0.04)	1.78(1.98)	-	-
48	0.22(0.04)	1.86(2.36)	-	-
49	0.38(0.04)	1.70(1.92)	-	-
50	0.28(0.04)	1.86(2.44)	-	-
51	0.32(0.04)	1.83(2.52)	-	-
52	-0.36(0.07)	0.91(0.94)	-0.41(0.07)	0.96(1.00)
53	-0.25(0.06)	0.92(0.86)	-0.25(0.07)	1.06(0.98)
54	-0.2(0.06)	0.91(0.85)	-0.20(0.06)	1.07(0.97)
55	0.01(0.05)	0.92(0.90)	0.07(0.06)	1.10(1.08)
56	-0.09(0.05)	0.95(0.95)	-0.04(0.06)	1.15(1.16)
57	0.28(0.11)	0.73(0.72)	0.43(0.12)	0.91(0.88)
58	0.27(0.10)	0.80(0.79)	0.41(0.11)	0.97(0.96)
59	0.37(0.10)	0.92(0.93)	0.52(0.11)	1.14(1.16)
60	0.44(0.04)	0.82(0.83)	0.62(0.05)	0.97(0.98)
61	0.19(0.05)	0.93(0.93)	0.31(0.06)	1.11(1.10)
62	-0.02(0.06)	0.98(0.96)	0.04(0.06)	1.16(1.13)
63	0.15(0.05)	1.01(1.01)	0.26(0.06)	1.20(1.20)
64	0.32(0.05)	1.12(1.13)	0.45(0.06)	1.39(1.42)
65	0.09(0.05)	0.92(0.92)	0.17(0.06)	1.08(1.06)

Residual analysis of the 57-item scale revealed that 43.7% of the variance was explained by the items together, which was a lot closer to the ideal value of 50%. The first contrast in residuals explained 5.0% of the variance (eigenvalue=5.1), which was marginally within the tolerant range of unidimensionality. The 57-item version of C CP QOL-Child parent proxy questionnaire was therefore considered meeting the unidimensionality nature of measuring QOL of children with CP.

3.5. Reliability and separation

In the original 65-item scale, the person and item reliability were 0.91 and 0.95, respectively. In the modified 57-item scale, the person reliability was 0.95, which was 0.04 higher than the original one; and its item reliability remained the same (0.95). Item difficulty estimates ranged from -0.59 to 0.49 (SD=0.28) in the original scale, and -0.72 to 0.62 (SD=0.32) in the modified scale. The index of person separation in the original scale and the modified scale were 3.17 and 4.30, respectively; the index of item separation in the original scale and the modified scale were 4.37 and 4.43, respectively. All these results indicated that the item categories in the modified scale were separated better than the original scale, and the respondent differences in the modified scale were distinguished better than the original scale (Table 6.4).

Table 6.4

Separation reliability and cumulative explained variance of C-CP QOL-Child Scale

	Person		Item		Cumulative Explained Variance (%)
	Reliability	Separation	Reliability	Separation	
Original 65-item scale	0.91	3.17	0.95	4.37	56.4
Modified 57-item scale	0.95	4.3	0.95	4.43	62.6

CHAPTER 4

DISCUSSION

The CP QOL-Child was the first instrument specifically developed to assess QOL of children with CP. Its Chinese version is useful in a different cultural setting, i.e., among Chinese speaking populations. Successful validation of the Chinese version would further justify that CP QOL-Child can in fact serve as an international-based measure. The reliability and validity of C CP QOL-Child has been assessed using CTT (Wang, et al., 2010). The current study further validated and refined the instrument using item response theory (IRT), more specifically the Rasch model.

4.1. CTT versus IRT (Reliability and Validity)

Results from the CTT analysis showed that the internal consistency (Cronbach's alpha) of the 7 dimensions were 0.82 (Swb), 0.90 (Fun), 0.91 (Part), 0.82 (Ewb), 0.86(Acc), 0.78(Pain) and 0.89 (Fam), respectively. Since a simple way to obtain a higher level of reliability is to

increase the number of items in the scale (Streiner & Norman, 2008), the Spearman-Brown prophecy formula was used to calculate the test length increment (%) from using the CTT to IRT approach. The results indicated that except Dimension 7 (Fam), which exhibits no test length increment, the test length increments for Dimension 1 to 6 were 98%, 259%, 88%, 192%, 698% and 6% respectively (Table 6.5). Thus the test length increment demonstrates a better measurement efficiency of the IRT over the CTT approach.

Table 6.5

Reliability comparisons with test length increment

	Dimension						
	1 (Swb)	2 (Fun)	3 (Part)	4 (Ewb)	5 (Acc)	6 (Pain)	7 (Fam)
item(n)	12	12	11	6	12	8	4
IRT							
M	0.56	0.55	0.32	0.57	0.51	0.14	0.29
SD	0.40	0.36	0.34	0.52	0.33	0.10	0.21
Reliability: 7-dimension	0.90	0.97	0.95	0.93	0.98	0.79	0.89
Test length increment (%)	98	259	88	192	698	6	0
CTT							
M	6.9	5.8	6.0	6.6	5.7	5.4	5.5
SD	0.7	1.2	1.3	1.0	1.3	1.5	1.5
α	0.82	0.90	0.91	0.82	0.86	0.78	0.89

Swb : social well-being and acceptance
Fun : functioning
Part : participation and physical health
Ewb : emotional well-being
Acc : access to services
Pain : pain and impact of disability
Fam : family health

When using the CTT approach (Wang, et al., 2010), results of the principal components analysis followed by varimax rotation showed that the 62-item scale (in the Wang et al study, three items in the domain access to services had been excluded from the principal component analysis as many participants did not provide a response) cumulatively accounted for 57.7% of the variance. In this study using the IRT approach, results of the residual analysis for the original 65-item scale showed that the cumulative variance explained was 56.4%, which was 1.3% lower than the CTT 62-item scale. However, the cumulative variance explained for the modified 57-item scale was 62.6%, which was 4.9% higher than the CTT 62-item scale, and 6.2% higher than the IRT 65-item scale (Table 6.4).

In terms of the length of the instrument, some items had similar levels of difficulty [i.e., item 15 (-0.07), 16 (-0.08), 17 (-0.07) and 18 (-0.05)]. On the one hand, this could imply the need to remove some items. On the other hand, an argument for retention of these items could be put forward, as different life domains (eg, social well-being and participation) could be expected to evoke similar levels of life well-being or happiness. This debate is worth further investigation.

As indicated in the variable map, a few more “difficult” items and a few “easy” items could be inserted into the scale in order to differentiate or distinguish better among children with

CP with highest and lowest QOL; thus producing a more robust assessment tool for evaluating QOL in such children. All in all, the 65 items can serve as an item bank for further studies.

4.2. Practical Interpretations and Implications

The Rasch item difficulty estimates revealed that the items demonstrated an overall item hierarchy, which is useful for interpreting self-perceived well-being in children with CP and has clinical implications. To be more specific, hierarchically (in order of progressive challenge to well-being on QOL), item 10 (“going out on trips with the family?”) and item 11 (“how they are accepted by their family?”) on the scale were associated with highest level (the least challenge) of well-being, and item 46 (“Does your child worry about who will take care of them in the future?”) was associated with lowest level (the most challenge) of well-being for this sample. In terms of the 7 domains, 10 of the 12 “Swb” items and 5 of the 6 “Ewb” items locate toward the high well-being end of the hierarchy. It is evident that the caregivers perceived their child to be most satisfied with these two life domains. The 11 “Part” items and the 12 “Acc” items span the item-difficulty hierarchy, while most of the “Fam” items and “Fun” items are moderate to low level of well-being. Moreover, all of the 8 “Pain” items group at the low well-being end of the hierarchy. It indicates that the “Pain” domain

was perceived as least satisfied by the caregivers' child. Our results are generally consistence with, but are more specific than, previous research findings (Arnaud et al., 2008; Swiggum, Hamilton, Gleeson, & Roddey, 2010; Vargus-Adams & Martin, 2011). Certainly, children with CP and their parents are concerned most often with pain, daily function, participation and community life. As well, the social acceptance, emotions, and family functioning are the discerned areas of life which are associated with QOL of children with CP (Arnaud, et al., 2008; Swiggum, et al., 2010). The item-difficulty hierarchy found in this study reflects the caregivers' perception of the life well-being situation and it also delineate the needs for help for their child with CP. Identification of these hierarchical challenges to well-being or difficulty in life domains is an specific opportunity of the on-going improvement in the QOL of children with CP through appropriate stepwise intervention programs.

The hierarchical structure of the Chinese CP-QOL Child provides the evidence for its clinical usefulness. Clinician and therapists can expect a pattern of performance by a child with CP that is based on the established order of item difficulty found in this study. For example, a child with an average level of self-perceived well-being would generally be more satisfied with "easy" items (i.e., items on the lower end shown in Figure 6.2) and least satisfied with more "difficult" items on the upper end of the scale. The more "difficult" items may need closer evaluation because these items may indicate the children's areas of particular

limitation or need for help. Identifying specific patterns of QOL in terms of domains can therefore potentially serve as a guide to the progression of treatments and to target specific area for intervention. Future studies are needed to examine and identity factors associated with these life domains. Barriers to well-being, particularly those affecting self-perceived feeling and manner must be further investigated.

4.3. Sample Size and Sampling Issues

According to Wright and Stone(1979)'s "Best Test Design", if item calibrations stable within ±0.5 logit, with 99% confidence, the minimum sample size ranges from 108 to 243 (Wright & Stone, 1979). On the other hand, for a ±1 logit interval, 0.38 SE, and 95% confidence, approximately 2 in the person-item ratio is recommended for obtaining useful, relatively stable and precise estimates in item calibrations of Rasch analysis (Linacre, 1994, 2002). In this dataset, the sample size was 145, and the person-item ratio of the caregiver proxy scale was 2.23; it seems that the sample size we have is adequate for item calibration purposes. We have not included in this study an analysis of the child self-report scale. Comparison between caregiver proxy scale and child self-report scale therefore cannot be conducted either.

The sample was recruited in Taiwan; and children in this sample included a wide range of motor disability severity and type of CP. However, given that the health care system in mainland China, Hong King, and Macau are different from that in Taiwan, and that the culture between these places are not entirely the same, the usefulness of the C CP-QOL Child in those places may need to be further investigated.

4.4. Conclusion

The application of Rasch analysis to the Chinese version of CP QOL-Child parent proxy questionnaire has provided further evidence that it is a reliable and valid measure of the quality of life for children with CP. The analysis showed that after dropping all the 8 items in the dimension “pain and impact of disability” in the originally 65-item scale, the newly 57-item scale adequately exhibits the nature of unidimensionality. From the hierarchical structure identified, clinician and therapists can expect a pattern of performance by a child with CP that is based on the established order of item difficulty; and can design tailor-made intervention programmes for individuals accordingly.

CHATPER 5

SUMMARY

Background: The Chinese Cerebral Palsy Quality of Life for Children (C CP QOL-Child) is the first instrument developed to measure quality of life of (QOL) children with cerebral palsy in Chinese speaking populations.

Objective: The aim of the study was to examine the psychometric properties of C CP QOLChild using Item Response Theory Models. We were particularly interested to know how intervention strategies could be designed for individuals based on the item scores.

Methods: 145 primary caregivers (mostly mothers; mean age: 39.2) of children with cerebral palsy aged 4e12 were invited to complete the 65-item C CP QOL-Child questionnaire. Data were analyzed using Rasch analysis.

Results: Item difficulty estimates were aligned with person ability values, indicating that the items in the scale generally demonstrated an appropriate depth and width for measuring QOL of persons in the target population. The results also showed that after dropping the 8 items in the dimension pain and impact of disability in the 65-item scale, the revised 57-item scale exhibits unidimensionality (separation index = 4.43, r = 0.95); hence the total score computed

from the 57 items adequately reflects the level of QOL of the child as perceived by the caregiver. We further found that the Rasch item difficulty estimates demonstrated an overall item hierarchy; hence therapists can expect a pattern of performance by a child with CP that is based on the established order of item difficulty.

Conclusions: The hierarchical structure identified in the study may be useful for designing tailor-made interventions with an aim of improving QOL.

Remarks: Although the revised 57-item scale exhibits the characteristics of unidimensionality, the subscales of access to services (Acc) and family health (Fam) may somehow be considered as external factors while comparing to the between-dimension correlations within the scale.

Note:

The Part II study "Item Hierarchy of the Chinese version of Cerebral Palsy Quality of Life for Children" has been accepted for publication in European Journal of Paediatric Neurology (EJPN) in 2 June 2012.

APPENDICES

Appendix A: PART III: Software development for Computerized Adaptive Testing (CAT)

Part III

Software development for Computerized Adaptive Testing (CAT)

CHAPTER 1

ITEM RESPONSE TEHORY – AN APPLICATION APPROACH

Item Response Theory (IRT) is widely used in standardized testing programs such as the Graduate Management Admission Test (GMAT), the Graduate Record Examination (GRE), the Test of English as a Foreign Language (TOEFL), and the Scholastic Assessment Test (SAT) (Barrada, Olea, Ponsoda, & Abad, 2010; Belov & Armstrong, 2009; Jang & Roussos, 2007; Wainer & Wang, 2000).

Item Response Theory applies a set of mathematical models to indicate the interaction between a person's ability or a composite of abilities and the characteristics of items in a test (McCloy & Gibby, 2011). In IRT models, person ability is used to estimate person's level on a certain latent trait or skill, which is measured by test items (McCloy & Gibby, 2011).

In IRT, as a person's ability increases so does the probability of answering an item correctly. The probability of a person answering an item correctly in a logistic IRT model (3-parameter) can be defined as

$$P_i(\theta_j) = c_i + \frac{1 - c_i}{1 + e^{-Da_i(\theta_j - b_i)}}$$

where e is the base of the natural logarithms and equals 2.71828…; i, indexes test item (i=1, 2, 3, …, n); j, indexes examinee (j=1, 2, 3, …, n); a_i is the item discrimination index for item i, that is proportional to the slope of the item response function at the point θ_j=b_i; b_i is the item difficulty index for item i, that is the point on the ability scale at which an examinee has (1+c)2 probability of answering item i correctly; c_i is the lower asymptote parameter of the item response function for item i, that represents the probability of examinees with very low ability correctly answering the item; θ_j represents the ability of examinee j; $P_i(\theta_j)$ is the probability of examinee j with ability level θ answering item i correctly; and D is a scaling factor that equals 1.702, by which the values of $P_i(\theta_j)$ for the logistic and the normal ogive models are comparable.

Because of the mathematical convenience of the logistic models and because of their similarity to the normal ogive models, the logistic models are the most used in practice. However, the normal ogive models are occasionally used in testing programs. It can be expressed by

$$P_i(\theta_j) = c_i + (1 - c_i) \int_{-\infty}^{a_i(\theta_j - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2 2} dz$$

where $\int_{-\infty}^{a_i(\theta_j - b_i)}$ is the integral or the area under the contiuous curve of $\frac{1}{\sqrt{2\pi}} e^{-z^2 2}$; π is a constant that equals 3.14…; z is a variable that ranges from $-\infty$ to $a_i(\theta_j - b_i)$; e is the base of the natural logarithms and equals 2.71828…; i, indexes test item (i=1, 2, 3, …, n); j, indexes examinee (j=1, 2, 3, …, n); a_i is the item discrimination index for item i, that is proportional to the slope of the item response function at the point θ_j=b_i; b_i is the item difficulty index for item i, that is the point on the ability scale at which an examinee has (1+c)2 probability of answering item i correctly; c_i is the lower asymptote parameter of the item response function for item i, that represents the probability of examinees with very low ability correctly answering the item; θ_j represents the ability of examinee j; and $P_i(\theta_j)$ is the probability of examinee j with ability level θ answering item i correctly.

IRT includes a group of assumptions about the data to which the models apply. One assumption is called the assumption of unidimensionality, which means that only one ability or one composite of multiple abilities is measured by a test. The second assumption of IRT models is called the assumption of local independence. This assumption means that an examinee's performance on one item will not affect his/her response to the other items in a

test (Chen, 2010; McCloy & Gibby, 2011; van der Linden, Ariel, & Veldkamp, 2006; van der Linden & Veldkamp, 2004; Ware Jr et al., 2003).

Another assumption of IRT models is the monotonicity assumption, which implies that the probability of a correct response to a test item increases as the measured underlying trait increases (McCloy & Gibby, 2011; van der Linden & Veldkamp, 2004). These assumptions are considered strong assumptions, because in practice it is very difficult to satisfy all these assumptions. For example, the local independence assumption will be violated for the i^{th} item whenever

$$P(u_1, u_2, u_3, \dots, u_i|\theta_j) \neq P(u_1|\theta_j)\, P(u_2|\theta_j) \dots P(u_i|\theta_j)$$

where u_i represents examinee's response to test item i (i.e., with 1 denoting a correct response and 0 an incorrect one) and *P* is the probability of an examinee response given an ability level θ_j.

This violation can happen, for example, when one item provides a hint for a subsequent item, or supplies additional information that may assist an examinee to be better assured the correct response (McCloy & Gibby, 2011). The estimation of item parameters or examinee ability

may be affected when the assumptions are violated. Researchers have indicated that the model-data fit may be poor if the assumptions of an IRT model cannot be satisfied (Barrada, et al., 2010; Barrada, Veldkamp, & Olea, 2009; Belov & Armstrong, 2009; Chen, 2010; van der Linden, et al., 2006; van der Linden & Veldkamp, 2004).

Several major aspects of IRT models were important for the study in item levels, such as information function, test characteristic curve, and methods for estimating item and ability parameters. Those issues will be discussed below.

1.1. Information Function

One of the fundamental notions of IRT is the item information function. It has various applications in the measurement field, such as in test construction, item selection, assessment of precision of measurement, and comparison of scoring methods. In a psychometric or statistical sense, information is the precision with which things are being estimated. Fisher (1922) defined statistical information as the reciprocal of the standard error with which a parameter can be estimated (Fisher, 1922). In IRT, information given by an item can be estimated at each point along the θ continuum. This represents the information that an item would provide for estimating examinee ability at each ability level (Barrada, et al., 2009;

McCloy & Gibby, 2011). Therefore, the greater the information there is at a given θ level, the more precise the measurement will be at that θ. The item information function can be expressed as

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

where $I_i(\theta)$ indexes the information provided by item i at a given θ; $P_i(\theta)$ is the probability of an examinee with an ability level of θ answering item i correctly on the θ continuum; $P_i'(\theta)$ is the first derivative of $P_i(\theta)$ with respect to given θ; $Q_i(\theta)$ equals $1 - P_i(\theta)$, that is the probability of an examinee answering item i incorrectly at a given θ; and i indexes test item (i=1, 2, 3,…, n).

The test information function shows the maximum amount of information that is available from a test composed of a particular set of items. The information provided by a test is simply the sum of item information functions at a given θ. The test information function, expressed as $I(\theta)$ is given by

$$I(\theta) = \sum_{i=1}^{n} I_i(\theta)$$

where i and $I_i(\theta)$ have the same meanings as in the above equations, and $I(\theta)$ represents the test information function.

The information function of a test can also be described as the following when the first derivative of $P_i(\theta)$ is solved

$$I(\theta) = D^2 \sum_{i=1}^{n} \frac{a_i^2 Q_i (P_i - c_i)^2}{P_i (1 - c_i)^2}$$

where all the notations were defined the same as above.

The standard error of estimation, stated as $SE(\hat{\theta})$, of a test is inversely related to the test information function along the θ continuum. The $SE(\hat{\theta})$ can be denoted as

$$SE(\hat{\theta}) = \frac{1}{\sqrt{I(\theta)}}$$

In the IRT framework, $SE(\hat{\theta})$ serves the same role as the standard error of measurement in classical test theory. Based on the $SE(\hat{\theta})$, a confidence interval for a given θ on the θ continuum can be established by

$$\hat{\theta} - z_{\alpha/2}SE(\hat{\theta}) \leq \theta \leq \hat{\theta} + z_{\alpha/2}SE(\hat{\theta})$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile point of the standard normal curve; $SE(\hat{\theta})$ represents the standard error of estimation at a given θ; and $\hat{\theta}$ is the estimated ability value.

1.2. Test Characteristic Curve

The test characteristic curve is another important concept in IRT that is related to the item response function. The test characteristic curve is the average of item response functions (i.e., the probability of correctly answering items in a test), and is expressed as

$$TCC = \frac{1}{n}\sum_{i=1}^{n} P_i(\theta)$$

where i and $P_i(\theta)$ have the same notations as in the above equations, and *TCC* represents the test characteristic curve along the θ continuum.

1.3. Parameter Estimation

In practice, both item and ability parameters are unknown, while only examinees' test responses (i.e., 0/1s) are known. Based on examinees' responses to test items, item parameters and ability parameters (a, b, and c) and ability parameters (θ) are estimated. Several procedures are currently available to estimate item and ability parameters in IRT, such as joint maximum likelihood, conditional maximum likelihood, marginal maximum likelihood, and joint marginal Bayesian estimation procedures. These methods all use maximum likelihood or Bayesian estimation procedures to estimate item parameters and/or an examinee's ability. Computer programs that accomplish these estimation methods have also been developed accordingly. WINSTEPS is one of the computer programs, in which the joint maximum likelihood estimation (JMLE) procedure is used to estimate parameters.

CHAPTER 2

COMPUTERIZED ADAPTIVE TESTING

Computerized adaptive testing (CAT) is beginning to be viewed as a practical alternative to traditional paper-and-pencil tests. Test specialists have investigated the use of CAT from

different perspectives for their testing purposes. CAT administers items based on examinees' responses to previously administered items, which may result in a more precise or accurate estimate of examinee's proficiency on the underlying scale. More specifically, the first item that will be administered to an examinee has about medium difficulty for the total population. Those who answer correctly will receive a more difficult item; and those who answer incorrectly will obtain an easier item. There are different ways of selecting items to be administered. One of the most frequently used methods of item selection is the maximum information approach proposed by Marco (1977). The primary advantage of selecting items using the maximum information method is that $\hat{\theta}$ values can be equally precise across the θ continuum. After each item response, the examinee's ability will be provisionally estimated. The process of item administration continues until there is enough information to give a final ability estimate with a specified level of accuracy, or until a previously specified number of items have been administered (Marco, 1977). Also, some alternative item selection procedures have also been studied in the past few years, which include the self-adaptive testing, the global information procedure, the maximum information procedure, and the Bayesian item selection. In self-adaptive testing, examinees decide the difficulties of items to be administered, whereas in the others, the administered item is based on the previous response of the examinees. The major difference between the global information and the maximum information procedure is that the global information selects the most informative

item with respect to a broad range around $\hat{\theta}$, and the maximum information bases the choice on the information provided at the point estimate of θ. The Bayesian item selection procedure selects the test items that minimize the variance of the posterior distribution of an examinee's ability (McCloy & Gibby, 2011; Van der Linden & Hambleton, 1997). While the item selection procedure is only one of the major components in CAT, in order to illustrate the whole picture of CAT application, we will discuss the details of each component below.

There are seven major components in a typical CAT administration (Chien, Wang, Huang, Lai, & Chow, 2011; Chien, Wang, Wang, & Lin, 2009; McCloy & Gibby, 2011; Pesudovs, 2010; van der Linden, et al., 2006; van der Linden & Veldkamp, 2004; Ware Jr, et al., 2003): 1) build an item pool. This is a set of items from which test items are selected. Many researchers indicated that a main element for a good CAT is a large and well-distributed item pool with well-calibrated item parameters; 2) choose an item response model. CAT has benefited from and depended on IRT as the underlying test model; 3) determine a test entry point. The difficulty of the initial item administered may be any value, but usually an item with medium difficulty will be given to an examinee first; 4) adopt an item selection procedure. This is the procedure for selecting the next item to be administered. An item selection algorithm should include certain restrictions in terms of how to select the next item, such as content balancing, exposure control, and item overlap constraint; 5) choose a method

of estimating the provisional ability. The provisional ability estimation can be achieved by using different ability estimation procedures; 6) determine a test stopping rule. The CAT can be terminated based on a previously specified criterion. Typically, this is based on a level of accuracy (variable length) or a fixed number of test items (fixed length); and 7) develop a method for computing the final ability. As the CAT terminates, a final ability estimate will be obtained.

Ability estimation and item selection algorithms are two essential components of a CAT administration. The ability estimation procedure is used to obtain a provisional ability estimate based on an examinee's test responses after each administered item, and a final ability estimate when the test terminates. The combinations of different ability estimation procedures can be used as the provisional and final ability estimates. Then, item selection algorithms provide efficient choice of items according to the provisional estimate of examinee's ability (Barrada, et al., 2010; McCloy & Gibby, 2011; van der Linden, et al., 2006; Van der Linden & Hambleton, 1997).

The decision of implementing a fixed or a variable length CAT depends on the testing purpose. Each type of CAT has certain advantages or disadvantages. Fixed length CAT, as the name implies, administers the same number of items to all the examinees. It is easy to

justify and explain to examinees and to the public, and easy to set a time allotment. A fixed length CAT can be inefficient, however, and may generate different levels of precision for different examinees (Barrada, et al., 2010; McCloy & Gibby, 2011; van der Linden & Veldkamp, 2004). On the other hand, a variable length CAT has advantages of providing a constant level of precision and giving an efficient test. The disadvantages of variable length CAT are that it can be difficult to justify or explain, and it is difficult to set a time allotment (Barrada, et al., 2010; Barrada, et al., 2009).

The popularity of CAT is related to the advantage of administering a test according to examinees' responses to previously administered test items. The advantages of CAT are documented in many of the literatures (Chien, et al., 2011; Chien et al., 2009; Chien, Wang, Wang, et al., 2009; Chien, Wu, Wang, Castillo, & Chou, 2009; Mitchell, et al., 2011; van der Linden, et al., 2006; Van der Linden & Hambleton, 1997; Ware Jr, et al., 2003). Research has shown that an equally reliable score can be obtained in CAT with approximately half the items required in paper-and-pencil tests (McCloy & Gibby, 2011). A second advantage of CAT is the fact that CAT can save on costs related to shipping and printing test materials (i.e., because CAT stores all the test items in a computer system). Another advantage is that examinees are usually challenged but not discouraged due to the nature of CAT (i.e., tailored test for each individual). The other advantages of CAT include frequent and convenient test

scheduling, immediate scoring, online/computer-collection of data, and presenting items in a multimediated way (i.e., with animation, sound track, etc.) (Chien, et al., 2011; Chien, Wu, et al., 2009; McCloy & Gibby, 2011; Ware Jr, et al., 2003).

2.1. Entry point item

Perhaps the most common technique in CATs is to get started with an item which is usually of moderate difficulty. Generally if the examinee gets it right, the difficulty level will be automatically increased. Many test developers have evolved useful rules regarding the ordering of items within a particular test. In fact, tests are often designed to start out easy and end up hard. Such tests are designed in this way for several reasons: first, persons with lower proficiency are generally encouraged by initial success therefore will likely to work harder at the more difficult items; second, the algorithm of selections for CATs oftentimes violate the ordering rules, hence the initial item within a CAT is typically one in middle difficulty; third, if a person gets these items correct, a quite more difficult items will probably be chosen. Certainly, we may choose a lower starting point instead, but a lower starting point will minimize the ordering effect, and also will limit the efficiency gains from the adaptive testing.

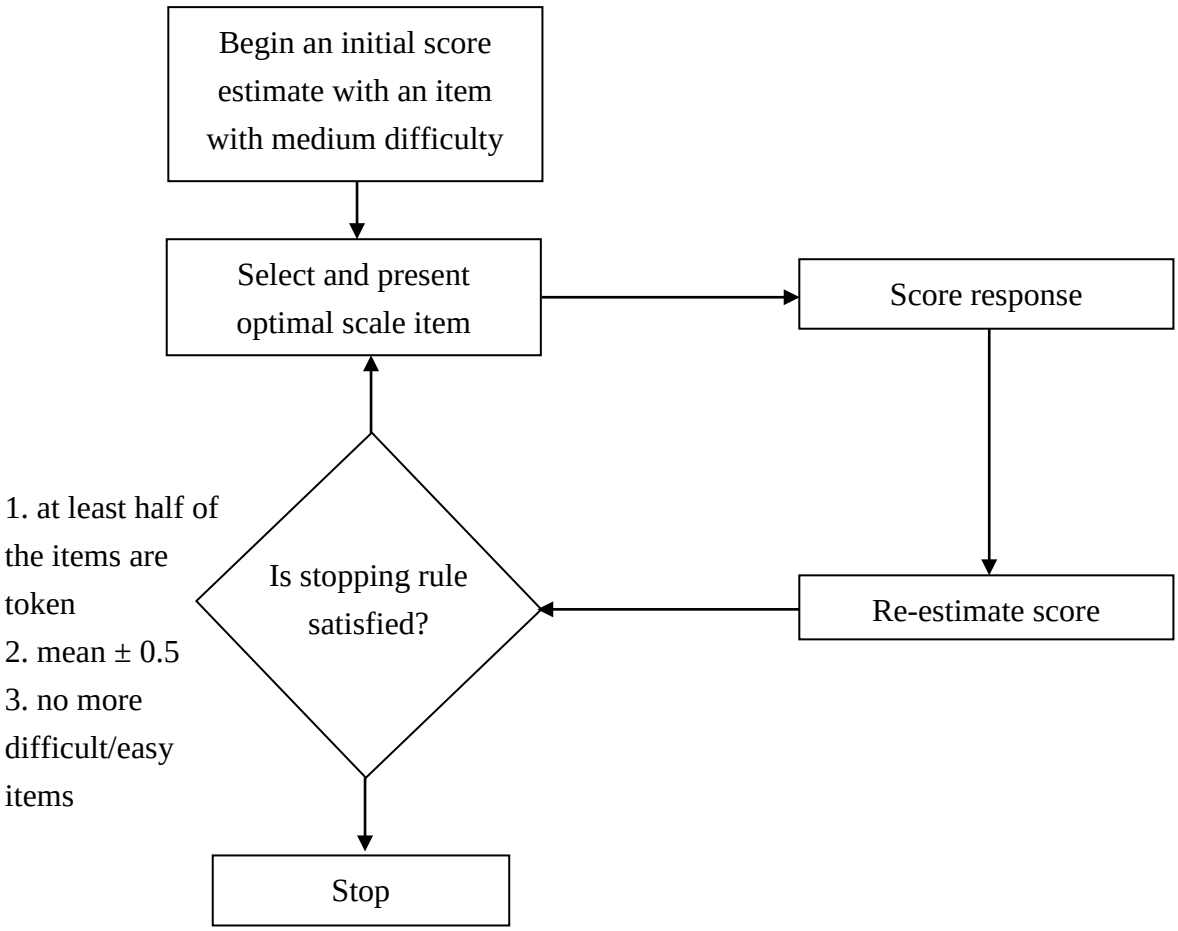
2.2. Limitation of middle entry point

For some examinees whose quality of life/happiness levels are below or above middle will waste time responding to questions that are not appropriate for their current level. In addition, it will increase the number of items required to achieve precise measurement.

2.3. Item bank preparation

Item pool was prepared by the results of the full scale Non-Adaptive Testing (NAT) study. Once the item bank is ready, it can be loaded into the CAT program.

2.4. Logic of CAT



2.5. CAT programing code

A complier, Visual Basic 6 with Service Pack 6 (Microsoft Corporation, USA), was utilized to transform source code into machine code and create an executable program for CAT application. The designed program code with annotation was shown below (Table 7.1).

Table 7.1

Program code overview

Script	Action/Description

`Dim a(9) As Integer, b(9) As Double,`
`c(9) As Boolean, t(9) As String`
`Dim x, y, z As Double`
`Dim m, n, o, p As Integer`
	Define the variable types

`Private Sub Form_Activate()`
 	
`Command4.Enabled = False`
	While the program window was activated, run the script below

`Option1.Caption = "Strongly Agree"`
`Option2.Caption = "Agree"`
`Option3.Caption = "Neutral"`
`Option4.Caption = "Disagree"`
`Option5.Caption = "Strongly`
`Disagree"`
 	
`Command1.Caption = "Next Item"`
`Command4.Caption = "Summary"`
`Command2.Caption = "Reset"`
`Command3.Caption = "Exit"`
	Define the caption of all the buttons

`n = 9` `m = 0`	Total number of items within the item pool

`For i = 1 To n`
`  a(i) = 0`
`  b(i) = 0`
`  c(i) = False`
`Next i`
	Return to zero setting

`t(1) = "1. i am extremely easy."`
`t(2) = "2. i am very easy."`
`t(3) = "3. i am easy."`
`t(…) = "……"`
	Load the item content into the physical memory in arrays

Table 7.1 *(continued)*

Script	Action/Description

`z = 0`
`p = 5`
`Print t(p)`
`c(p) = True`
 	
`End Sub`
	Determine the starting point and print out the first item

`Private Sub Command2_Click()`
`Cls`
`Command1.Enabled = True`
`Call Form_Activate`
`End Sub`
	"Reset" button
	Clear the screen
	"Next Item" button enable

`Private Sub Exit_Click()`
`End`
`End Sub`
	"Exit" option under the menu bar
	End up the program

`Private Sub Command1_Click()`
 	
`If p > 9 Then`
`  Print "No more difficult items"`
`  Command4.Enabled = True`
`  Command1.Enabled = False`
`  Exit Sub`
`ElseIf p < 1 Then`
`  Print "No more easy items"`
`  Command4.Enabled = True`
`  Command1.Enabled = False`
`  Exit Sub`
`End If`
	While the "Next Item" button was on click, run the script below
	
	Test if the selected item exceeds the boundary of the current item pool or not

Table 7.1 *(continued)*

Script	Action/Description

`If Option1.Value = True Then x = 2`
`If Option2.Value = True Then x = 1`
`If Option3.Value = True Then x = 0`
`If Option4.Value = True Then x = -1`
`If Option5.Value = True Then x = -2`
	Detect responses and store the value into the physical memory

`m = m + 1`
`z = z + x`
`y = z / m`
`a(m) = p`
`b(m) = x`
	Accumulate the respondent scores and calculate the initial mean

`If m > 4 And (y > -0.5 And y < 0.5) Then`
`  Command4.Enabled = True`
`  Command1.Enabled = False`
`  Exit Sub`
`End If`
	Test if the selected item number exceeds half of the total item number or not, then test the initial mean is within the range of ±0.5 or not

`Select Case x`
`  Case 2`
`  p = p + 1`
`  While c(p) = True`
`  p = p + 1`
`  Wend`
`  If p < 10 Then`
`  Print t(p)`
`  c(p) = True`
`  End If`
	Analyze the initial response, estimate the examinee's ability, and then choose an appropriate following item for the examinee by utilizing the case-by-case condition function

Table 7.1 *(continued)*

Script	Action/Description

Case 1
p = p + 1
While c(p) = True
p = p + 1
Wend
If p < 10 Then
Print t(p)
c(p) = True
End If	

Case 0
Randomize
*o = Int(Rnd * 2) + 1*
If o = 1 Then
p = p + 1
While c(p) = True
p = p + 1
Wend
End If
If o = 2 Then
p = p - 1
While c(p) = True
p = p - 1
Wend
End If
If p < 10 And p > 0 Then
Print t(p)
c(p) = True
End If
	If a zero score (neutral option) received, determine either level go up or go down by a randomize function

Table 7.1 *(continued)*

Script	Action/Description

``
    Case -1
        p = p - 1
        While c(p) = True
            p = p - 1
        Wend
        If p > 0 Then
            Print t(p)
            c(p) = True
        End If
 	
    Case -2
        p = p - 1
        While c(p) = True
            p = p - 1
        Wend
        If p > 0 Then
            Print t(p)
            c(p) = True
        End If
End Select
 	
If m = 9 Then
  Command4.Enabled = True
  Command1.Enabled = False
End If
 	
End Sub
``
	If all the items in the item pool were chosen, change the status of button "Next Item" to disable, and enable the button "Summary"

Table 7.1 *(continued)*

Script	Action/Description
`Private Sub Command4_Click()`	"Summary" button
`Print`	
`Print "mean = "; y`	Show the mean score of the initial person
`Print "number of item responsed = "; m`	Show the total number of items chosen
`Print`	
`Print "Item Number", "Response"`	
`For j = 1 To 9`	Print the field heading "Item Number" and "Response"
`If a(j) > 9 Or a(j) < 1 Then`	
`a(j) = 0`	Test the array and screen out those items did not choose
`b(j) = 0`	
`End If`	
`If a(j) <> 0 Then Print a(j), b(j)`	
`Next j`	Print out the initial results
`Command4.Enabled = False`	
`End Sub`	Change the status of button "Summary" to disable
`Private Sub Save_Click()`	"Save " option under the menu bar
`Open "save.csv" For Output As #1`	Create a file naming "save.csv"
`Write #1, "Item", "Response"`	Write the field heading "Item" and "Response"
`For k = 1 To 9`	
`If a(k) <> 0 Then Write #1, a(k), b(k)`	Write the scoring record into the file
`Next k`	
`Close #1`	
`End Sub`	

2.6. User manual of CAT program

Step 1: Prepare an item pool

Description of the sample item set

- 5-point Likert scale, from “Strongly Disagree” to “Strongly Agree”

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Score	-2	-1	0	1	2
	Level down			Level up	

Step 2: Prepare an item difficulty list

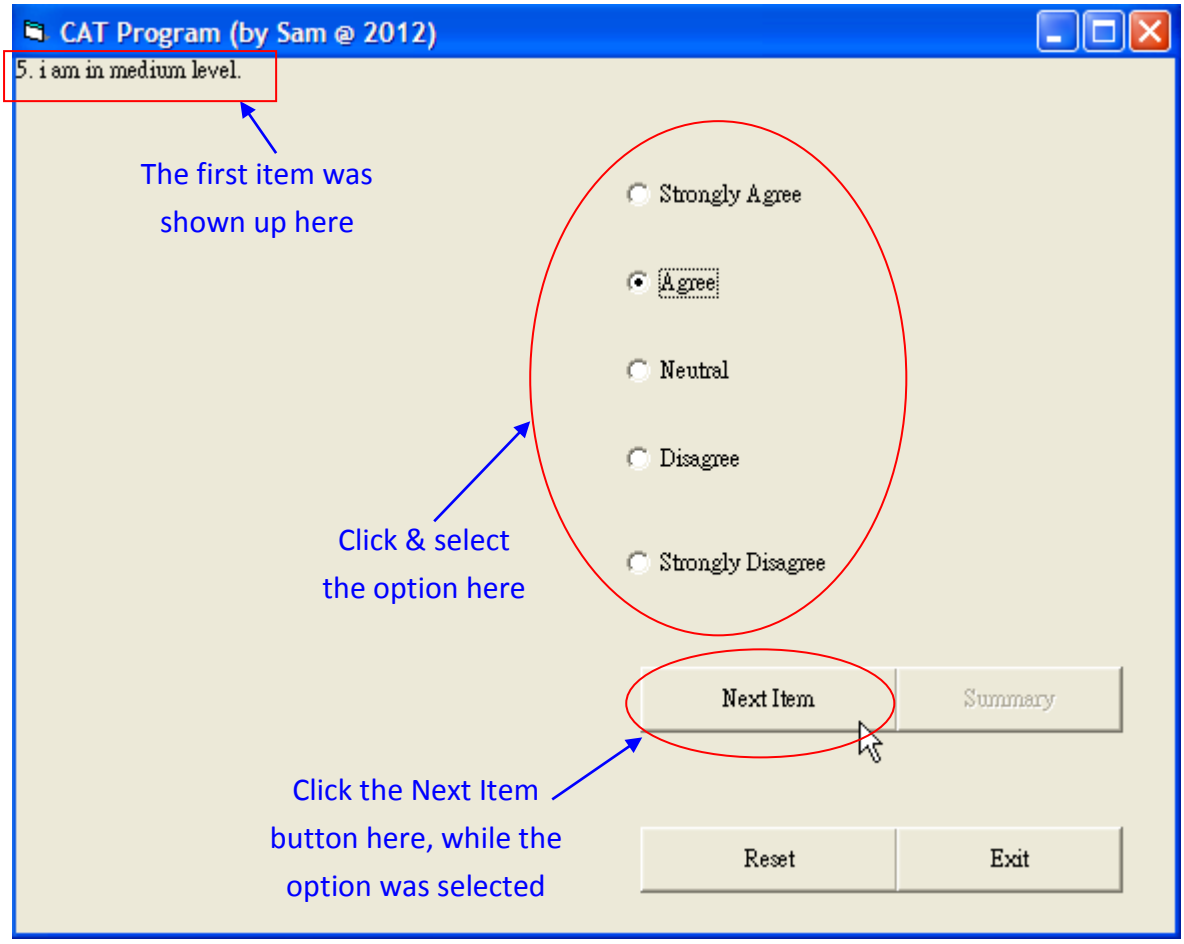
- person-item map

x	2	9	i am extremely difficult.
xxx			
x xxx		8	i am very difficult.
xxxxxxx			
xxxxxxx xxxx	1	7	i am difficult.
xxxxxxxxxxxxxxxxx			
xxxxxxxxxxx xxxxxx xxxx		6	i am little bit difficult.
xxxxxxxxxxxxxxxxxxxxxxxx			
xxxxxxxxx xxxxxx xxxxxx xxxxxx	0	5	i am in medium level.
xxxxxxxxxxxxxxxxxxxxxxx			
xxxxxxxxxxx xxxxxx xxx		4	i am little bit easy.
xxxxxxxxxxxxxxxxx			
xxxxxxxxxxxx	-1	3	i am easy.
xxxxxxxx			
xxxx		2	i am very easy.
xxx			
x	-2	1	i am extremely easy.

Person – Item Map

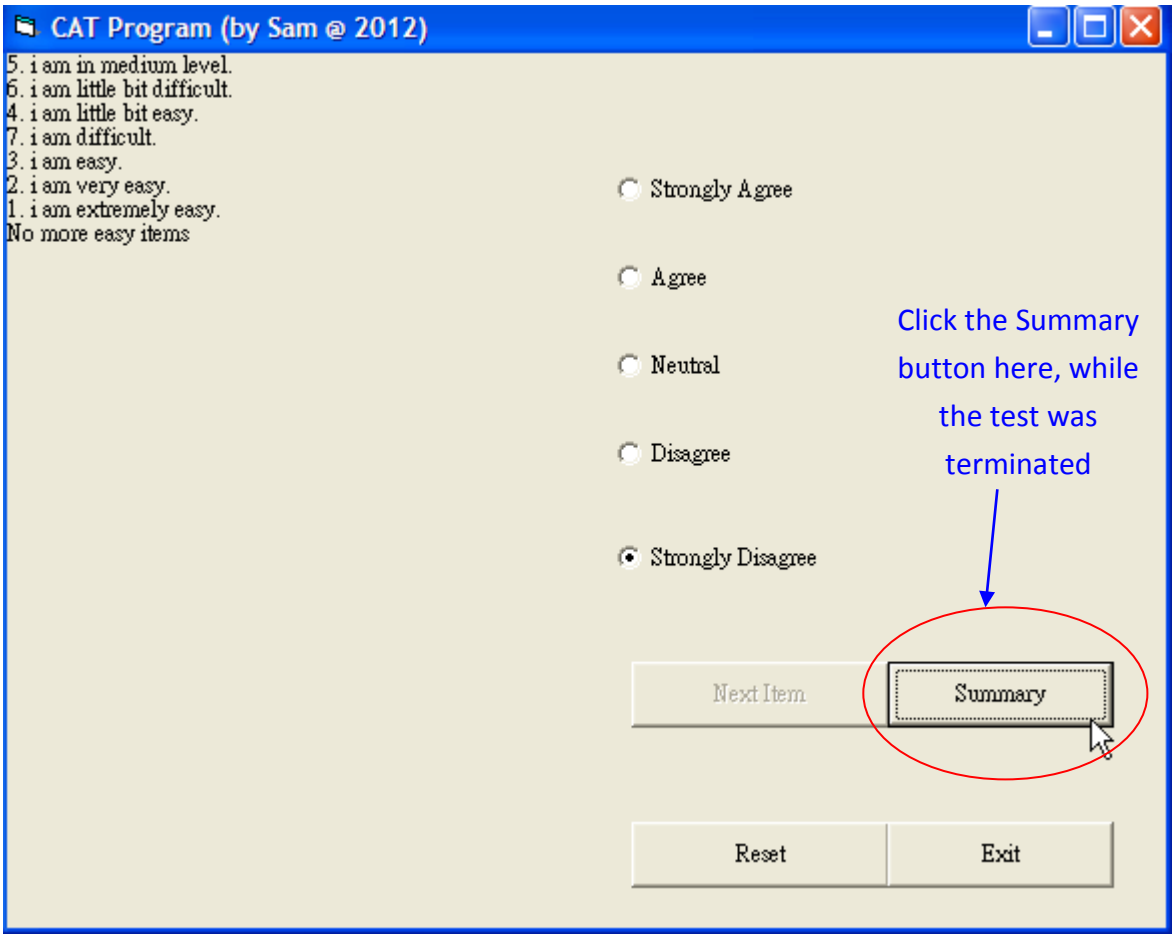
Step 3: Choose an appropriate stopping criterion

- item > 4
- mean ± 0.5
- no more difficult item
- no more easy item



When the program was activated, a window was displayed as above

- item(s) will be shown up at the left corner of the window
- response to the item by selecting the category in the middle of the window, then click the "Next Item" button for generating another item based on the respondent score



When the stopping criteria were met, the button “Next Item” will be disabled, and at the same time, the button “Summary” will be enabled

- click the button “Summary” to display the statistic history

CHAPTER 3

DEVELOPMENT OF AN ON-SITE CP QOL-CHILD CAT PROGRAM

The basic framework of the program was interpreted in the previous section, from now on we attempted to establish a CAT system for the CP QOL-Child scale.

3.1. Item pool

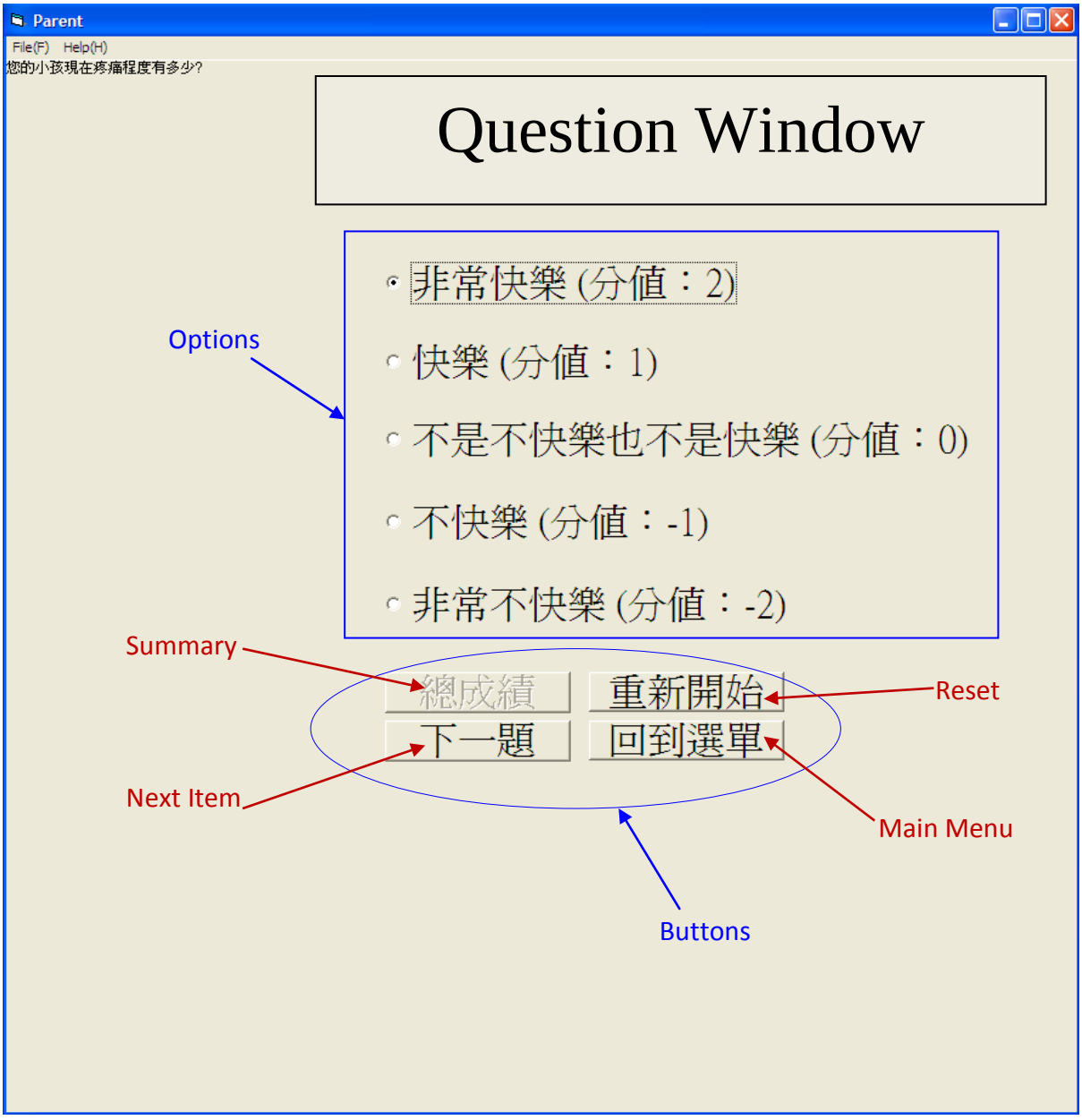
For implementation purposes, the first thing we need to do is to build an item pool for the item selection procedure. According to the results of Rasch Analysis (Part I) before, we can see the list of relative difficulties of each item within the scale, and this difficulty list can be implanted into the program after an appropriate arrangement (sort in descending order), this can perform as the test item pool.

Content	Item	Content	Item	Score	(SE)
the way you get along with people, generally?	1	你平常與人相處的方式?	1	-0.08	0.07
the way you get along with the person who looks after you?	2	你與照顧你的人相處的方式?	2	-0.32	0.07
the way you get along with your brothers and sisters?*	3	你與家裡的兄弟姊妹相處的方式?*	3	-0.14	0.06
The way you get along with other children at school?	4	你在學校與其他學童相處的方式?	4	-0.14	0.06
the way you get along with other children outside of school?	5	你在學校以外與其他小朋友相處的方式?	5	0	0.06
the way you get along with adults?	6	你與大人相處的方式?	6	0	0.07
the way you get along with your teachers and/or careers?	7	你與學校老師或照顧者相處的方式?	7	-0.34	0.07
your ability to play on your own?	8	你一個人玩的能力?	8	0	0.06
your ability to play with friends?	9	你與朋友一起玩的能力?	9	-0.26	0.05
going out on trips with your family?	10	你與家人外出旅遊的情形?	10	-0.59	0.07
how you are accepted by your family?	11	你被家人接納的情形?	11	-0.42	0.07
how you are accepted by other children at school?	12	你被學校學童接納的情形?	12	-0.18	0.06
how you are accepted by other children outside of school?	13	你被學校以外的小朋友接納的情形?	13	-0.06	0.06
how you are accepted by adults?	14	你被大人接納的情形?	14	-0.17	0.07
how you are accepted by people in general?	15	你被一般大眾接納的情形?	15	-0.04	0.06
being able to do the things you want to do?	16	你可以做到你喜歡做的事?	16	-0.14	0.05

For preparing the item pool, please refer to Figure A1, Table 4.1a, and Figure B1, Table 4.2a.

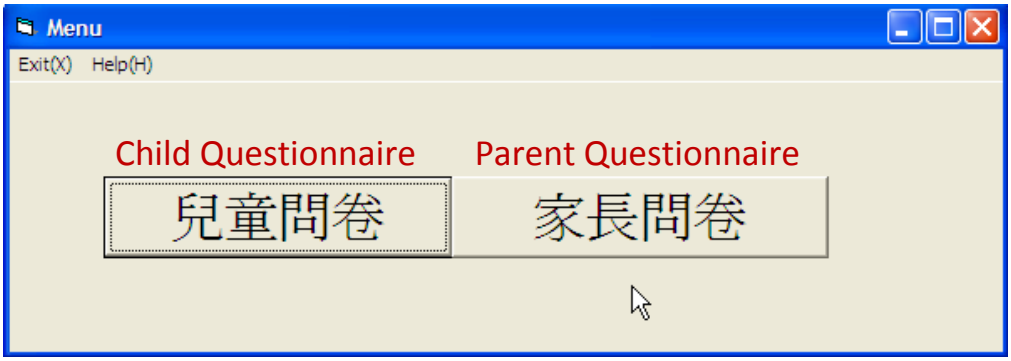
3.2. Entry point

In this program, we adopted a middle entry point technique, the first item shown was an item of intermediate difficulty, but those settings can be changed with modifying the program code.



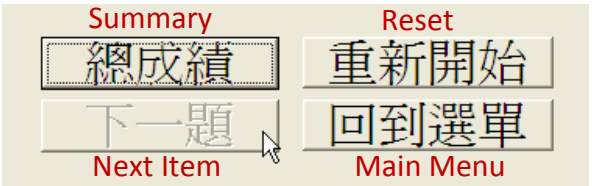
3.3. Main menu

After entering the main menu of the program, we can see the Child Questionnaire and the Parent Questionnaire buttons, users can click on the Child Questionnaire for entering into the child self report scale, or click on the Parent Questionnaire for entering into the parent proxy report scale.

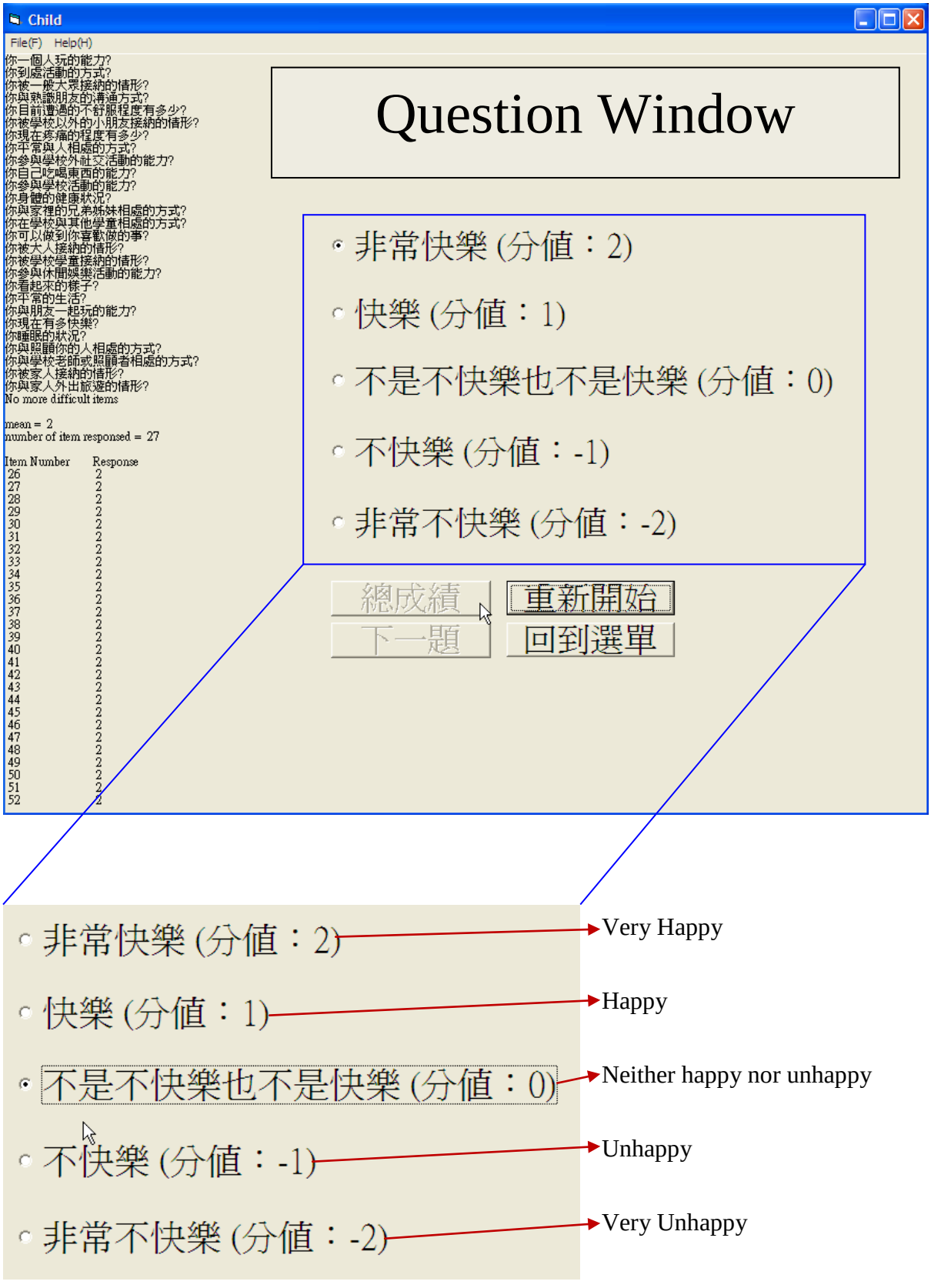


3.4. Stopping rule

In this program, in order to achieve a reasonable degree of reliability, the examinees are required to response at least half of the total number of the items, and when this requirement was fulfilled, the new stopping rule will be the mean scores within the range of ± 0.5 or no more higher or lower level of difficulty item for them to response.

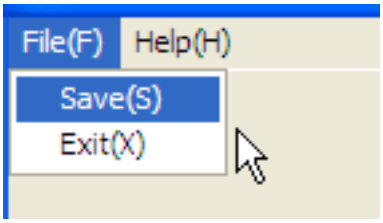


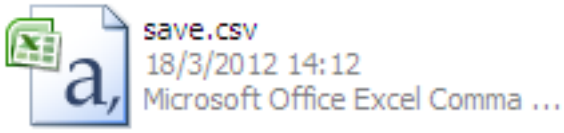
When the stopping criteria were reached, the button Next Item will then be disabled.

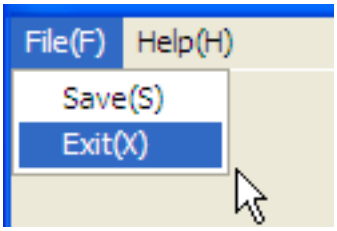


3.5. Save the results

After clicking on the drop-down menu File(F), we can see the Save(S) and Exit(X) options. Click on the Save(S) option (or use the hotkey Alt + S), all the current results will be saved into a file named “save.csv”; Click on the Exit(X) option (or use the hotkey Alt + X), the program will be terminated immediately.







Exit: close the window and quit the program.

This CAT program can be run on some currently popular mobile devices, such as slim notebooks and tablets, for some extends, there may require some code switching or transferring for the issue of compatibilities.

Supplementary Tables and Figures

Child self-report

Table A1

CP-QOL Child: Child Scale (Original English Version)

Item	Content
1	the way you get along with people, generally?
2	the way you get along with the person who looks after you?
3	the way you get along with your brothers and sisters?*
4	The way you get along with other children at school?
5	the way you get along with other children outside of school?
6	the way you get along with adults?
7	the way you get along with your teachers and/or careers?
8	your ability to play on your own?
9	your ability to play with friends?
10	going out on trips with your family?
11	how you are accepted by your family?
12	how you are accepted by other children at school?
13	how you are accepted by other children outside of school?
14	how you are accepted by adults?
15	how you are accepted by people in general?
16	being able to do the things you want to do?
17	your ability to participate at school?
18	your ability to participate in recreational activities?
19	your ability to participate in sporting activities?
20	your ability to participate in social events outside of school?
21	your ability to participate in your community?
22	the way you communicate with people you know well?
23	the way you communicate with people you don't know well?
24	the way other people communicate with you?
25	your physical health?
26	the way you get around?
27	how you sleep?
28	the way you look?
29	your ability to keep up academically with your peers?
30	your ability to keep up physically with your peers?
31	your life in general?
32	yourself?

Table A1 *(Continued)*

Item	Content
33	your future?
34	your opportunities in life?
35	the way you use your arms?
36	the way you use your legs?
37	the way you use your hands?
38	your ability to dress yourself?
39	your ability to eat or drink independently?
40	your ability to use the toilet by yourself?
41	the special equipment you have at home?*
42	the special equipment you have at your school?*
43	the special equipment that is available in the community?*
44	are you bothered by hospital visits?
45	are you bothered when you miss school for health reasons?
46	are you bothered by being handled by other people?
47	do you worry about who will take care of you in the future?
48	are you concerned about having cerebral palsy?
49	how much pain do you have?
50	how do you feel about the amount of pain you have?
51	how much discomfort do you experience?
52	how happy are you?

Note: * may contain missing data.

Table A2

C CP-QOL Child: Child Scale (Chinese Translated Version)

Item	Content
1	你平常與人相處的方式?
2	你與照顧你的人相處的方式?
3	你與家裡的兄弟姊妹相處的方式?*
4	你在學校與其他學童相處的方式?
5	你在學校以外與其他小朋友相處的方式?
6	你與大人相處的方式?
7	你與學校老師或照顧者相處的方式?
8	你一個人玩的能力?
9	你與朋友一起玩的能力?
10	你與家人外出旅遊的情形?
11	你被家人接納的情形?
12	你被學校學童接納的情形?
13	你被學校以外的小朋友接納的情形?
14	你被大人接納的情形?
15	你被一般大眾接納的情形?
16	你可以做到你喜歡做的事?
17	你參與學校活動的能力?
18	你參與休閒娛樂活動的能力?
19	你參與運動的能力?
20	你參與學校外社交活動的能力?
21	你參與社區活動的能力?
22	你與熟識朋友的溝通方式?
23	你與不熟識的人的溝通方式?
24	其他人與你溝通的方式?
25	你身體的健康狀況?
26	你到處活動的方式?
27	你睡眠的狀況?
28	你看起來的樣子?
29	你在學業上能跟上同學的能力?
30	你在體能上能跟上同學的能力?
31	你平常的生活?
32	你自己?
33	你的未來
34	你在生活中的機會?
35	你使用手臂的方式?

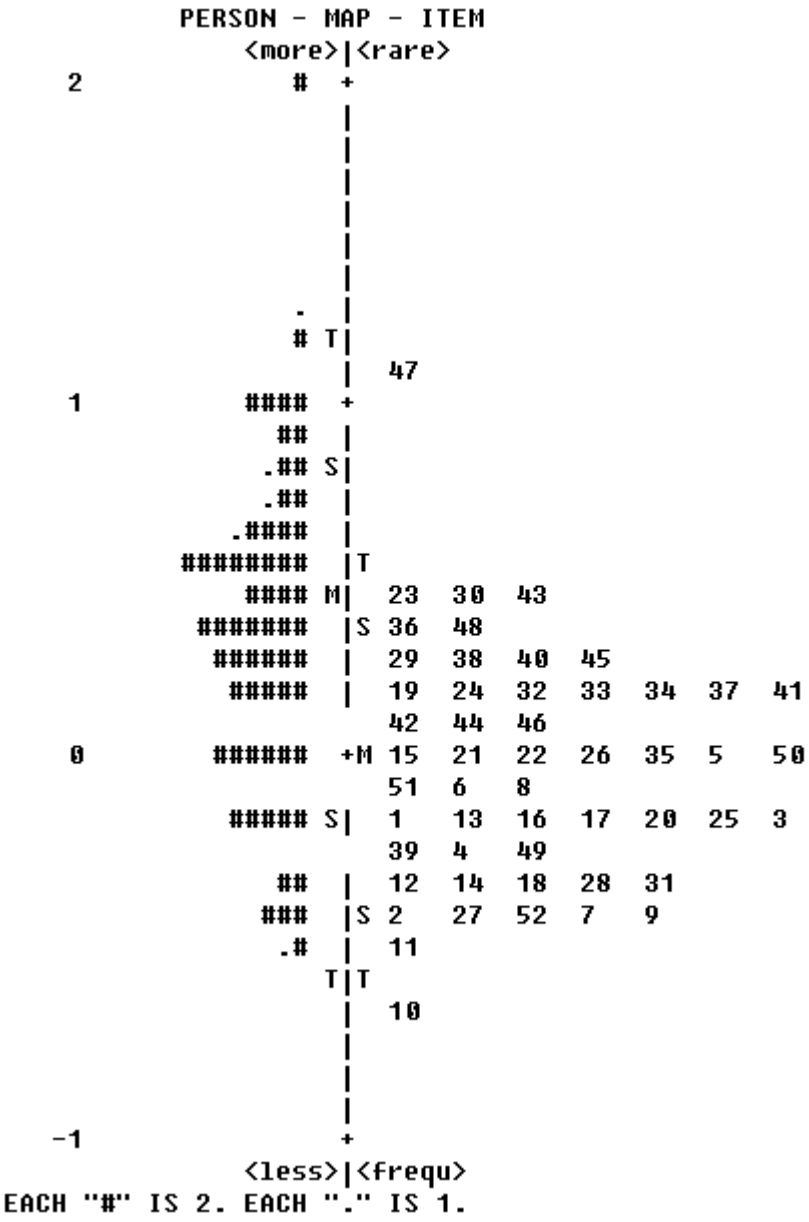
Table A2 *(Continued)*

Item	Content
36	你使用腳的方式?
37	你使用手的方式?
38	你自己穿衣服的能力?
39	你自己吃喝東西的能力?
40	你自己上廁所與盥洗的能力?
41	你使用家裡的特殊設備?*
42	你使用學校的特殊設備?*
43	你使用社區中的特殊設備?*
44	去醫院就醫時是否感到困擾?
45	因健康因素無法上學時是否感到困擾?
46	當你被別人扶持移動身體時是否感到困擾?
47	你是否會擔心未來是誰來照顧你?
48	你在意自己患有腦性麻痺嗎?
49	你現在疼痛的程度有多少?
50	你對於自己目前所承受的疼痛程度感受如何？
51	你目前遭遇的不舒服程度有多少？
52	你現在有多快樂？

Note: * may contain missing data.

Figure A1

C CP-QOL Child: Child Scale Variable Map



Parent proxy report

Table B1

CP-QOL Child: Parent Scale (Original English Version)

Item	Content
1	the way they get along with people, generally?
2	the way they get along with you?
3	the way they get along with their brothers and sisters?*
4	the way they get along with other children at preschool or school?*
5	the way they get along with other children outside of preschool or school?
6	the way they get along with adults?
7	the way they get along with their teachers and/or careers?
8	their ability to play on their own?
9	their ability to play with friends?
10	going out on trips with the family?
11	how they are accepted by their family?
12	how they are accepted by other children at preschool or school?*
13	how they are accepted by other children outside of preschool or school?
14	how they are accepted by adults?
15	how they are accepted by people in general?
16	being able to do the things they want to do?
17	their ability to participate at preschool or school?*
18	their ability to participate in recreational activities?
19	their ability to participate in sporting activities?
20	their ability to participate in social events outside of preschool or school?
21	their ability to participate in their community?
22	the way they communicate with people they know well?
23	the way they communicate with people they don't know well?
24	the way other people communicate with them?
25	their physical health?
26	the way they get around?
27	how they sleep?
28	the way they look?
29	their ability to keep up academically with their peers?
30	their ability to keep up physically with their peers?
31	their life in general?
32	themselves?
33	their future?

Table B1 *(Continued)*

Item	Content
34	their opportunities in life?
35	the way they use their arms?
36	the way they use their legs?
37	the way they use their hands?
38	their ability to dress themselves?
39	their ability to eat or drink independently?
40	their ability to use the toilet by themselves?
41	the special equipment they have at home?*
42	the special equipment they have at their school?*
43	the special equipment that is available in the community?*
44	is your child bothered by hospital visits?
45	is your child bothered when they miss school for health reasons?
46	is your child bothered by being handled by other people?
47	does your child worry about who will take care of them in the future?
48	is your child concerned about having cerebral palsy?
49	how much pain does your child have?
50	how does your child feel about the amount of pain they have?
51	how much discomfort does your child experience?
52	how happy is your child?
53	your child's access to treatment?
54	your child's access to therapy?
55	your child's access to specialised medical or surgical care?
56	your ability to get advice from a paediatrician?
57	your access to respite care?*
58	the amount of respite care you receive?
59	how easy it is to get respite?
60	your child's access to community services and facilities?
61	your child's access to extra help with learning at preschool or school?
62	your physical health?
63	your work situation?
64	your family's financial situation?
65	how happy are you?

Note: * may contain missing data.

Table B2

C CP-QOL Child: Parent Scale (Chinese Translated Version)

Item	Content
1	他平常與人相處的方式？
2	他與您相處的方式？
3	他與家中兄弟姊妹相處的方式？*
4	他在幼托園或學校與其他學童相處方式?*
5	他在幼托園或學校以外與其他小孩相處的方式？
6	他與大人相處的方式？
7	他與學校老師或照顧者相處的方式？
8	他自己一個人玩的能力？
9	他與朋友一起玩的能力？
10	他與家人外出旅遊的情形？
11	他被家人接納的情形？
12	在幼托園或學校被學童接納的情形?*
13	他被幼托園或學校以外其他小朋友接納的情形？
14	他被大人接納的情形？
15	他被一般社會大眾接納的情形？
16	他可以去做他喜歡做的事？
17	他參與幼托園或學校活動的能力?*
18	他參與休閒娛樂活動的能力？
19	他參與運動的能力？
20	他參與幼托園或學校外社交活動的能力？
21	他參與社區活動的能力？
22	他與熟識朋友的溝通方式?
23	他與不熟識的人的溝通方式?
24	其他人與他溝通的方式？
25	他身體的健康狀況？
26	他到處活動的方式？
27	他睡眠的狀況？
28	他看起來的樣子？
29	他在學業上能跟上同學的能力?
30	他在體能上能跟上同學的能力?
31	他平常的生活？
32	他自己？
33	他的未來？
34	他在生活中的機會？
35	他使用手臂的方式？

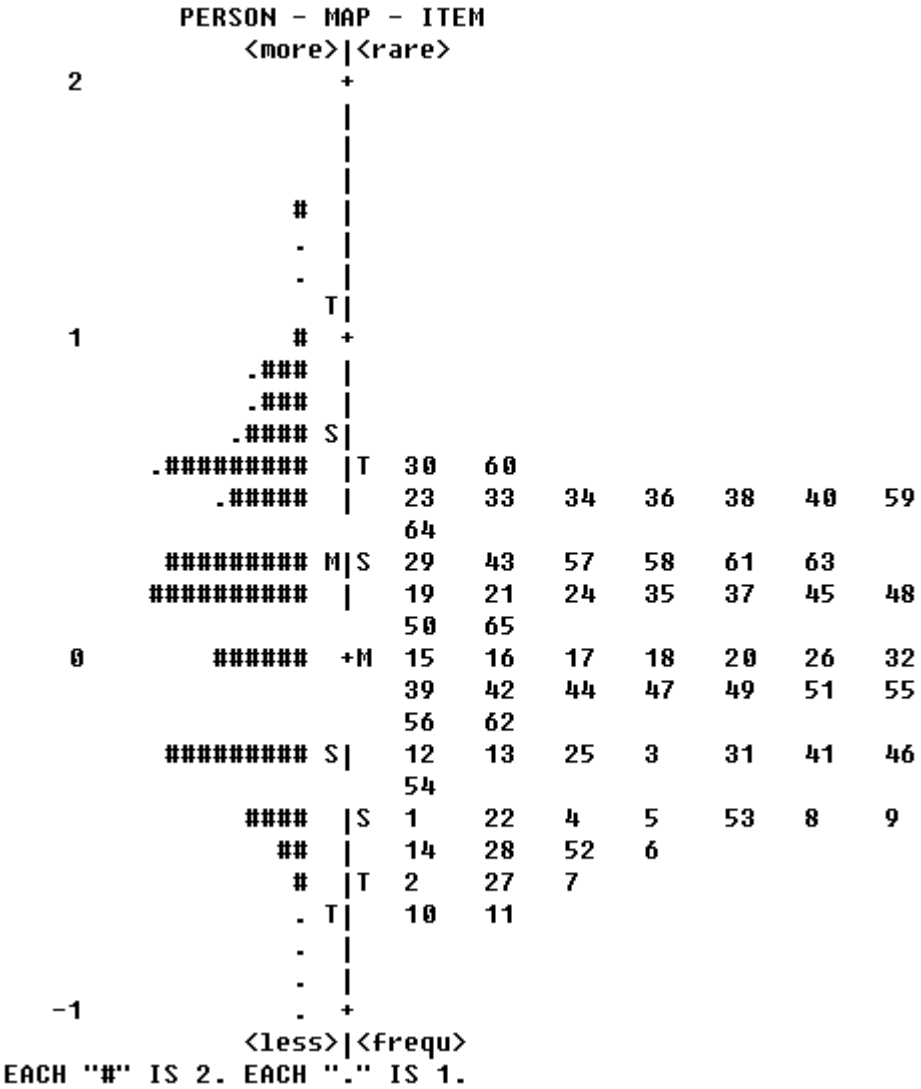
Table B2 *(Continued)*

Item	Content
36	他使用腳的方式？
37	他使用手的方式？
38	他自己穿衣服的能力？
39	他自己吃喝東西的能力？
40	他自己上廁所與盥洗的能力？
41	他使用家裡的特殊設備?*
42	他使用學校的特殊設備?*
43	他使用社區中的特殊設備?*
44	您的小孩去醫院就醫時是否感到困擾？
45	您的小孩因健康因素無法上學時是否感到困擾？
46	您的小孩被別人扶持移動身體時是否感到困擾？
47	您的小孩是否會擔心未來是誰來照顧他們？
48	您的小孩在意他患有腦性麻痺嗎？
49	您的小孩現在疼痛程度有多少？
50	您的小孩對於目前所承受的疼痛程度感受如何？
51	您的孩子目前遭遇的不舒服程度有多少？
52	您的孩子現在有多快樂？
53	您的小孩取得治療的方便性？
54	您的小孩取得復健治療的方便性?
55	您的小孩取得專科醫療或手術處理的方便性？
56	您可從小兒科醫師獲得建議訊息的方便性？
57	您取得喘息照顧服務的方便性?*
58	您接受喘息服務的量？
59	取得喘息服務的容易程度？
60	您的孩子獲得社區服務與設施的方便性?
61	您的孩子在幼托園或學校學習上獲得額外幫助的方便性?
62	您的身體健康？
63	您的工作狀況？
64	您家庭的經濟狀況？
65	您有多快樂？

Note: * may contain missing data.

Figure B1

C CP-QOL Child: Parent Scale Variable Map



Child self report item characteristic curve

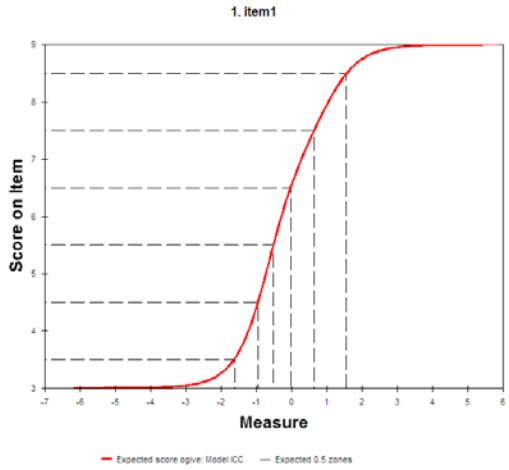


Figure C1: ICC of item 1

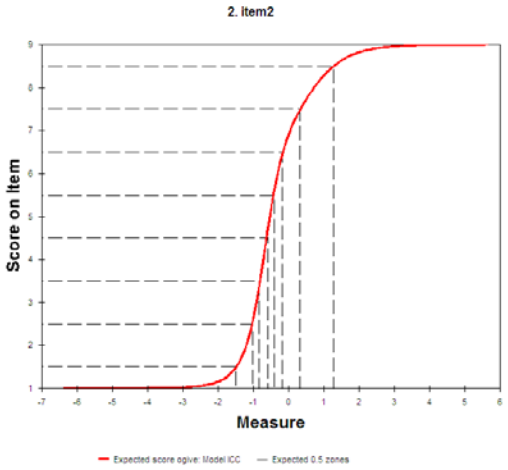


Figure C2: ICC of item 2

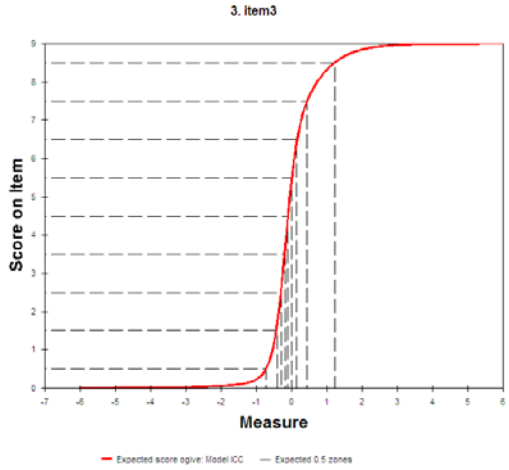


Figure C3: ICC of item 3

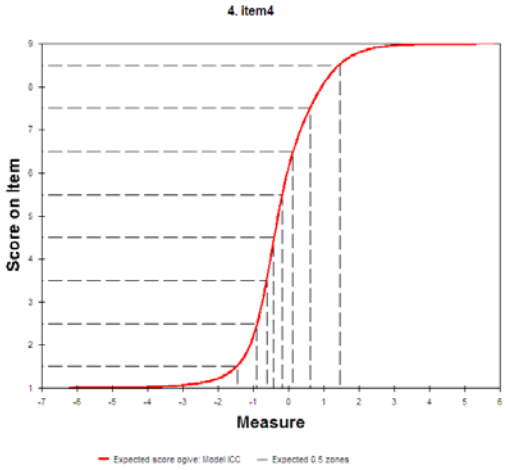


Figure C4: ICC of item 4

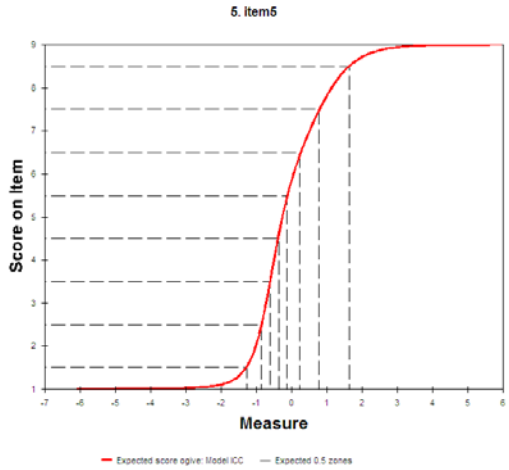


Figure C5: ICC of item 5

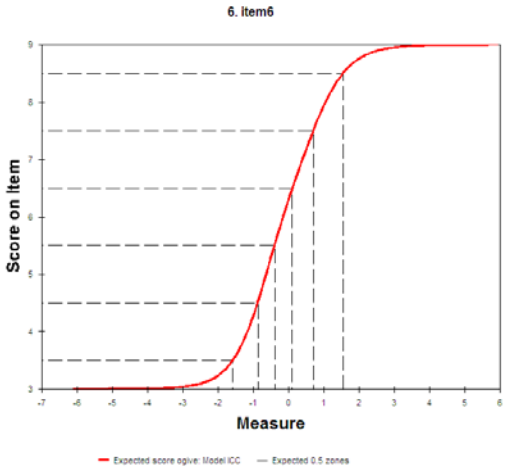


Figure C6: ICC of item 6

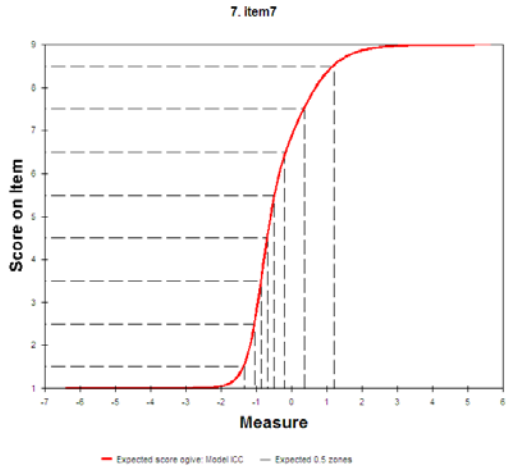


Figure C7: ICC of item 7

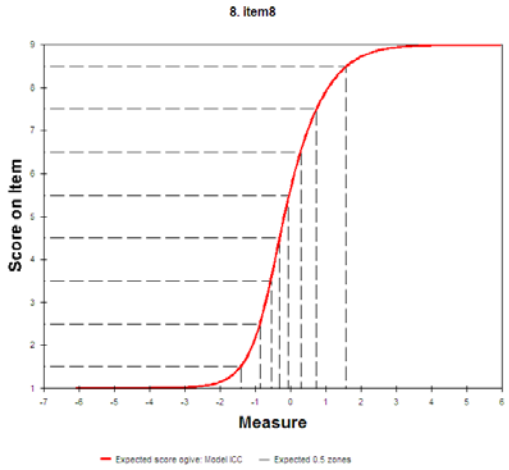


Figure C8: ICC of item 8

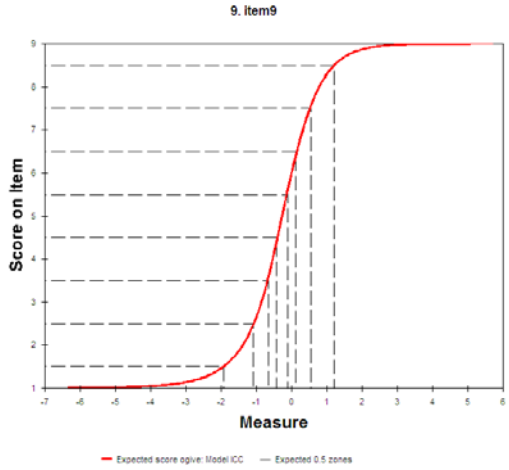


Figure C9: ICC of item 9

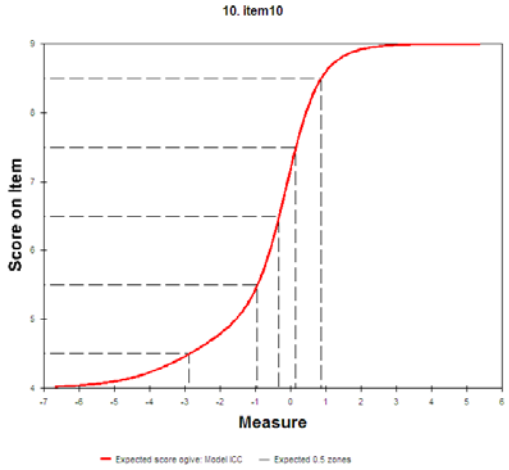


Figure C10: ICC of item 10

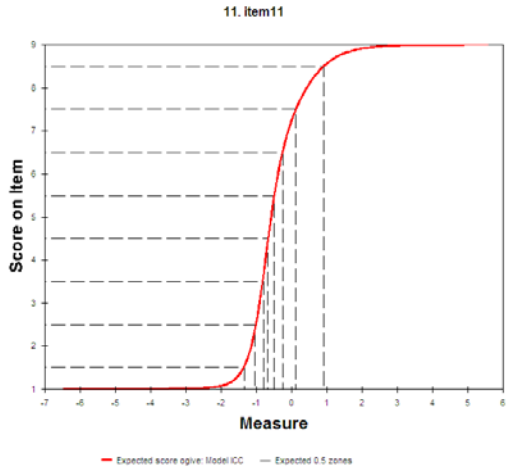


Figure C11: ICC of item 11

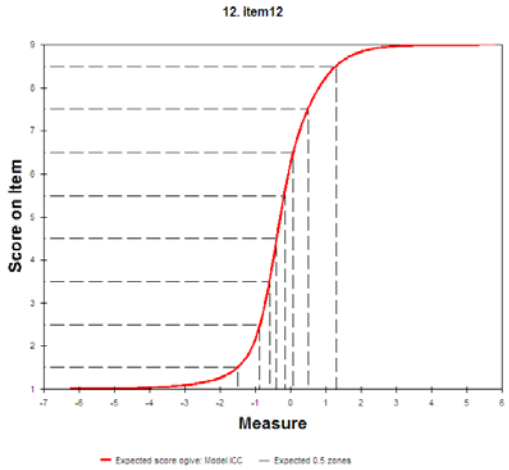


Figure C6: ICC of item 12

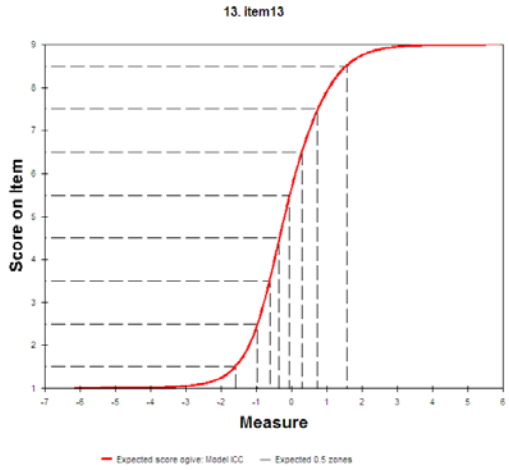


Figure C13: ICC of item 13

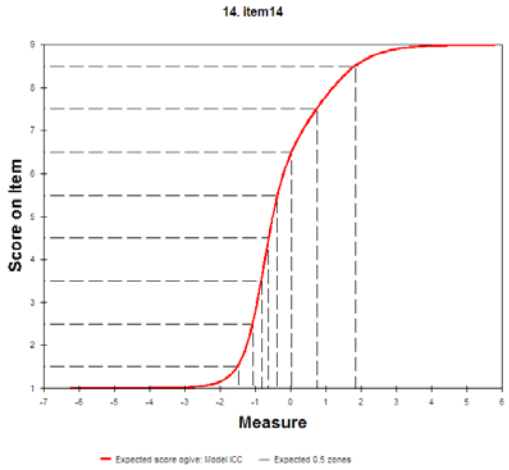


Figure C14: ICC of item 14

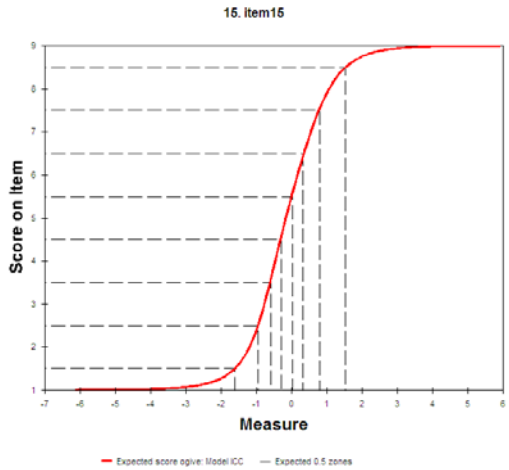


Figure C15: ICC of item 15

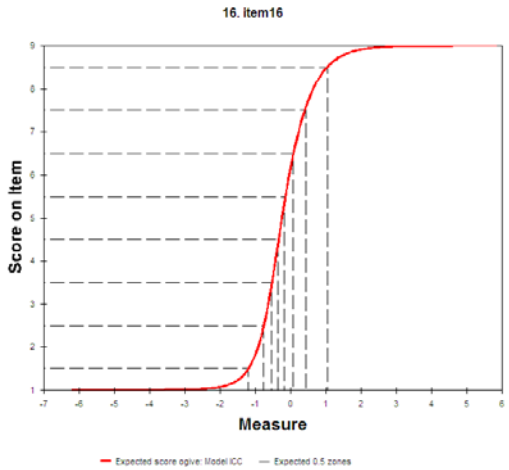


Figure C16: ICC of item 16

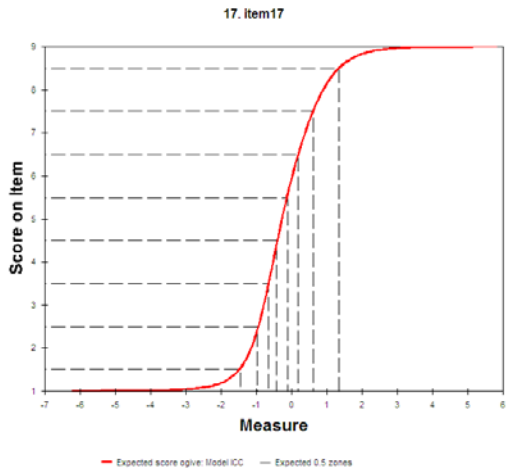


Figure C17: ICC of item 17

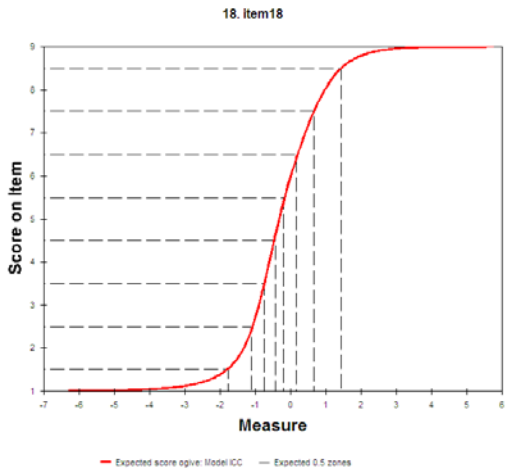


Figure C18: ICC of item 18

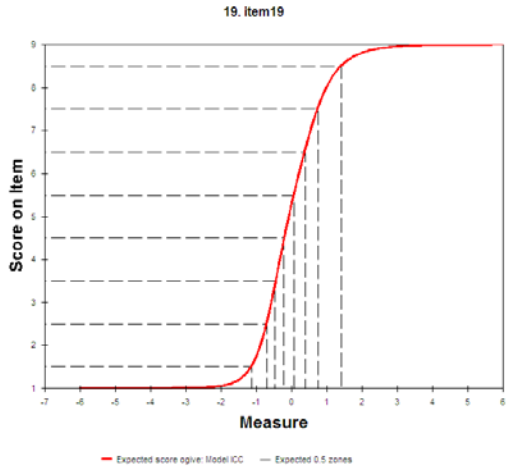


Figure C19: ICC of item 19

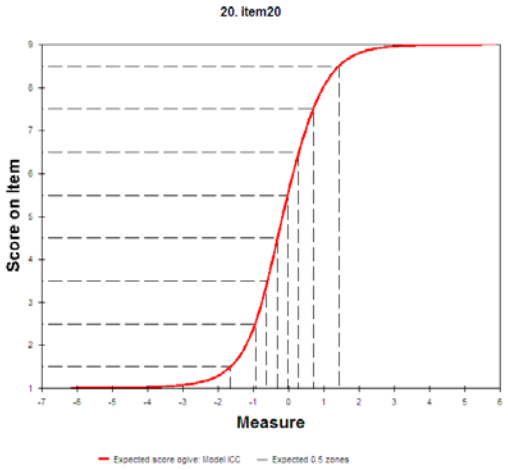


Figure C20: ICC of item 20

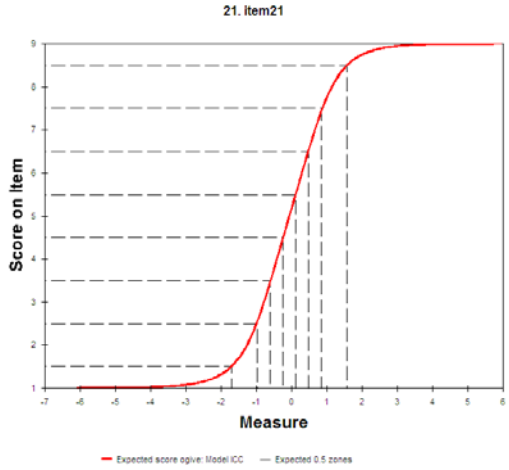


Figure C21: ICC of item 21

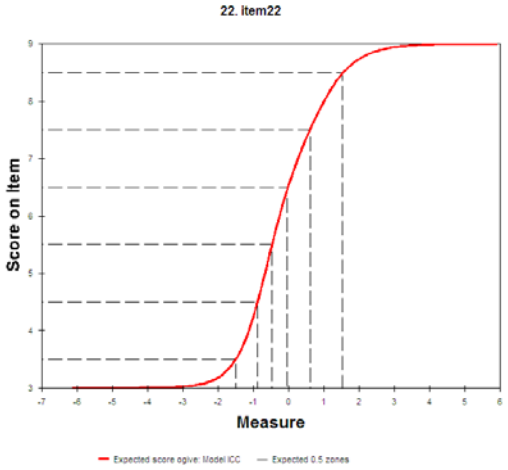


Figure C22: ICC of item 22

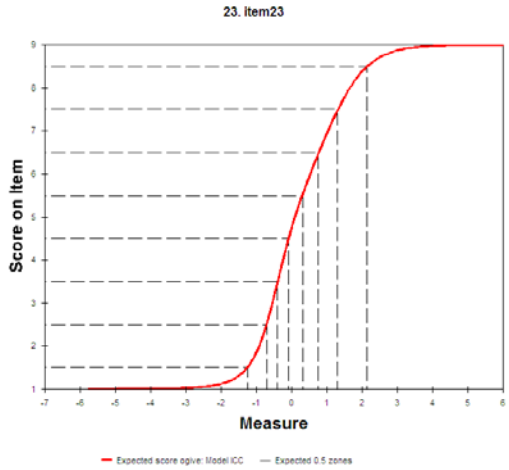


Figure C23: ICC of item 23

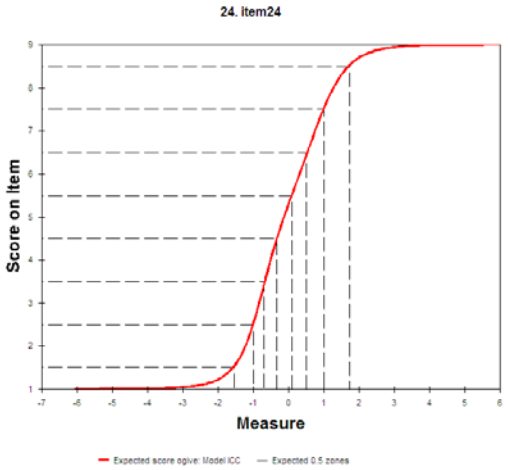


Figure C24: ICC of item 24

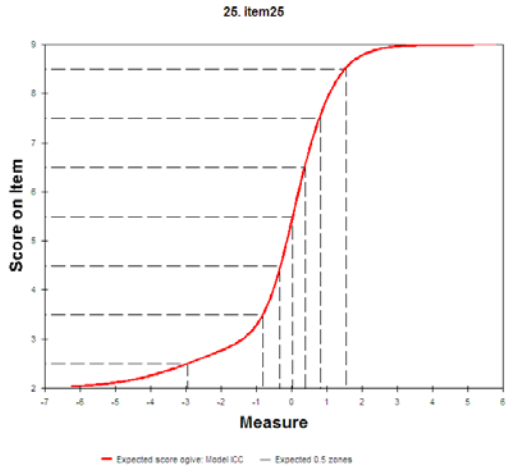


Figure C25: ICC of item 25

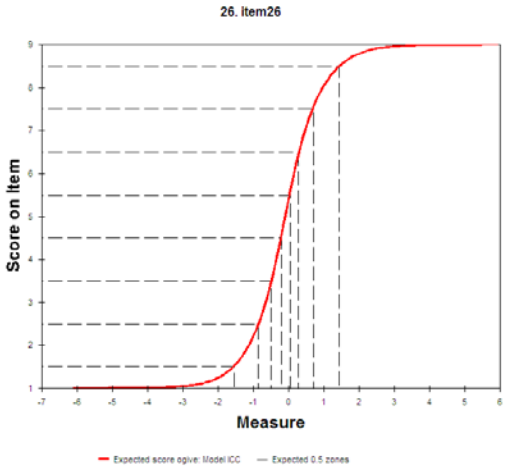


Figure C26: ICC of item 26

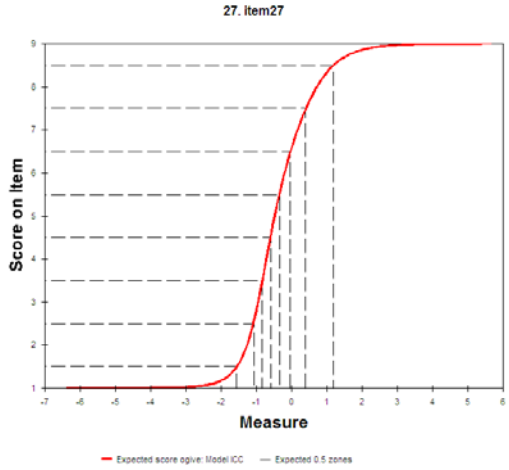


Figure C27: ICC of item 27

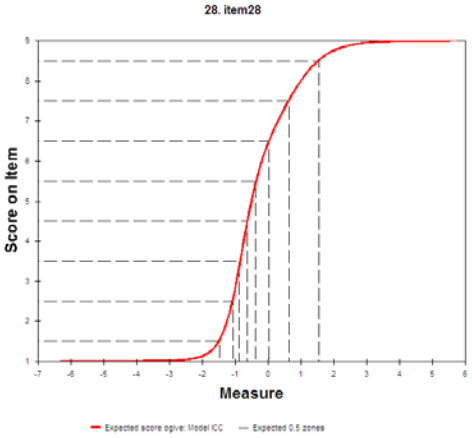


Figure C28: ICC of item 28

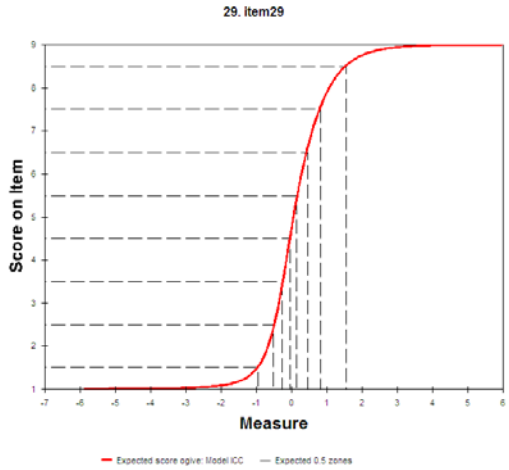


Figure C29: ICC of item 29

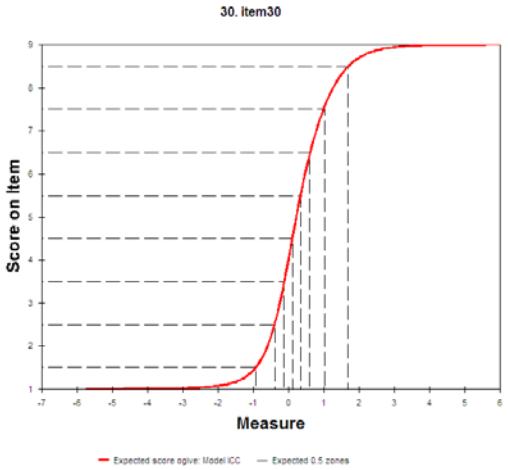


Figure C30: ICC of item 30

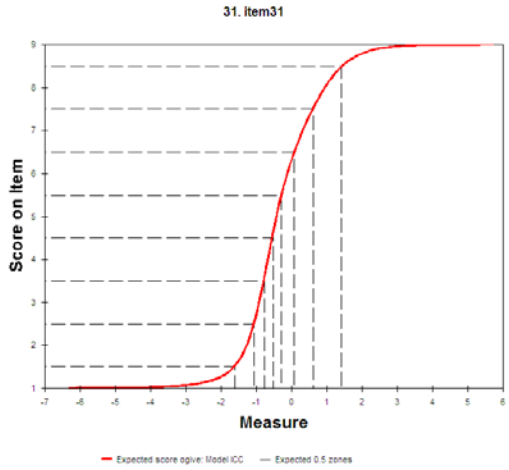


Figure C31: ICC of item 31

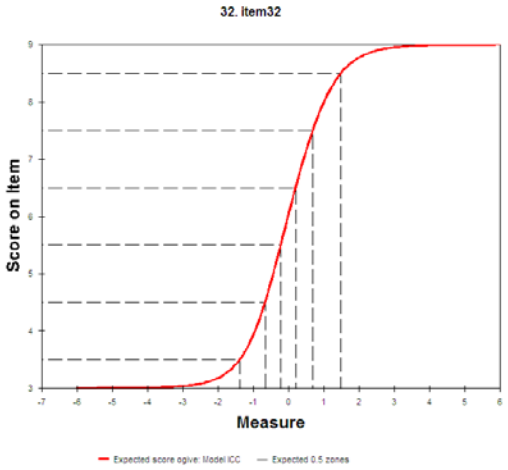


Figure C32: ICC of item 32

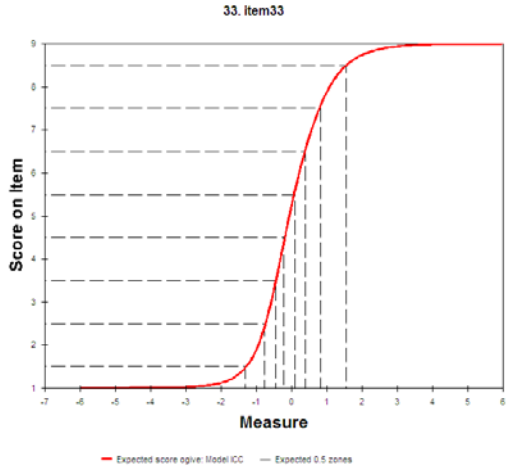


Figure C33: ICC of item 33

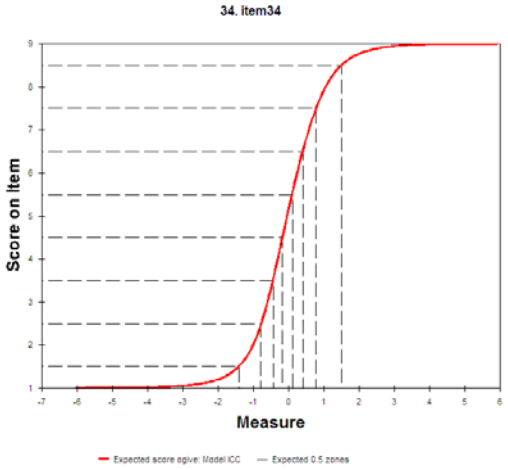


Figure C34: ICC of item 34

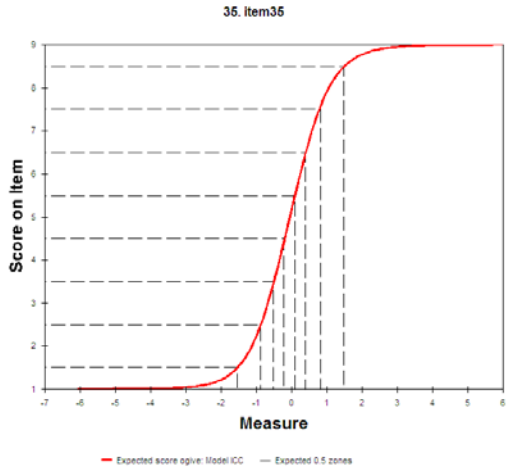


Figure C35: ICC of item 35

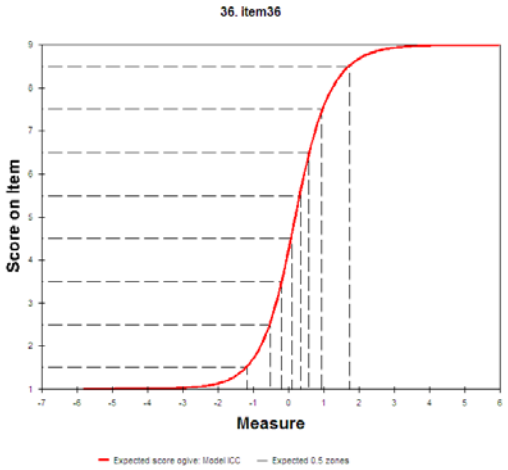


Figure C36: ICC of item 36

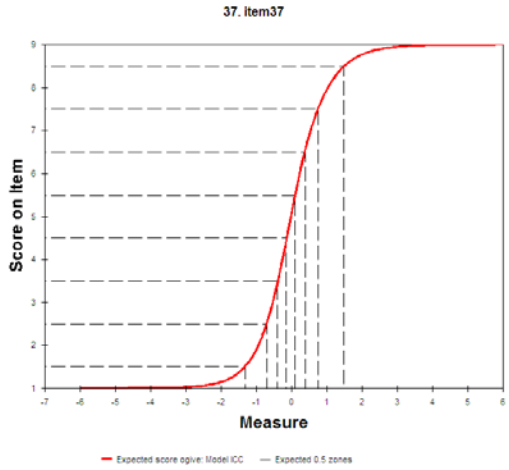


Figure C37: ICC of item 37

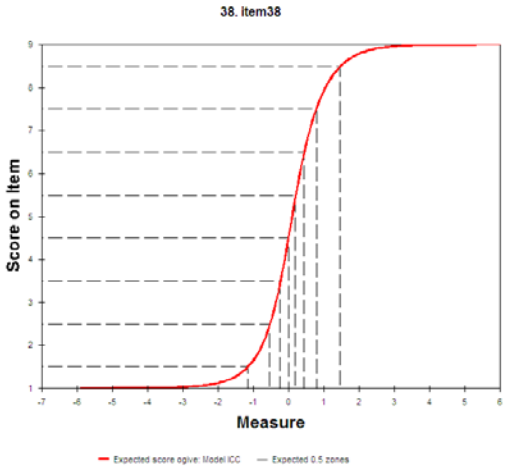


Figure C38: ICC of item 38

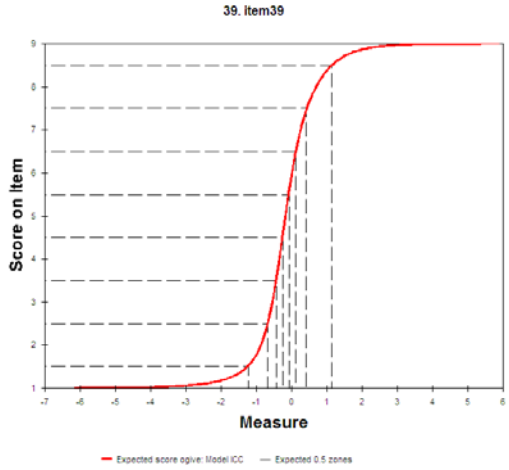


Figure C39: ICC of item 39

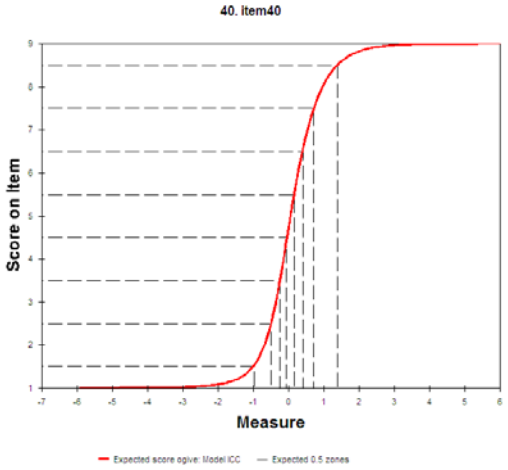


Figure C40: ICC of item 40

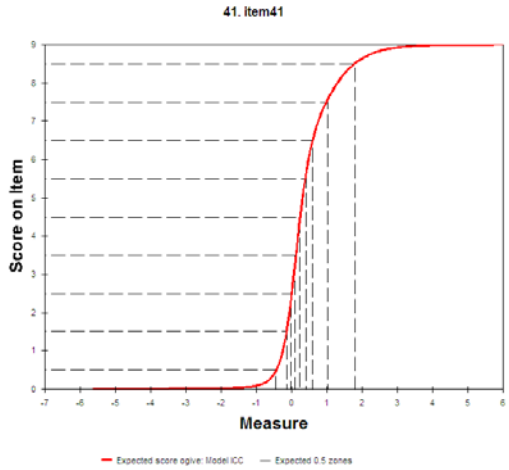


Figure C41: ICC of item 41

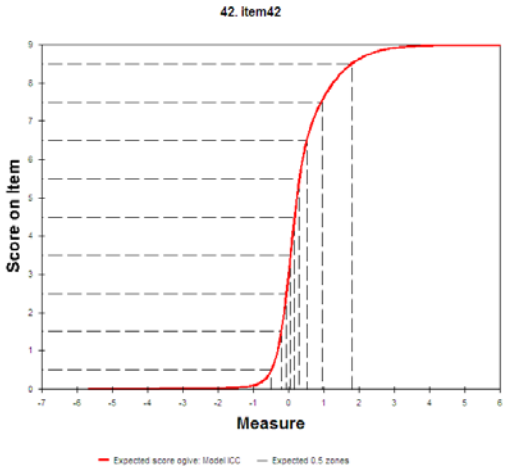


Figure C42: ICC of item 42

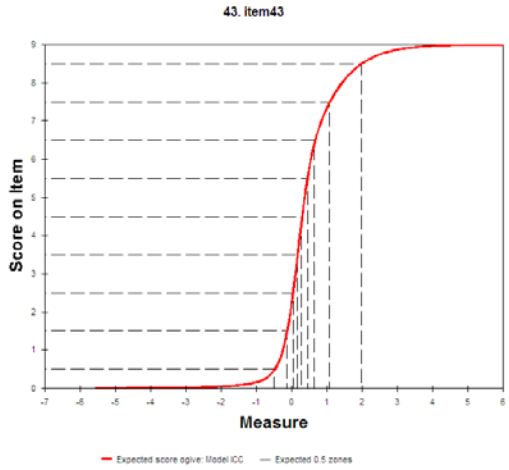


Figure C43: ICC of item 43

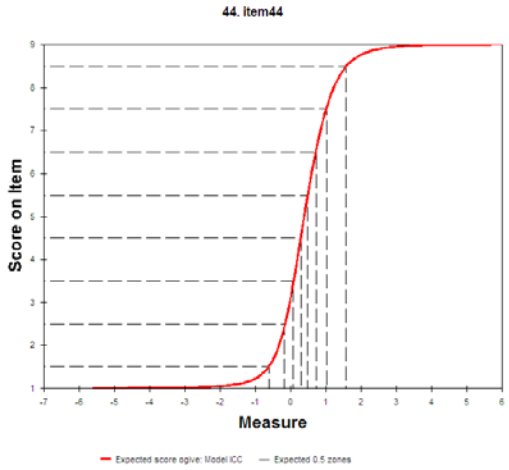


Figure C44: ICC of item 44

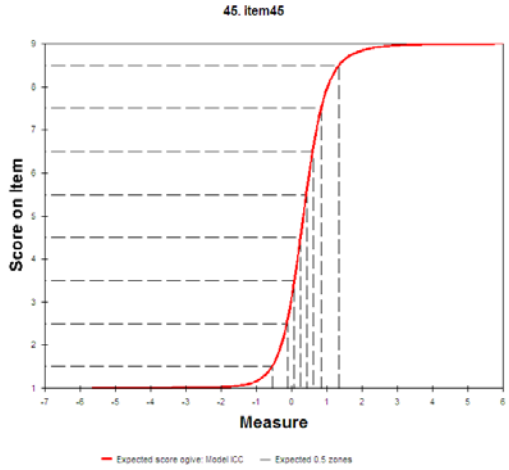


Figure C45: ICC of item 45

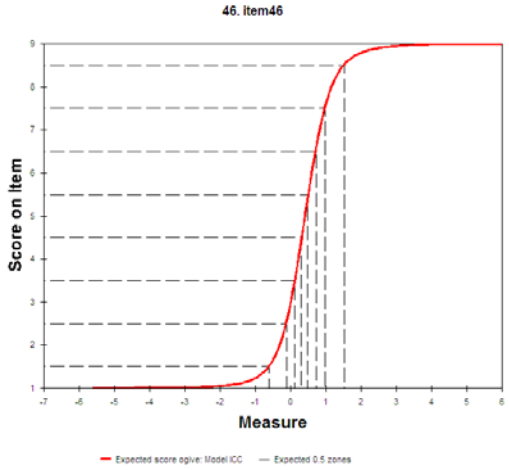


Figure C46: ICC of item 46

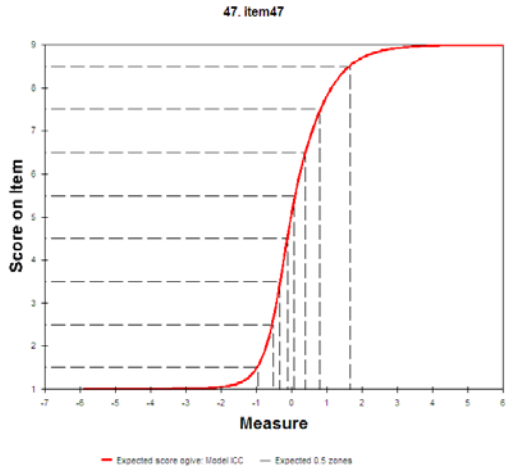


Figure C47: ICC of item 47

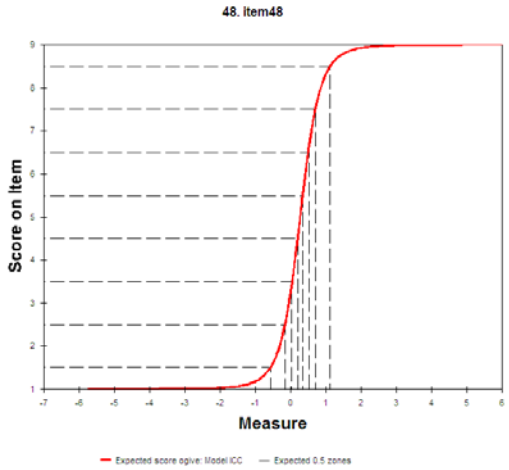


Figure C48: ICC of item 48

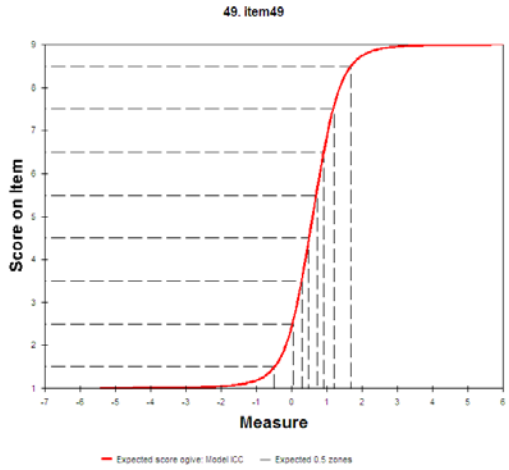


Figure C49: ICC of item 49

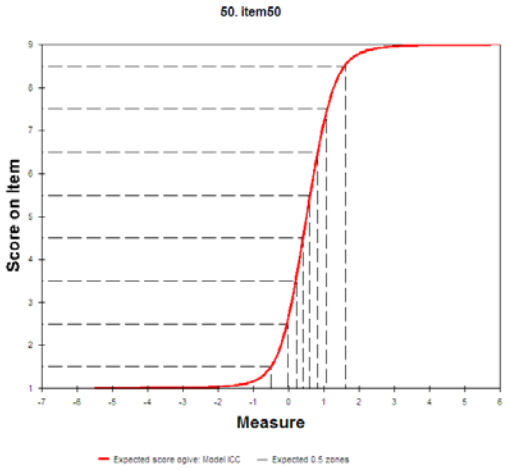


Figure C50: ICC of item 50

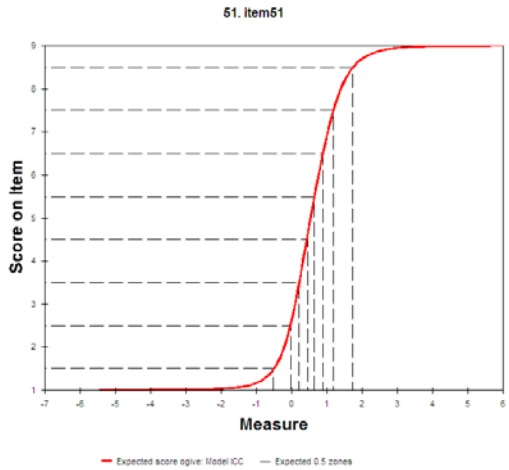


Figure C51: ICC of item 51

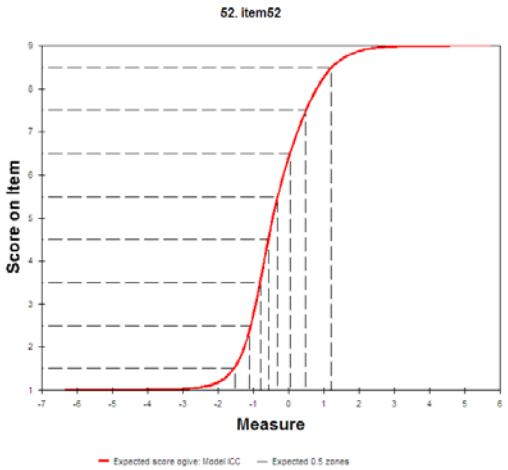


Figure C52: ICC of item 52

Parent proxy report item characteristic curve

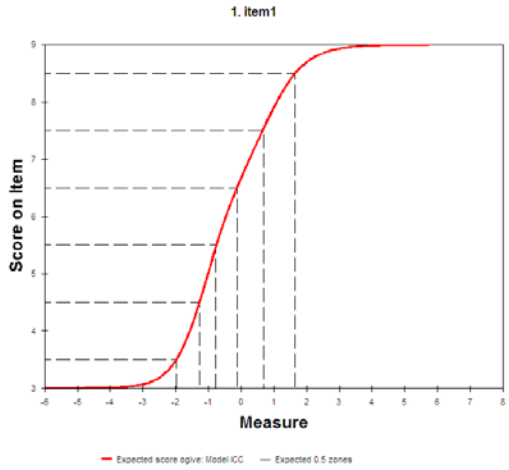


Figure D1: ICC of item 1

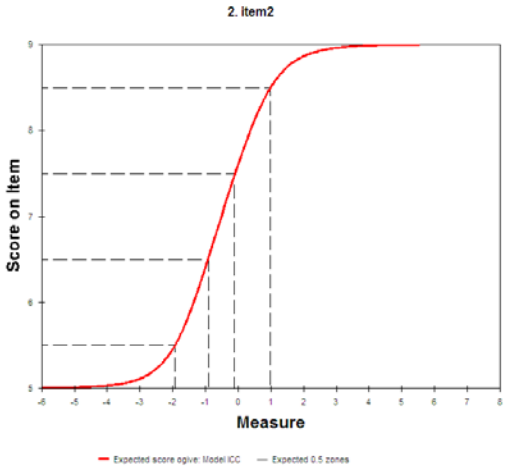


Figure D2: ICC of item 2

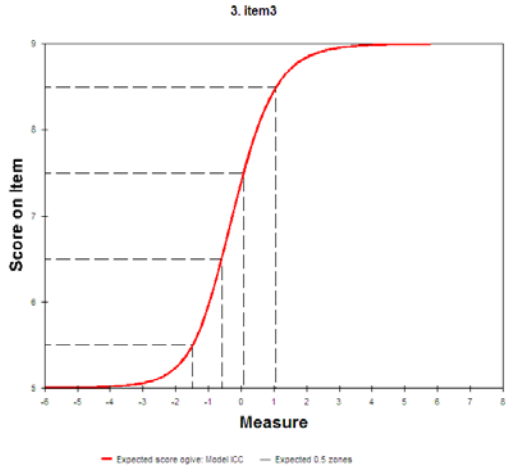


Figure D3: ICC of item 3

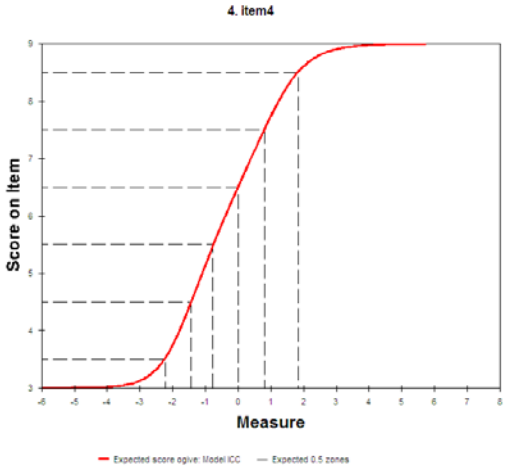


Figure D4: ICC of item 4

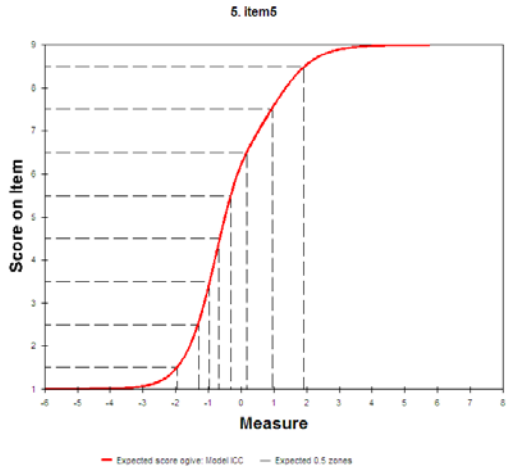


Figure D5: ICC of item 5

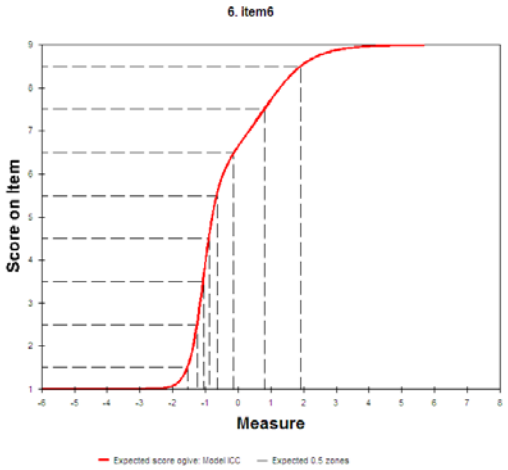


Figure D6: ICC of item 6

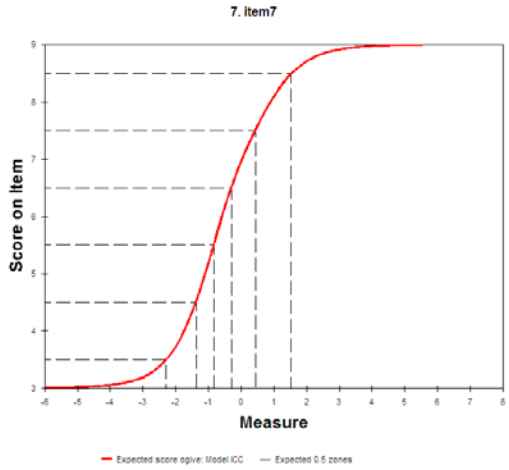


Figure D7: ICC of item 7

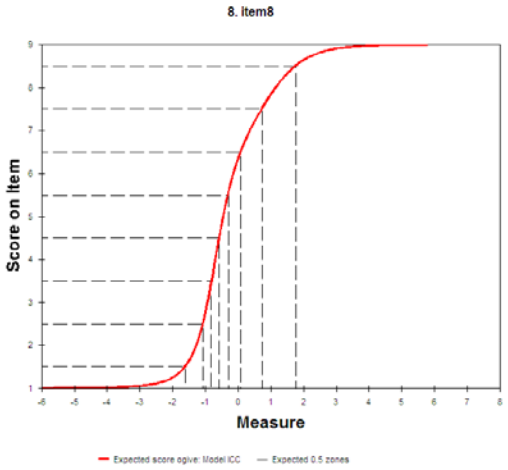


Figure D8: ICC of item 8

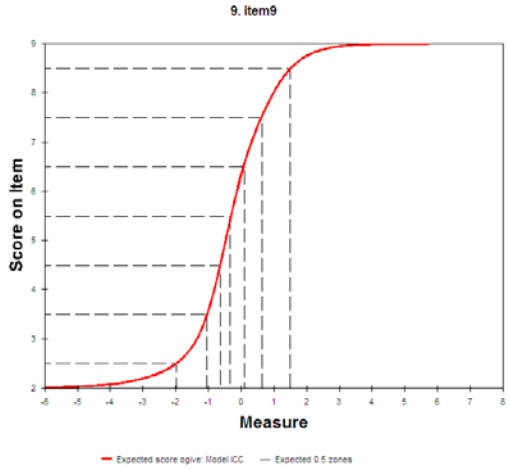


Figure D9: ICC of item 9

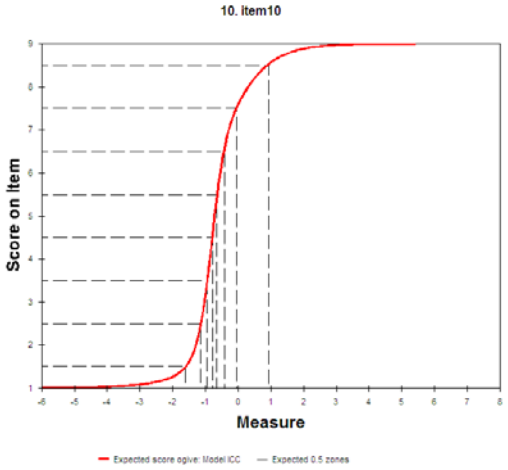


Figure D10: ICC of item 10

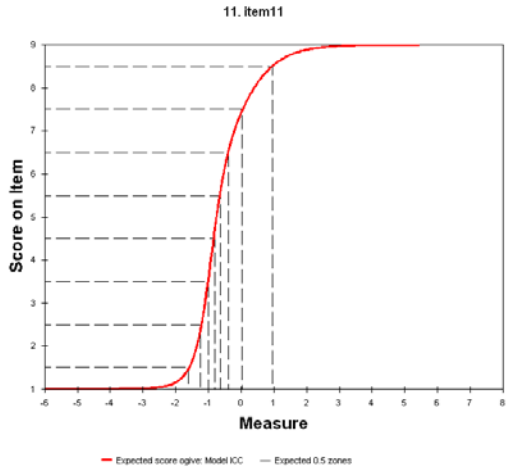


Figure D11: ICC of item 11

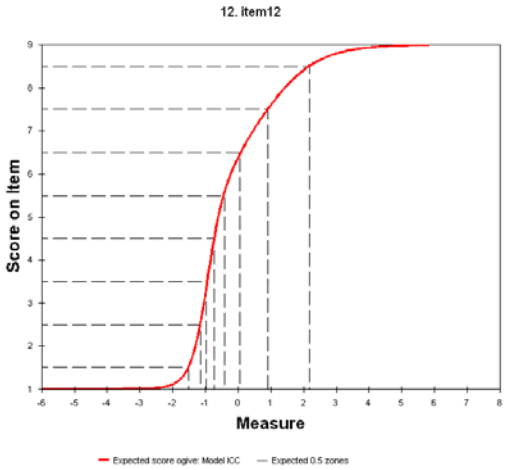


Figure D12: ICC of item 12

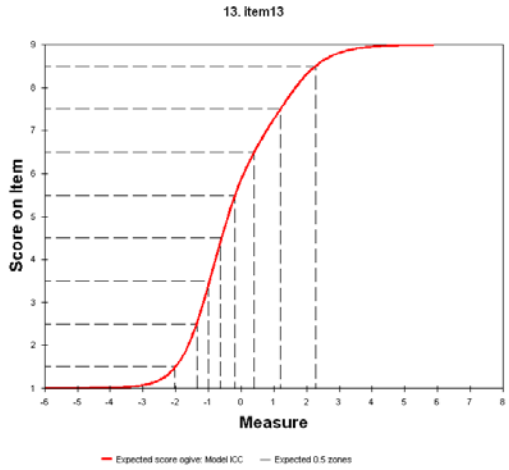


Figure D13: ICC of item 13

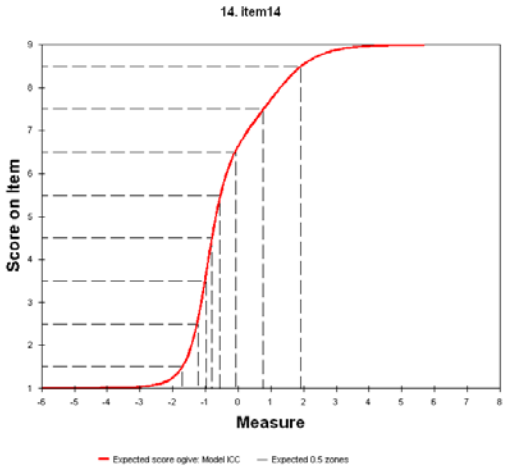


Figure D14: ICC of item 14

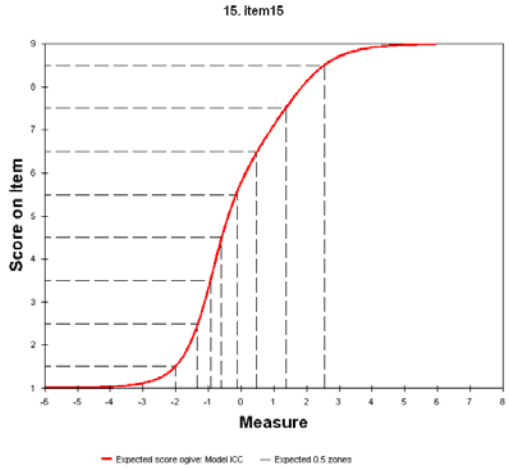


Figure D15: ICC of item 15

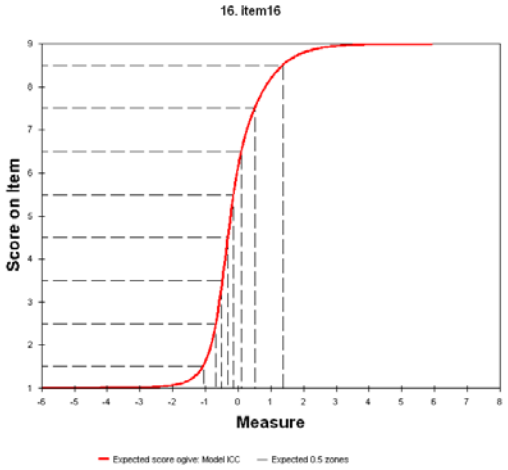


Figure D16: ICC of item 16

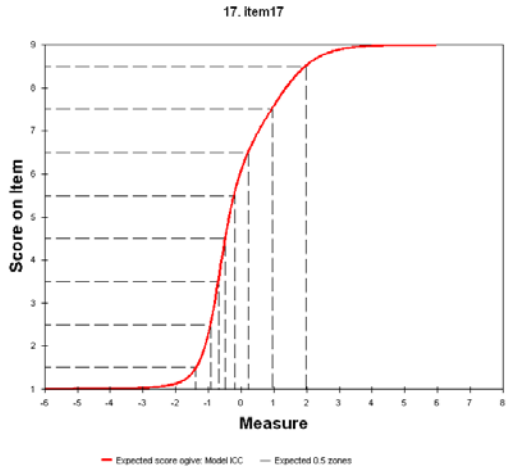


Figure D17: ICC of item 17

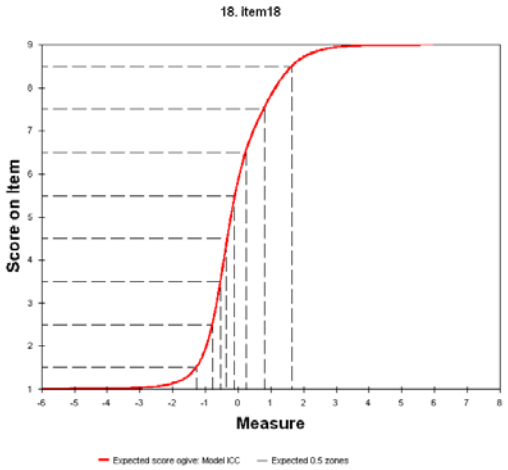


Figure D18: ICC of item 18

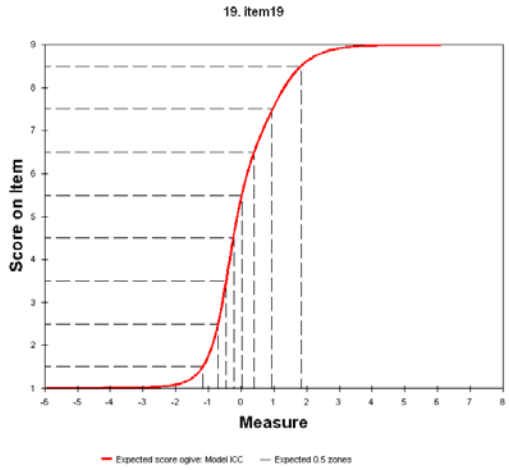


Figure D19: ICC of item 19

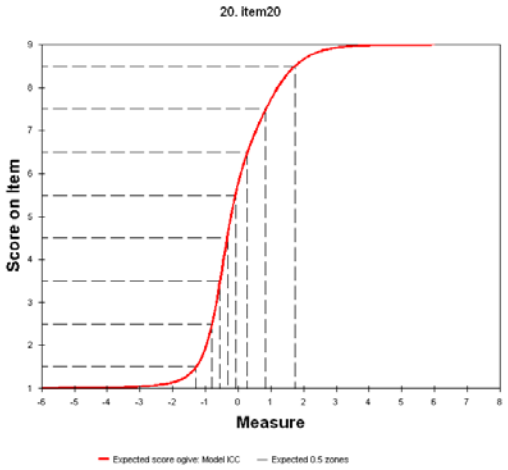


Figure D20: ICC of item 20

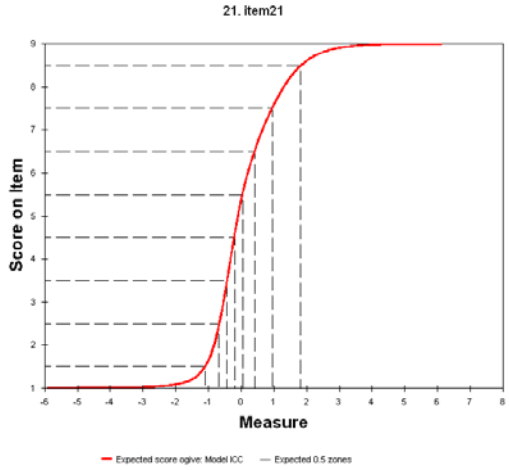


Figure D21: ICC of item 21

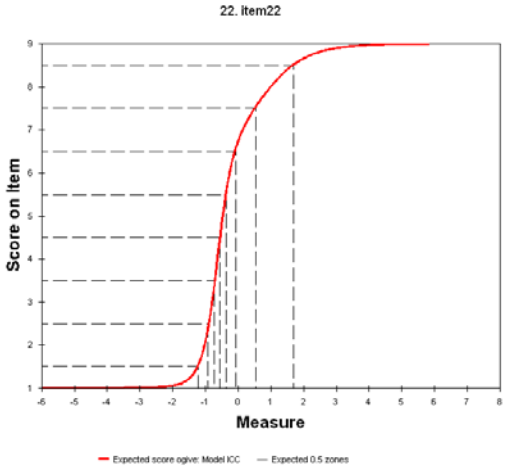


Figure D22: ICC of item 22

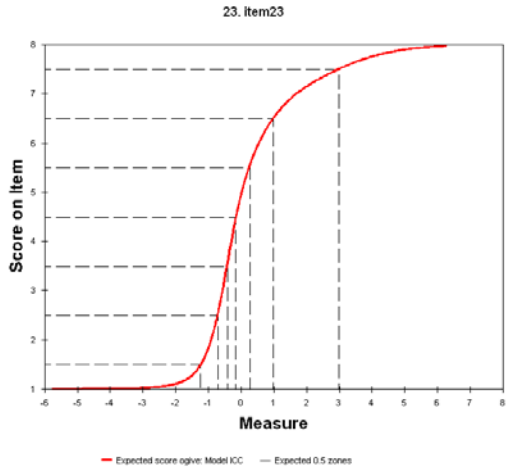


Figure D23: ICC of item 23

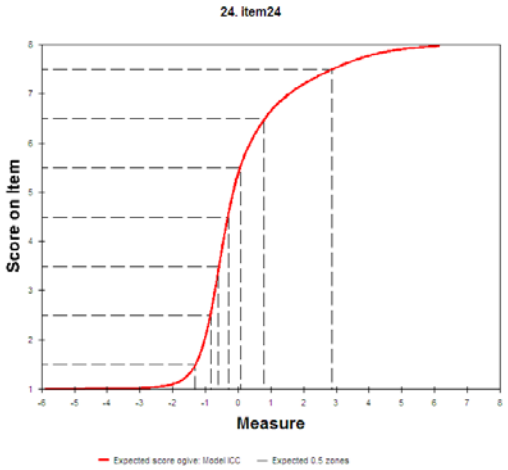


Figure D24: ICC of item 24

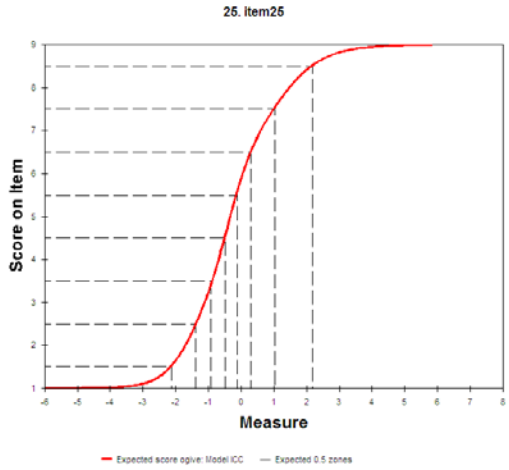


Figure D25: ICC of item 25

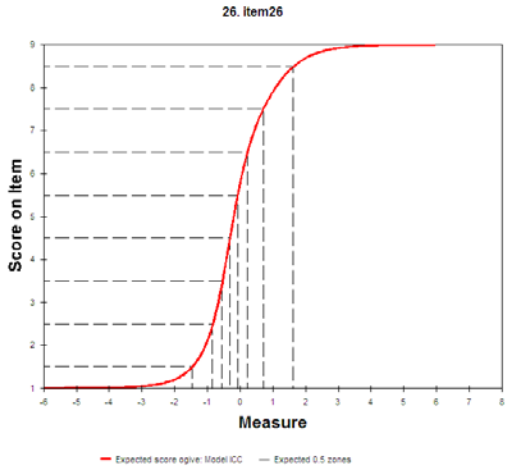


Figure D26: ICC of item 26

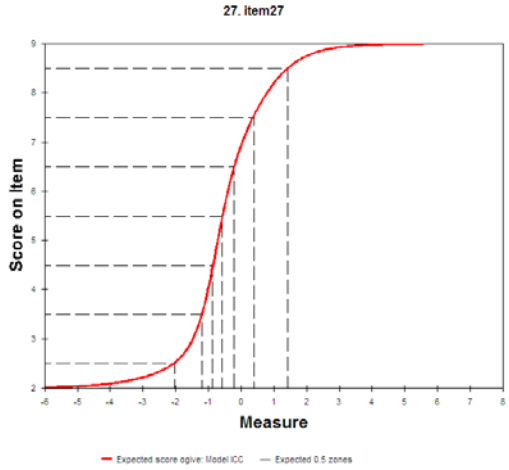


Figure D27: ICC of item 27

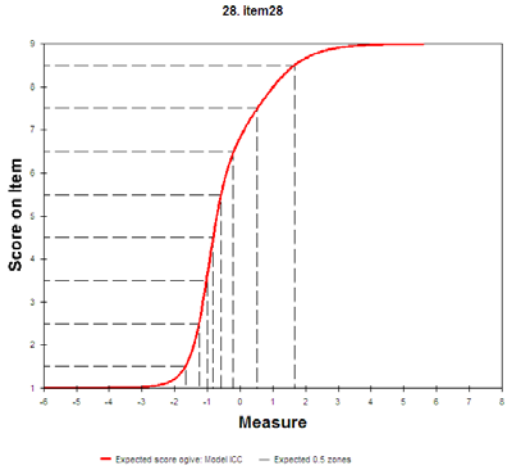


Figure D28: ICC of item 28

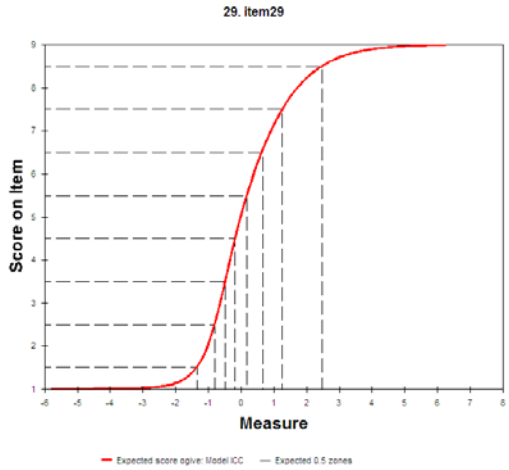


Figure D29: ICC of item 29

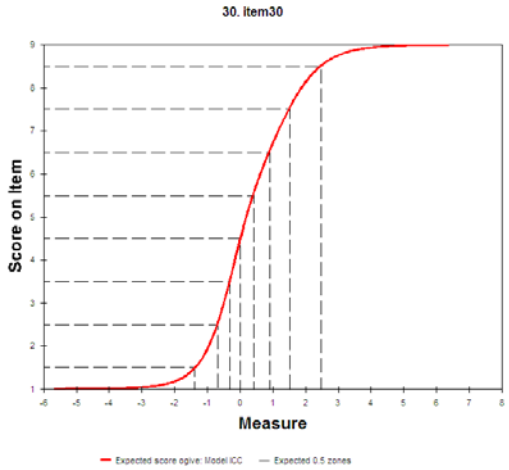


Figure D30: ICC of item 30

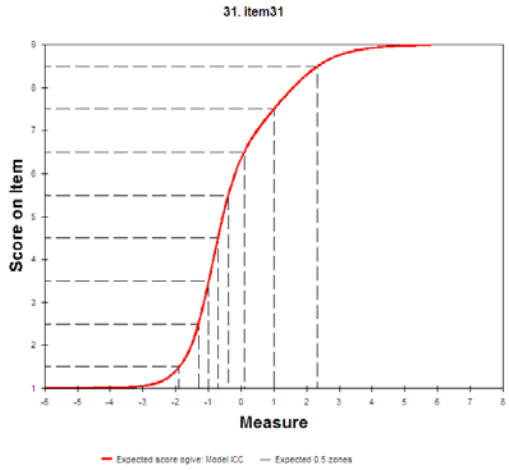


Figure D31: ICC of item 31

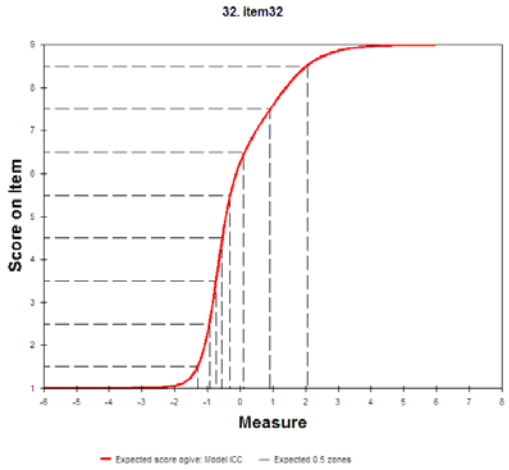


Figure D32: ICC of item 32

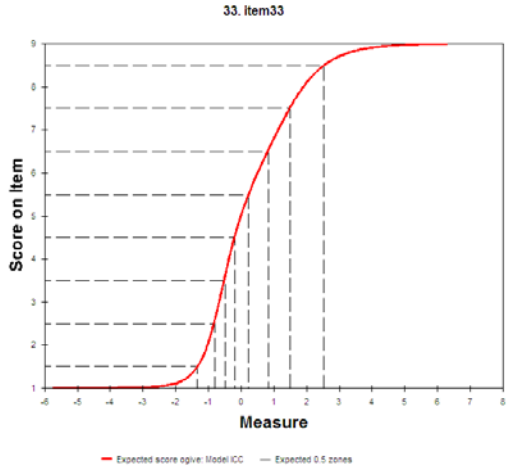


Figure D33: ICC of item 33

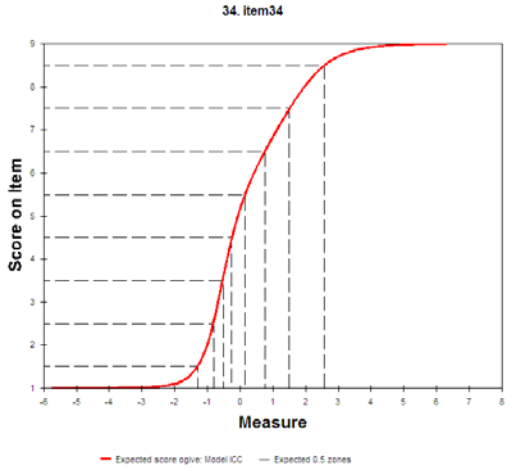


Figure D34: ICC of item 34

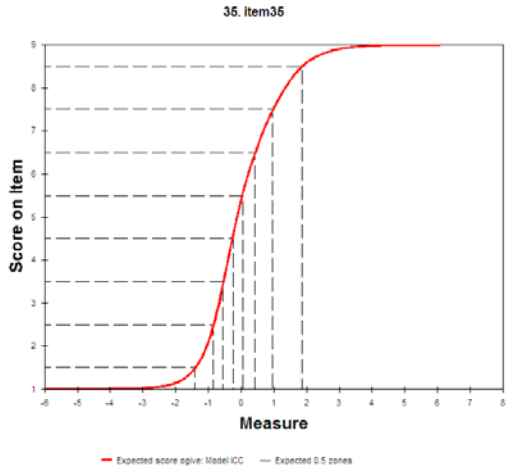


Figure D35: ICC of item 35

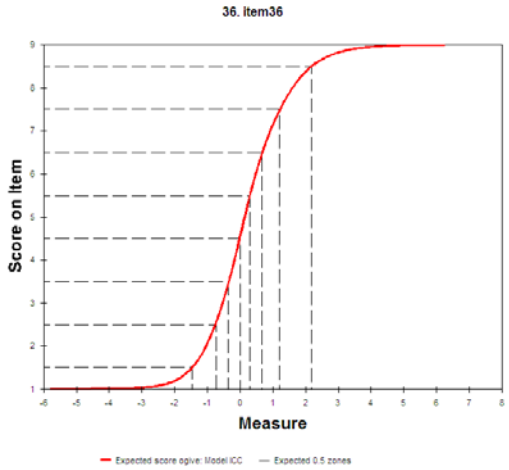


Figure D36: ICC of item 36

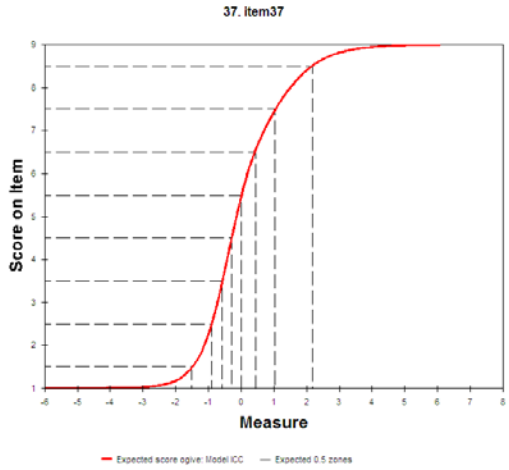


Figure D37: ICC of item 37

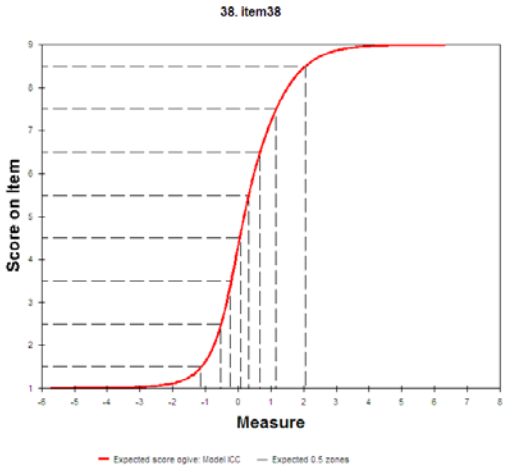


Figure D38: ICC of item 38

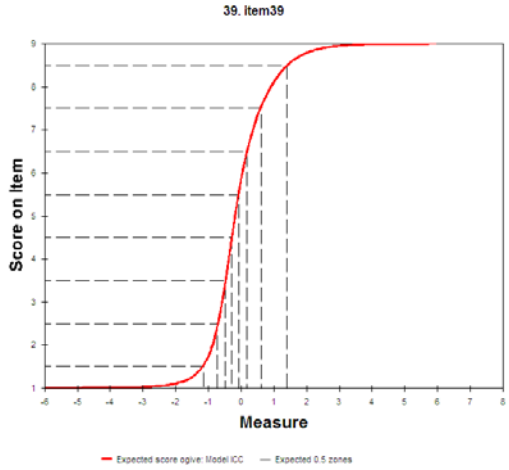


Figure D39: ICC of item 39

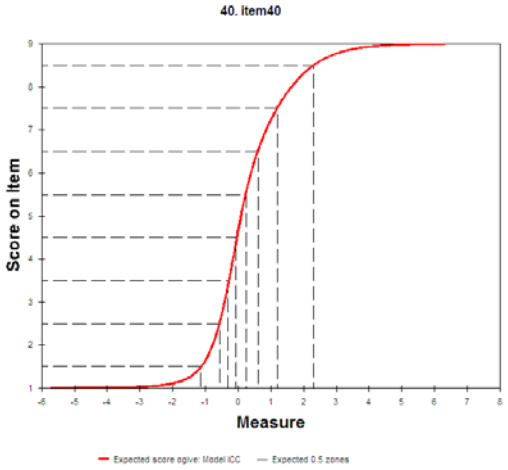


Figure D40: ICC of item 40

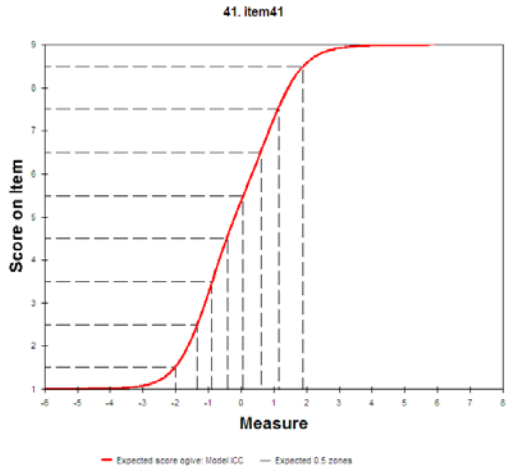


Figure D41: ICC of item 41

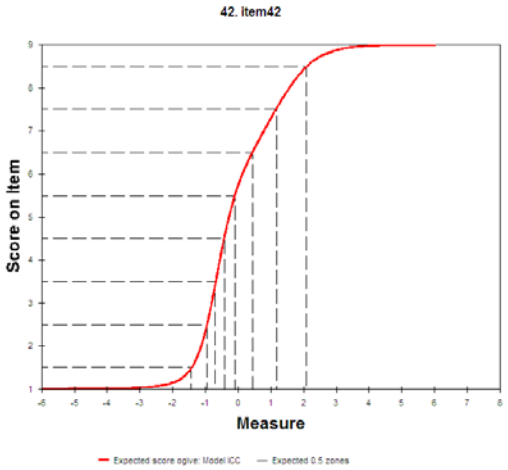


Figure D42: ICC of item 42

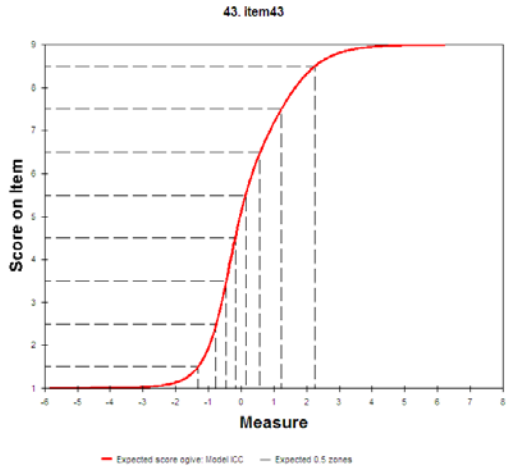


Figure D43: ICC of item 43

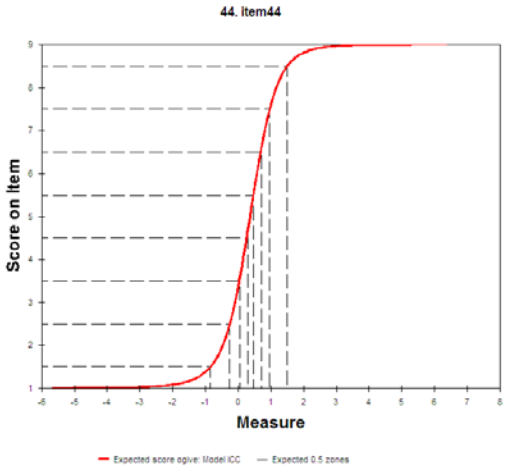


Figure D44: ICC of item 44

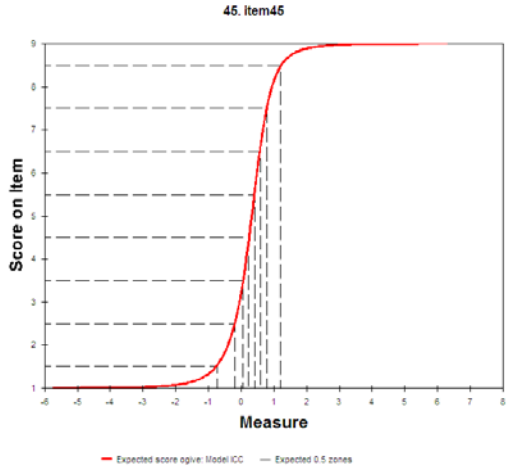


Figure D45: ICC of item 45

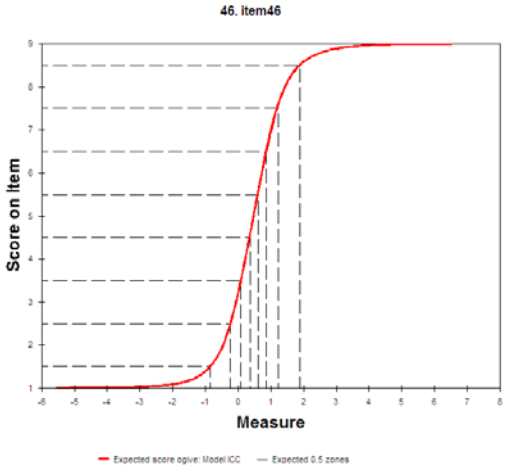


Figure D46: ICC of item 46

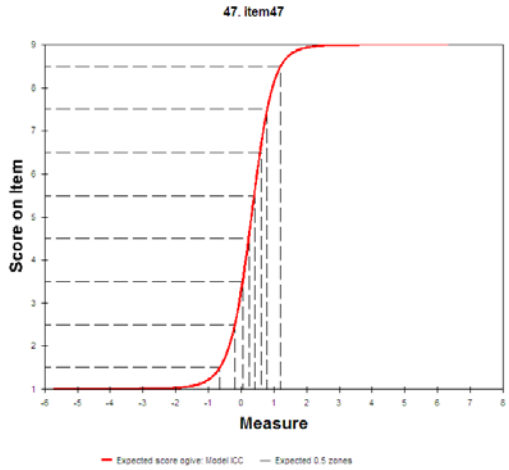


Figure D47: ICC of item 47

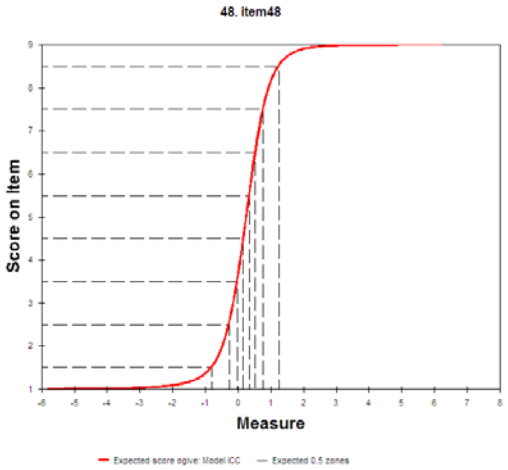


Figure D48: ICC of item 48

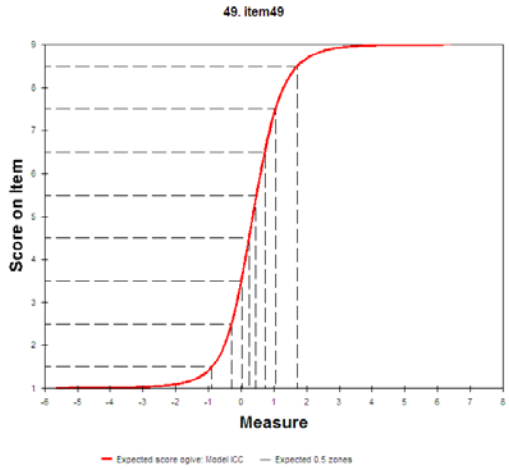


Figure D49: ICC of item 49

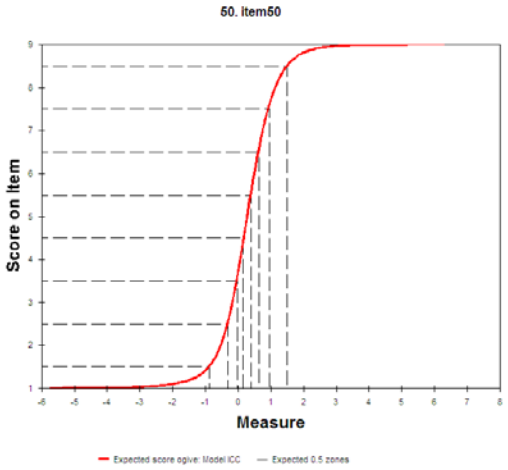


Figure D50: ICC of item 50

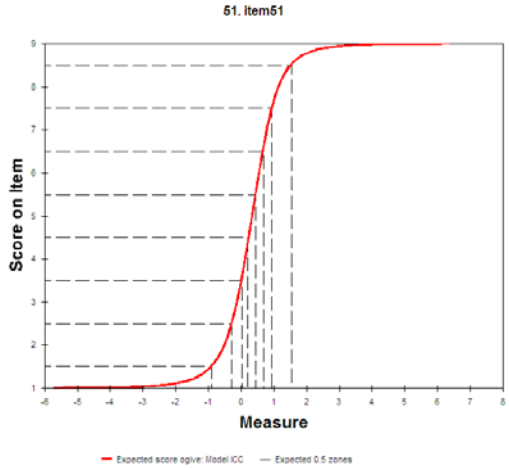


Figure D51: ICC of item 51

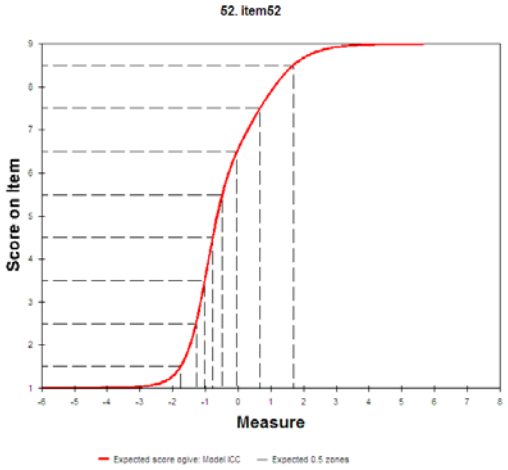


Figure D52: ICC of item 52

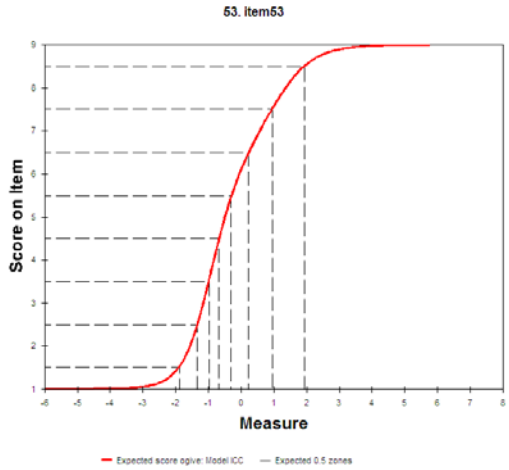


Figure D53: ICC of item 53

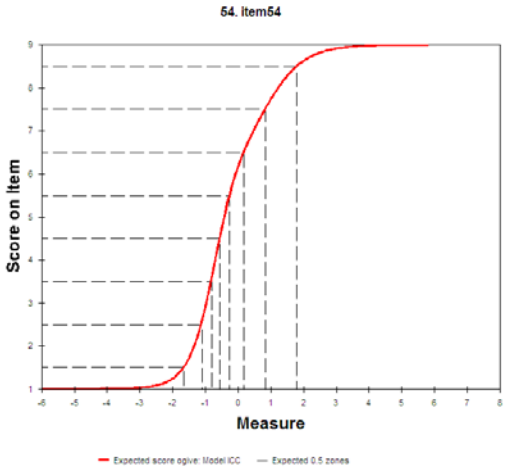


Figure D54: ICC of item 54

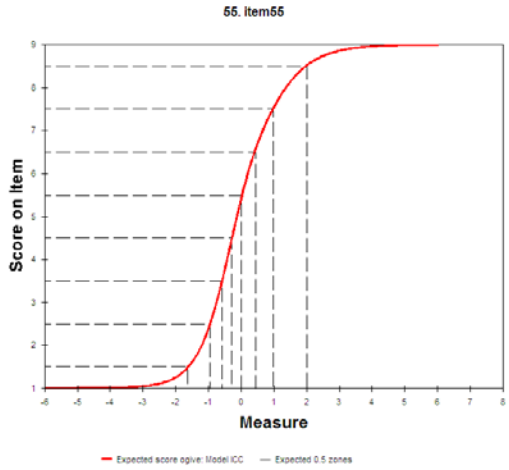


Figure D55: ICC of item 55

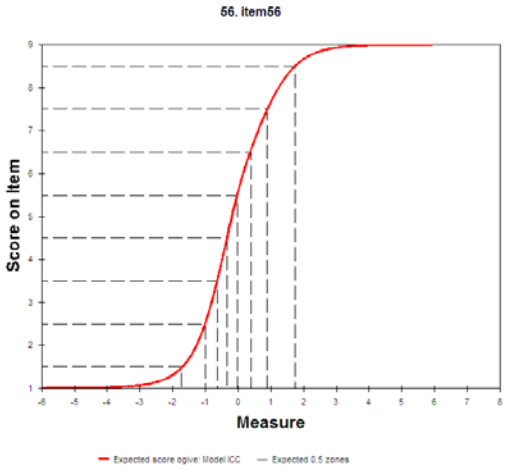


Figure D56: ICC of item 56

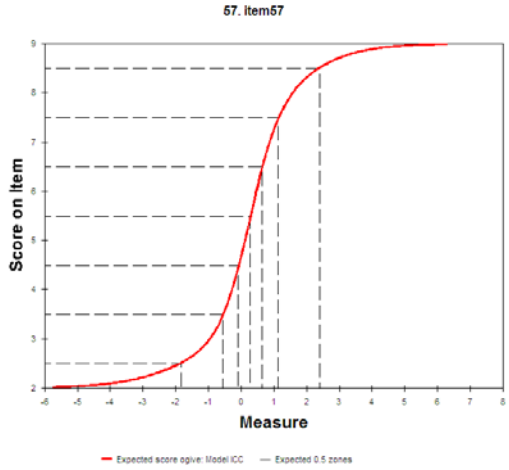


Figure D57: ICC of item 57

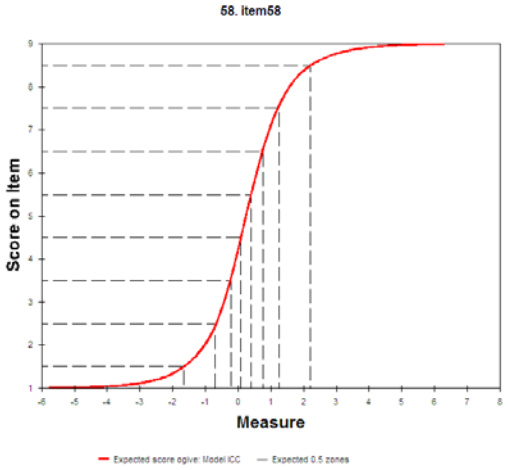


Figure D58: ICC of item 58

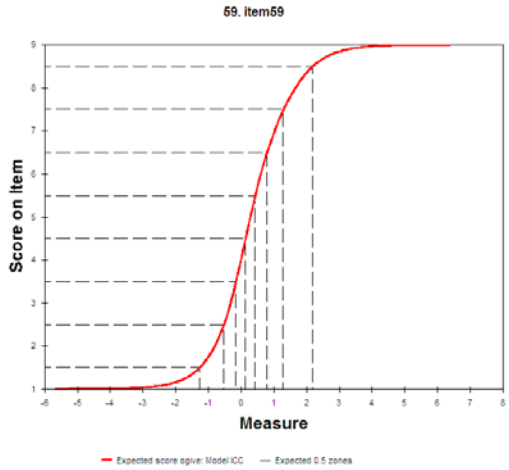


Figure D59: ICC of item 59

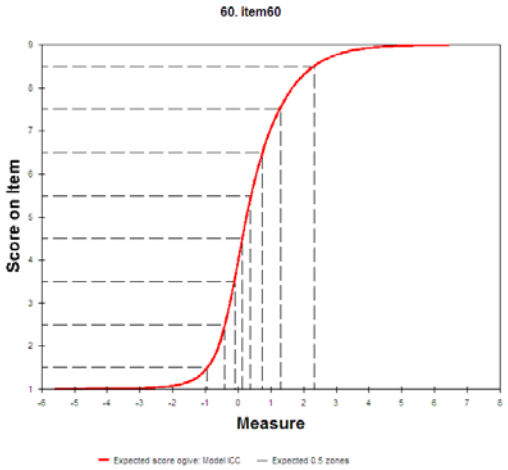


Figure D60: ICC of item 60

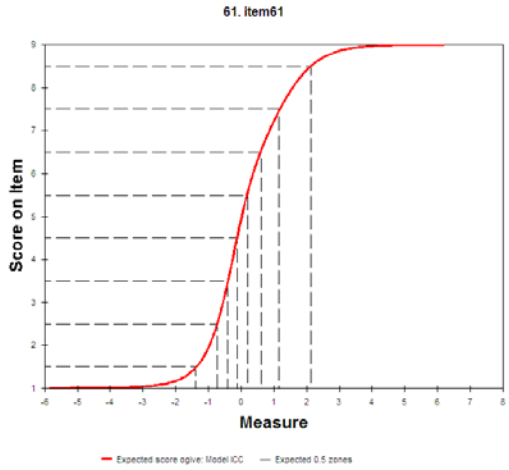


Figure D61: ICC of item 61

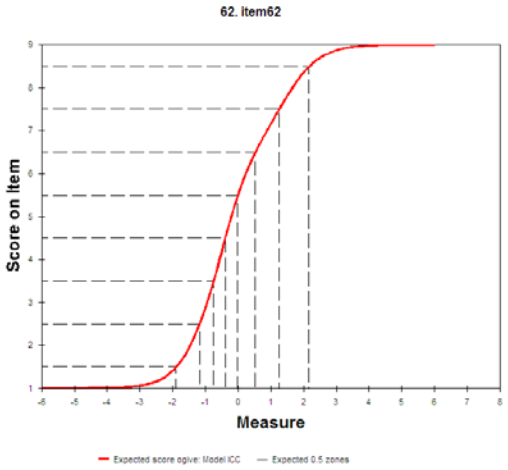


Figure D62: ICC of item 62

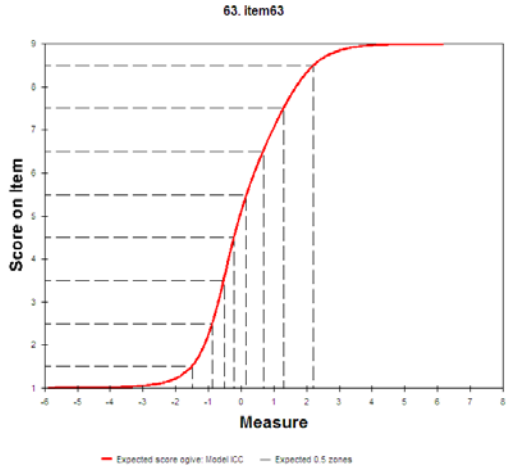


Figure D63: ICC of item 63

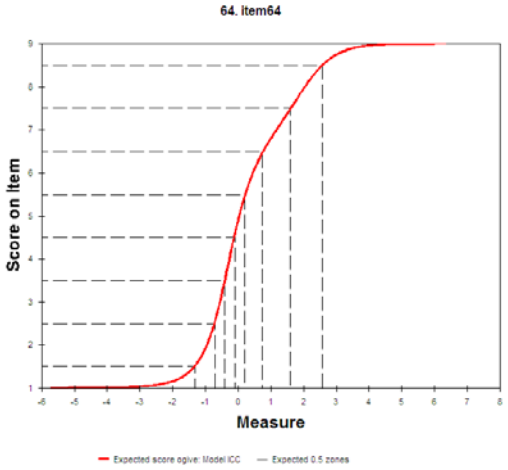


Figure D64: ICC of item 64

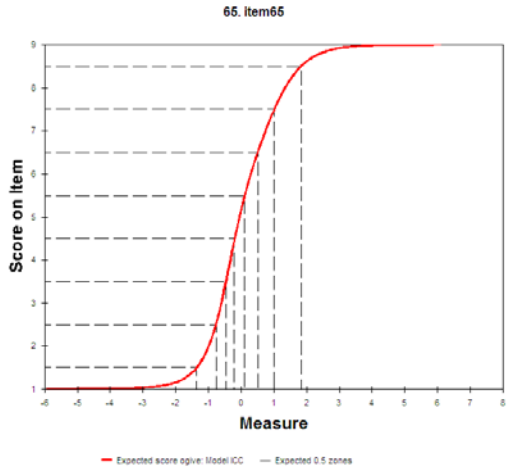


Figure D65: ICC of item 65

BIBLIOGRAPHY

References

Amichai, T., Harries, N., Dvir, Z., Patish, H., & Copeliovitch, L. (2009). The effects of femoral derotation osteotomy in children with cerebral palsy: an evaluation using energy cost and functional mobility. *Journal of Pediatric Orthopedics, 29*(1), 68-72. doi: 10.1097/BPO.0b013e3181924331

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561-573. doi: 10.1007/bf02293814

Arnaud, C., White-Koning, M., Michelsen, S. I., Parkes, J., Parkinson, K., Thyen, U., . . . Colver, A. (2008). Parent-Reported Quality of Life of Children With Cerebral Palsy in Europe. *Pediatrics, 121*(1), 54-64. doi: 10.1542/peds.2007-0854

Avery, L. M., Russell, D. J., Raina, P. S., Walter, S. D., & Rosenbaum, P. L. (2003). Rasch analysis of the gross motor function measure: validating the assumptions of the rasch model to create an interval-level Measure. *Archives of Physical Medicine and Rehabilitation, 84*(5), 697-705. doi: 10.1016/s0003-9993(02)04896-7

Bagley, A. M., Gorton, G., Oeffinger, D., Barnes, D., Calmes, J., Nicholson, D., . . . Tylkowski, C. (2007). Outcome assessments in children with cerebral palsy, part II: discriminatory ability of outcome tools. *Developmental Medicine & Child Neurology, 49*(3), 181-186. doi: 10.1111/j.1469-8749.2007.00181.x

Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2010). A Method for the Comparison of Item Selection Rules in Computerized Adaptive Testing. *Applied Psychological Measurement, 34*(6), 438-452. doi: 10.1177/0146621610370152

Barrada, J. R., Veldkamp, B. P., & Olea, J. (2009). Multiple Maximum Exposure Rates in Computerized Adaptive Testing. *Applied Psychological Measurement, 33*(1), 58-73. doi: 10.1177/0146621608315329

Belov, D. I., & Armstrong, R. D. (2009). Direct and Inverse Problems of Item Pool Design for Computerized Adaptive Testing. *Educational and Psychological Measurement, 69*(4), 533-547. doi: 10.1177/0013164409332224

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: fundamental measurement in the human sciences* (2 ed.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Burro, R., Sartori, R., & Vidotto, G. (2011). The method of constant stimuli with three rating categories and the use of Rasch models. *Quality & Quantity, 45*(1), 43-58. doi: 10.1007/s11135-009-9282-3

Carlon, S., Shields, N., Yong, K., Gilmore, R., Sakzewski, L., & Boyd, R. (2010). A systematic review of the psychometric properties of Quality of Life measures for school aged children with cerebral palsy. *BMC Pediatrics, 10(1),* 81. doi: 10.1186/1471-2431-10-81

Chen, C.-M., Chen, C.-Y., Wu, K. P., Chen, C.-L., Hsu, H.-C., & Lo, S.-K. (2011). Motor Factors Associated with Health-Related Quality-of-Life in Ambulatory Children with

Cerebral Palsy. *American Journal of Physical Medicine & Rehabilitation, 90*(11), 940-947 910.1097/PHM.1090b1013e3182240d3182254.

Chen, S.-Y. (2010). A Procedure for Controlling General Test Overlap in Computerized Adaptive Testing. *Applied Psychological Measurement, 34*(6), 393-409. doi: 10.1177/0146621610367788

Cheng, Y.-Y., Wang, W.-C., & Ho, Y.-H. (2009). Multidimensional rasch analysis of a psychological test with multiple subtests: a statistical solution for the bandwidth-fidelity dilemma. *Educational and Psychological Measurement, 69*(3), 369-388. doi: 10.1177/0013164408323241

Chien, T.-W., Hsu, S.-Y., Tai, C., Guo, H.-R., & Su, S.-B. (2008). Using Rasch Analysis to Validate the Revised PSQI to Assess Sleep Disorders in Taiwan's Hi-tech Workers. *Community Mental Health Journal, 44*(6), 417-425. doi: 10.1007/s10597-008-9144-9

Chien, T. W., Wang, W. C., Huang, S. Y., Lai, W. P., & Chow, J. C. (2011). A web-based computerized adaptive testing (CAT) to assess patient perception in hospitalization. [Research Support, Non-U.S. Gov't]. *Journal of medical Internet research, 13*(3), e61. doi: 10.2196/jmir.1785

Chien, T. W., Wang, W. C., Lin, S. B., Lin, C. Y., Guo, H. R., & Su, S. B. (2009). KIDMAP, a web based system for gathering patients' feedback on their doctors. [Research Support, Non-U.S. Gov't]. *BMC Medical Research Methodology, 9(1)*, 38. doi: 10.1186/1471-2288-9-38

Chien, T. W., Wang, W. C., Wang, H. Y., & Lin, H. J. (2009). Online assessment of patients' views on hospital performances using Rasch model's KIDMAP diagram. [Research Support, Non-U.S. Gov't]. *BMC Health Services Research, 9(1)*, 135. doi: 10.1186/1472-6963-9-135

Chien, T. W., Wu, H. M., Wang, W. C., Castillo, R. V., & Chou, W. (2009). Reduction in patient burdens with graphical computerized adaptive testing on the ADL scale: tool development and simulation. [Comparative Study, Validation Studies]. *Health and Quality of Life Outcomes, 7(1)*, 39. doi: 10.1186/1477-7525-7-39

Damsg, rd, E., Fors, T., Anke, A., & e, C. (2008). The Tampa Scale of Kinesiophobia: A Rasch Analysis of its Properties in Subjects with Low Back and more Widespread Pain. *Journal of Rehabilitation Medicine, 39*(9), 672-678. doi: 10.2340/16501977-0125

Davis, E., Davies, B., Wolfe, R., Raadsveld, R., Heine, B., Thomason, P., . . . Graham, H. K. (2009). A randomized controlled trial of the impact of therapeutic horse riding on the quality of life, health, and function of children with cerebral palsy. *Developmental Medicine & Child Neurology, 51*(2), 111-119. doi: 10.1111/j.1469-8749.2008.03245.x

Davis, E., Mackinnon, A., & Waters, E. (2011). Parent proxy-reported quality of life for children with cerebral palsy: is it related to parental psychosocial distress? *Child: Care, Health and Development*, no-no. doi: 10.1111/j.1365-2214.2011.01267.x

Davis, E., Shelly, A. M. Y., Waters, E., & Davern, M. (2010). Measuring the quality of life of children with cerebral palsy: comparing the conceptual differences and psychometric properties of three instruments. *Developmental Medicine & Child Neurology, 52*(2), 174-180. doi: 10.1111/j.1469-8749.2009.03382.x

de Morton, N. A., Keating, J. L., & Davidson, M. (2008). Rasch Analysis of the Barthel Index in the Assessment of Hospitalized Older Patients After Admission for an Acute Medical Condition. *Archives of Physical Medicine and Rehabilitation, 89*(4), 641-647. doi: 10.1016/j.apmr.2007.10.021

Dieruf, K., Burtner, P. A., Provost, B., Phillips, J., Bernitsky-Beddingfield, A., & Sullivan, K. J. (2009). A pilot study of quality of life in children with cerebral palsy after intensive body weight-supported treadmill training. *Pediatric Physical Therapy, 21*(1), 45-52. doi: 10.1097/PEP.0b013e31818ec835

Dougherty, B. E., Nichols, J. J., & Nichols, K. K. (2011). Rasch Analysis of the Ocular Surface Disease Index (OSDI). *Investigative Ophthalmology & Visual Science, 52*(12), 8630-8635. doi: 10.1167/iovs.11-8027

Embretson, S. E., & Hershberger, S. L. (1999). *The new rules of measurement : what every psychologist and educator should know*. Mahwah, N.J.: L. Erlbaum Associates.

Fischer, A. R. H., & Frewer, L. J. (2009). Reliability of the Rasch Food Safety Practices scale. *Appetite, 53*(2), 241-244. doi: 10.1016/j.appet.2009.02.018

Fisher, R. A. (1922). The Goodness of Fit of Regression Formulae, and the Distribution of Regression Coefficients. *Journal of the Royal Statistical Society, 85*(4), 597-612.

Forster, J. (1976). *Great teachers : Burns, Shelley, Coleridge, Tennyson, Ruskin, Carlyle, Emerson, Browning*. Folcroft, Pa.: Folcroft Library Editions.

Gothwal, V. K., Wright, T. A., Lamoureux, E. L., & Pesudovs, K. (2009). Rasch analysis of the quality of life and vision function questionnaire. *Optometry and vision science : official publication of the American Academy of Optometry, 86*(7), E836-844. doi: 10.1097/OPX.0b013e3181ae1ec7

Houlihan, C. M., O'Donnell, M., Conaway, M., & Stevenson, R. D. (2004). Bodily pain and health-related quality of life in children with cerebral palsy. *Developmental Medicine & Child Neurology, 46*(5), 305-310. doi: 10.1111/j.1469-8749.2004.tb00490.x

Hoving, M. A., van Raak, E. P., Spincemaille, G. H., Palmans, L. J., Becher, J. G., & Vles, J. S. (2009). Efficacy of intrathecal baclofen therapy in children with intractable spastic cerebral palsy: a randomised controlled trial. *European Journal of Neurology, 13*(3), 240-246. doi: 10.1016/j.ejpn.2008.04.013

Jang, E. E., & Roussos, L. (2007). An Investigation into the Dimensionality of TOEFL Using Conditional Covariance-Based Nonparametric Approach. *Journal of Educational Measurement, 44*(1), 1-21. doi: 10.1111/j.1745-3984.2007.00024.x

Levitt, S. (2010). *Treatment of Cerebral Palsy and Motor Delay*. West Sussex: John Wiley & Sons.

Linacre, J. M. (1998). Detecting multidimensionality: which residual data-type works best? *Journal of outcome measurement, 2*(3), 266-283.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106.

Marco, G. L. (1977). Item characteristic curve solutions to three intractable testing problems 1. *Journal of Educational Measurement, 14*(2), 139-160. doi: 10.1111/j.1745-3984.1977.tb00033.x

McCloy, R. A., & Gibby, R. E. (2011). Computerized Adaptive Testing *Technology-Enhanced Assessment of Talent* (pp. 153-189): Jossey-Bass.

Mitchell, A. J., Smith, A. B., Al-salihy, Z., Rahim, T. A., Mahmud, M. Q., & Muhyaldin, A. S. (2011). Redefining diagnostic symptoms of depression using Rasch analysis: testing an item bank suitable for DSM-V and computer adaptive testing. *Australian and New Zealand Journal of Psychiatry, 45*(10), 846-852. doi: 10.3109/00048674.2011.596477

Odding, E., Roebroeck, M. E., & Stam, H. J. (2006). The epidemiology of cerebral palsy: incidence, impairments and risk factors. *Disability and Rehabilitation, 28*(4), 183-191. doi: 10.1080/09638280500158422

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine & Child Neurology, 39*(4), 214-223. doi: 10.1111/j.1469-8749.1997.tb07414.x

Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology, 46*(1), 1-18. doi: 10.1348/014466506x96931

Pesudovs, K. (2010). Item banking: a generational change in patient-reported outcome measurement. [Review]. *Optometry and vision science : official publication of the American Academy of Optometry, 87*(4), 285-293. doi: 10.1097/OPX.0b013e3181d408d7

Prieto, L., Alonso, J., & Lamarca, R. (2003). Classical Test Theory versus Rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes, 1*, 27. doi: 10.1186/1477-7525-1-27

Ryser, L., Wright, B. D., Aeschlimann, A., Mariacher-Gehler, S., & Stucki, G. (1999). A new look at the Western Ontario and McMaster Universities Osteoarthritis Index using

Rasch analysis. *Arthritis Care & Research, 12*(5), 331-335. doi: 10.1002/1529-0131(199910)12:5<331::aid-art4>3.0.co;2-w

Smith, A. B., Wright, E. P., Rush, R., Stark, D. P., Velikova, G., & Selby, P. J. (2006). Rasch analysis of the dimensional structure of the Hospital Anxiety and Depression Scale. *Psycho-Oncology, 15*(9), 817-827. doi: 10.1002/pon.1015

Stanley, F., & Blair, E. (2000). *Cerebral palsies: epidemiology and causal pathways*. Cambridge: Mac Keith.

Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford University Press.

Swiggum, M., Hamilton, M. L., Gleeson, P., & Roddey, T. (2010). Pain in Children with Cerebral Palsy: Implications for Pediatric Physical Therapy. *Pediatric Physical Therapy, 22*(1), 86-92 10.1097/PEP.1090b1013e3181cd1018a1097.

Tang, W. K., Wong, E., Chiu, H. F. K., & Ungvari, G. S. (2007). Rasch analysis of the scoring scheme of the HADS Depression subscale in Chinese stroke patients. *Psychiatry Research, 150*(1), 97-103. doi: 10.1016/j.psychres.2006.01.015

Tecklin, J. S. (2008). *Pediatric physical therapy* (4 ed.). Philadelphia: Lippincott Williams & Wilkins.

Tennant, A., McKenna, S. P., & Hagell, P. (2004). Application of Rasch analysis in the development and application of quality of life instruments. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research, 7 Suppl 1*, S22-26. doi: 10.1111/j.1524-4733.2004.7s106.x

van der Linden, W. J., Ariel, A., & Veldkamp, B. P. (2006). Assembling a Computerized Adaptive Testing Item Pool as a Set of Linear Tests. *Journal of Educational and Behavioral Statistics, 31*(1), 81-99. doi: 10.3102/10769986031001081

Van der Linden, W. J., & Hambleton, R. K. (1997). Item response theory: Brief history, common models, and extensions. *Handbook of modern item response theory*, 1-28.

van der Linden, W. J., & Veldkamp, B. P. (2004). Constraining Item Exposure in Computerized Adaptive Testing With Shadow Tests. *Journal of Educational and Behavioral Statistics, 29*(3), 273-291. doi: 10.3102/10769986029003273

Waclawski, E. R. (2010). Health Measurement Scales—A Practical Guide to Their Development and Use. *Occupational Medicine, 60*(2), 156. doi: 10.1093/occmed/kqp179

Wainer, H., & Wang, X. (2000). Using a New Statistical Model for Testlets to Score TOEFL. *Journal of Educational Measurement, 37*(3), 203-220. doi: 10.1111/j.1745-3984.2000.tb01083.x

Wake, M., Ba, L. S., & Reddihough, D. (2003). Health status of Australian children with mild to severe cerebral palsy: cross-sectional survey using the Child Health

Questionnaire. *Developmental Medicine & Child Neurology, 45*(3), 194-199. doi: 10.1111/j.1469-8749.2003.tb00930.x

Wang, H. Y., Cheng, C. C., Hung, J. W., Ju, Y. H., Lin, J. H., & Lo, S. K. (2010). Validating the Cerebral Palsy Quality of Life for Children (CP QOL-Child) questionnaire for use in Chinese populations. *Neuropsychological Rehabilitation, 20*(6), 883-898. doi: 10.1080/09602011.2010.509190

Wang, W. C., Chen, P. H., & Cheng, Y. Y. (2004). Improving measurement precision of test batteries using multidimensional item response models. *Psychological Methods, 9*(1), 116-136. doi: 10.1037/1082-989X.9.1.116

Ware Jr, J. E., Kosinski, M., Bjorner, J. B., Bayliss, M. S., Batenhorst, A., Dahlöf, C. G. H., . . . Dowson, A. (2003). Applications of computerized adaptive testing (CAT) to the assessment of headache impact. *Quality of Life Research, 12*(8), 935-952. doi: 10.1023/a:1026115230284

Waters, E., Davis, E., Mackinnon, A., Boyd, R., Graham, H. K., Kai Lo, S., . . . Reddihough, D. (2007). Psychometric properties of the quality of life questionnaire for children with CP. *Developmental Medicine & Child Neurology, 49*(1), 49-55. doi: 10.1017/S0012162207000126.x

Waters, E., Maher, E., Salmon, L., Reddihough, D., & Boyd, R. (2005). Development of a condition-specific measure of quality of life for children with cerebral palsy: empirical thematic data reported by parents and children. *Child: Care, Health and Development, 31*(2), 127-135. doi: 10.1111/j.1365-2214.2004.00476.x

Weiss, D. J., & Bock, R. D. (1983). *New horizons in testing : latent trait test theory and computerized adaptive testing*. New York: Academic Press.

World Health Organization. (1993). *Measuring quality of life: the development of the World Health Organization Quality of Life Instrument (WHOQOL)*. Geneva: Division of Mental Health, World Health Organization.

Wright, B. D. (1996). Comparing Rasch measurement and factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 3*(1), 3-24. doi: 10.1080/10705519609540026

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: Mesa Press.

Wright, B. D., & Mok, M. (2000). Rasch models overview. *Journal of applied measurement, 1*(1), 83-106.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: Mesa Press.

Yao, L., & Boughton, K. A. (2007). A Multidimensional Item Response Modeling Approach for Improving Subscale Proficiency Estimation and Classification. *Applied Psychological Measurement, 31*(2), 83-105. doi: 10.1177/0146621606291559