International Journal of School & Educational Psychology, *3*, 278–292, 2015
Copyright © International School Psychology Association
ISSN 2168-3603 print/ISSN 2168-3611 online
DOI: 10.1080/21683603.2015.1067873

Routledge
Taylor & Francis Group

Trinidad and Tobago National Standardization of the Adjustment Scales for Children and Adolescents

Paul A. McDermott
Graduate School of Education, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Marley W. Watkins
Department of Educational Psychology, Baylor University, Waco, Texas, USA

Anna M. Rhoad and Jessica L. Chao
Graduate School of Education, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Frank C. Worrell
Graduate School of Education, University of California, Berkeley, California, USA

Tracey E. Hall
Center for Applied Special Technology, Wakefield, Massachusetts, USA

Given relevant cultural distinctions across nations, it is important to determine the dimensional structure and normative characteristics of psychological assessment devices in each focal population. This article examines the national standardization and validation of the Adjustment Scales for Children and Adolescents (ASCA) with a nationally representative sample of Trinidad and Tobago schoolchildren ($N = 900$). ASCA is a 156-item teacher rating scale that measures sociobehavioral adjustment. Results from exploratory and confirmatory analyses yielded the same Overactivity and Underactivity dimensions observed in international samples. The dimensions were scaled using IRT and Bayesian scoring, with scores evincing expected moderate to strong relationships with other teacher observations and weaker relationships with parent observations and reading achievement. Population performance trends are explored and implications are discussed.

Keywords: Behavioral and social adjustment, psychopathology, item response theory, Trinidad and Tobago

Child and adolescent socioemotional health is an international concern. Worldwide, up to 20% of children and adolescents experience notable emotional or behavioral distress, with suicide as the third leading cause of adolescent death (World Health Organization [WHO], 2005). Although children with emotional and behavioral disorders are often first seen in the educational system (Burns et al., 1995), school-based personnel and services are inadequate in all but a few high-income countries (WHO, 2005). Worldwide, it has been estimated that at least half of child mental health training needs are unmet (WHO, 2005). As a result, the majority of youth with serious disorders go unrecognized and untreated (Morris et al., 2011). For example, the treated prevalence of children and adolescents internationally is 159 per 100,000 population compared to 664 per 100,000 for the adult population (Morris et al., 2011). Poor social-emotional adjustment in youth is linked to educational failure, substance abuse, violence, and other health problems, whereas social competency is related to greater well-being and higher academic achievement (Eisenberg, Fabes, & Spinrad, 2006; Guerra & Bradshaw, 2008; Masten

Correspondence should be addressed to Paul A. McDermott, Graduate School of Education, Quantitative Methods Division, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104-6216, USA. E-mail: drpaul4@verizon.net

& Coatsworth, 1998; Patel, Flisher, Hetrick, & McGorry, 2007; Weissberg & Greenberg, 1998).

Of course, children and adolescents with emotional and behavioral disorders must first be identified before they can be treated (United Nations Children's Fund, 2013). Identification is formally accomplished through assessment that includes a variety of methods, including interviews, observations, cognitive and academic tests, and behavior rating scales (McConaughy & Ritter, 2014). Behavior rating scales have become especially important because of their convenience, scope of coverage, naturalistic foundation, efficiency, ecological validity, standardization, and norms (Barry, Frick, & Kamphaus, 2013; Dowdy, Twyford, & Sharkey, 2013; Merrell, 2008).

Many behavior rating scales are based on the theory that all problem behavior can be reduced to two or more broadband behavioral dimensions drawn from Eysenck's (1953) extraversion-introversion versus neuroticism dichotomy (Kohn, 1977; Peterson, 1961; Rutter, 1967). Although variously labeled personality problems versus behavior problems (Peterson, 1961), overactivity versus underactivity (Stott, 1979), and undercontrolled versus overcontrolled (Achenbach & Edelbrock, 1978), usage of an internalizing versus externalizing problems designation (Achenbach, 1966) has become common (Merrell, 2008). The internalizing dimension is characterized by problems with inhibited and shy-anxious behavior and is related to disorders such as depression, anxiety, phobias, and panic (Markon, 2010). In contrast, the externalizing dimension is distinguished by disinhibition and acting out behaviors, and is linked to substance abuse and dependence, antisocial personality disorder, and conduct disorder.

Latent variable modeling procedures have shown that the two hierarchical constructs of internalizing and externalizing underlie many of the common *Diagnostic and Statistical Manuals of Mental Disorders* (DSM; American Psychological Association [APA], 1987) diagnoses (Wright et al., 2013). The internalizing-externalizing (IE) model has proven to be robust across age, sex, ethnicity, culture, informant type, instrument, and DSM Axes (Achenbach, 1966; De Clercq, De Fruyt, Van Leeuwen, & Mervielde, 2006; Eaton, Krueger, & Oltmanns, 2011; Forbush & Watson, 2013; Kramer, Krueger, & Hicks, 2008; Krueger, Capsi, Moffitt, & Silva, 1998; Krueger, Chentsova-Dutton, Markon, Goldberg, & Ormel, 2003; Lahey et al., 2008; Rescorla et al., 2011; Slade & Watson, 2006; Slobodskaya, 2014; van der Ende, Verhulst, & Tiemeier, 2012; Wright et al., 2013). In contrast, there has been variability when narrower dimensions of psychopathology have been hypothesized (Gomez & Vance, 2014; Goodman, Lamping, & Ploubidis, 2010; McConaughy & Ritter, 2014; Pendergast et al., 2014; Van Meter et al., 2014).

Given the robustness of the internalizing and externalizing constructs, they have been incorporated into many behavior rating scales (Dowdy et al., 2013; Merrell, 2008) and modern models of mental health promulgated by the WHO (2007) promote widespread application of behavior rating scales (Carlson, Benson, & Oakland, 2010). Unfortunately, most of those behavior rating scales have been standardized and normed in Western industrialized nations and few are available in other regions of the world (Mpofu, Oakland, Ntinda, Seeco, & Maree, 2014). International surveys have found that tests are frequently imported from other countries without appropriate normative data and absent requisite evidence of reliability and validity (Hu & Oakland, 1991; Oakland & Hu, 1993). These trends are especially prevalent outside Australia, Canada, Western Europe, and the United States and among developing countries (Oakland, 2004; Oakland, Wechsler, & Maree, 2013).

Cultural differences among nations and people make it important to normatively standardize instruments that are germane to psychological assessment in each population of interest (United Nations Children's Fund, 2013). It is a common misperception that psychometric properties (such as validity and reliability at the item and scale level) are preserved regardless of where and with whom an instrument is used (van Widenfelt, Treffers, de Beurs, Siebelink, & Koudijs, 2005). Moreover, not all items and constructs are universally meaningful (Hambleton & Patsula, 1998) and cultural norms can influence the respondent's judgment of the acceptability of different types of behaviors, interactions, and relationships (Rubin, 1998; United Nations Children's Fund, 2013). Emotional distress among children and adolescents appears to be universal, but the particulars of cultural context lead to variations of expression and necessitate examination of instruments for use in each nation (van Widenfelt et al., 2005).

Like other developing nations, the Republic of Trinidad and Tobago has no nationally normed and psychometrically sound behavior rating scales, although its high crime rates (Greenberg & Agozino, 2012) as well as economic disparities, harsh parental disciplinary practices, and other social challenges (Cappa & Khan, 2011; Krishnakumar, Narine, Roopnarine, & Logie, 2014; Roopnarine, Krishnakumar, Narine, Logie, & Lape, 2014; Williams, 2013) reflect a compelling need for assessment and intervention to improve child and adolescent socioemotional health. Well aware of those needs, Trinidad and Tobago has teamed with the United Nations and other global organizations to improve social, economic, and human conditions (United Nations, 2014). Education has been recognized as especially important and the Ministry of Education has undertaken efforts to identify and support children and adolescents at risk for academic and behavior problems (Johnstone, 2010; Watkins, Hall, & Worrell, 2014). Without the economic resources and human capital required for extensive test development (Oakland et al., 2013), the Ministry of Education elected to standardize and validate the Adjustment Scales for Children and Adolescents (ASCA;

McDermott, Stott, & Marston, 1993) in Trinidad and Tobago for the purpose of identifying students with sociobehavioral problems and determining the prevalence of maladjusted students (Watkins et al., 2014).

The ASCA is comprised of 156 descriptions of behavior in 29 social, recreational, or learning situations in which a teacher observes a youth's adjustment to authority, peers, other youth, and various tasks. Instead of relying on estimates of frequency or severity of behaviors, respondents identify specific behaviors with the severity of the behavioral disturbance determined by the pervasiveness of behaviors across situations. A national standardization and validation of ASCA was conducted in the United States with a norm sample of 1,400 students aged 5 through 17 years (McDermott, 1993; McDermott, Steinberg, & Angelo, 2005). The sample was stratified according to the U.S. Census by age, gender, academic level, ethnicity, disability, region, community size, and parent education. Exploratory and confirmatory components analyses revealed six core syndromes and two second-order factors that are generalizable across age, gender, and ethnicity. Core syndromes include attention-deficit hyperactive, solitary aggressive (provocative), solitary aggressive (impulsive), oppositional defiant, diffident, and avoidant, while second-order factors are overactivity and underactivity. The overactivity and underactivity dimensions have been found in other populations including Hispanic/Latino, Native American, and Canadian youth (Canivez & Beran, 2009; Canivez & Bohan, 2006; Canivez & Sprouls, 2005, 2010).

The purpose of this study was to examine the normative development, dimensionality, and validation of ASCA for national application in Trinidad and Tobago. It follows a pilot study assessing the construct validity of ASCA in Trinidad and Tobago (George, McDermott, Watkins, Worrell, & Hall, 2012). Exploratory and confirmatory components analyses are conducted to examine the dimensional structure. The prevalence of problem behaviors is examined at the item level and scaled scores are estimated using item response theory (IRT). Product–moment correlations are applied to determine the direction and magnitude of relationships between scores on each ASCA dimension and external criterion variable, and relationships are further assessed using hierarchical linear modeling.

METHOD

Setting

Trinidad and Tobago is the southernmost country in the Caribbean chain of islands. It encompasses the twin islands of Trinidad and Tobago. Postcolonial and English-speaking, the country gained independence from Britain in 1962 but remains a Commonwealth country. The population is around 1.3 million with approximately 34% of African ancestry, 35% of East Indian ancestry, 24% of mixed ancestry, and 7% other ancestry (Central Intelligence Agency, 2014). Education in Trinidad and Tobago is free and compulsory between the ages of 5 and 16 years. Order and discipline are the prevailing forms of parent-child interaction (Barrow, 2008). Rather than encouraging play, parents tend to expect children to assist in the home and often to care for siblings by the age of 5 (Barrow, 2008). Corporal punishment in homes is common and accepted as a cultural norm, where respect for authority is a universally emphasized value for children (Barrow, 2008; Cappa & Khan, 2011; Gopaul-McNicol, 1999; Roopnarine et al., 2014). Discipline becomes more emphasized as children approach school age, with disobedience attributed to parental leniency (Barrow, 2008; Gopaul-McNicol, 1993).

Sample and Participants

Participants were children aged 4–15 years ($M = 8$, $SD = 2$) drawn from government and assisted elementary schools nationwide. The national normative sample ($N = 700$) was blocked to be representative by grade and gender, and a supplemental validity sample was drawn by oversampling ($n = 200$), for a full sample of 900. The sample was 50.3% female and 49.7% male, with 39.9% of African descent, 38.3% East Indian descent, and 21.7% mixed descent.

Instruments

Classroom sociobehavioral adjustment

ASCA is a behaviorally based teacher rating scale consisting of 156 items describing positive and problem behavior in relation to 29 classroom situations. Classroom situations examples include playing fairly, getting along with peers, being truthful to the teacher, seeking teacher help, taking part in team games, and coping with new learning tasks. The teacher indicates a student's observed behavior over the past two months by marking any behavioral description pertinent to a given situation. More than one behavior can be chosen to describe the child within each social, play, or learning context. For example, the coping with new learning tasks context includes items such as, "Has a happy-go-lucky attitude to every problem," "Charges in without taking time to think or follow instructions," "Approaches a new task with caution but gives it a try," "Won't even attempt it if she senses a difficulty," "Likes the challenge of something difficult," and "Cannot work up the energy to face anything new."

The 29 positive items (prevalence greater than or equal to 50%) were included to reduce response bias by allowing teachers to identify children's behavioral strengths. In previous research, teachers found it easier to respond to item sets that included positive behaviors (McDermott,

1993). Examples of the ASCA score reliability and validity are documented in the instrument manual (McDermott, 1994), as well as in numerous studies. These include studies supporting the instrument's internal consistency (Canivez, 2004, 2006; Canivez & Bohan, 2006; McDermott, 1993, 1994), interrater agreement (Canivez & Watkins, 2002; Canivez, Watkins, & Schaefer, 2002; Watkins & Canivez, 1997), and short-term stability (Canivez, Perry, & Weller, 2001; McDermott, 1993, 1994). There is substantial evidence of convergent and divergent/discriminant validity (Canivez & Bordenkircher, 2002; Canivez, Neitzel, & Martin, 2005; Canivez & Rains, 2002; McDermott, 1994; McDermott et al., 1995), while factorial validity in other populations has been demonstrated (Canivez & Beran, 2009; Canivez & Bohan, 2006; Canivez & Sprouls, 2010; George et al., 2012).

Classroom learning behavior

The Learning Behaviors Scale (LBS; McDermott, 1999) is a teacher rating scale containing 29 items, including both negative and positive learning behaviors, and yielding both a total score and four subscale scores (Competence Motivation, Attitude Toward Learning, Attention-Persistence, Strategy/Flexibility). Teachers who have observed the child for at least 50 days rate the manifestation of each behavior on a three-point Likert scale. The LBS was standardized on 1,500 students aged 5–17 years, and blocked for age, gender, and grade level. The LBS was found to have a factor structure that is invariant across gender, age, and ethnic group. Studies have documented score internal consistency and interrater reliability (McDermott, 1999; Worrell, Vandiver, & Watkins, 2001), and structural validity (McDermott, 1999; Worrell et al., 2001). Convergent and divergent validity for the LBS have been documented with the Differential Ability Scales (Elliott, 1990), ASCA (McDermott et al., 1993) and three subtests from the Basic Achievement Skills Individual Screener (Psychological Corporation, 1983).

Classroom clinical behavior

The Disruptive Behavior Disorders Rating Scale (DBDRS; Pelham, Gnagy, Greenslade, & Milich, 1992) is a teacher rating scale that aids in classifying clinical disorders based on the 36 diagnostic criteria from the three disruptive behavior categories (Attention Deficit Hyperactive Disorder, Oppositional-Defiant Disorder, Conduct Disorder) described in the DSM-III-R (American Psychiatric Association, 1987). The scale is comprised of three factors: oppositional/defiant, inattention, and impulsivity/overactivity, with respective coefficient α's of .96, .95, and .95 (Pelham et al., 1992). Research with this scale has typically focused on males in regular and special education classrooms (Pelham, Evans, Gnagy, & Greenslade, 1992) and demonstrates adequate score stability and validity.

Home socioemotional behavior

The Adjustment Scales for Children and Adolescents-Home Edition (ASCA-H; Watkins & McDermott, 2002) is a behaviorally based parent rating scale containing 202 items in 34 situations. Similar to ASCA, it presents behaviors in a situational context, but the items are related to behaviors observable in the home. Situational contexts include parental correction, relationship with other adults, peers, care of belongings, chores, meal times, unorganized activities, homework, and so forth. The parent indicates a child's behavior by marking any behavioral description applied to a situation that they have observed over the past two months.

Although the ASCA-H is still in development, preliminary evidence of its structural validity and reliability has been provided by two pilot studies. The first study was conducted with a sample of 314 children aged 5 to 17 years from the mid-Atlantic region of the United States. Three replicable factors (Unsocialized, Avoidant, and Restless-Impulsive) were found that demonstrated internal consistency ($r = .85$, .79, and .84) and stability across a 4-week interval ($r = .65$, .74, and .86) as well as convergent and divergent validity with other parent rating scales (Mordell, 2001). The second pilot study was conducted with 426 children aged 5 to 14 years in the mid-Atlantic region of the United States. Four first-order factors (Aggressive-Oppositional, Attention-Seeking Impulsive, Detached, and Diffident) with internal consistency reliability coefficients ranging from .65 to .92 emerged (Coffey, 2006). Internal consistency reliability coefficients of the three factors (ADH, Conduct Problems, and Overactivity) derived from the Trinidad and Tobago sample were .82, .77, and .74 respectively.

Academic achievement

Oral Reading Fluency (ORF; Fuchs, Fuchs, Hosp, & Jenkins, 2001) is a curriculum-based academic measure that assesses reading fluency by way of the number of words a child can read correctly in one minute. Two ORF passages that ranged from 153 to 321 words in length were created for each grade from local grade-level reading texts. The ORF score for each participant was the average of the number of words read correctly on those two grade-appropriate passages. Typically, the growth curve for fluency is steeper in the primary grades and negatively accelerates thereafter (Fuchs et al., 2001). That pattern was repeated in this sample: There was an average increase in ORF scores of 43% for each of the first two years assessed and an average increase of 12% for each of the final two years assessed. Correlations between ORF and other curriculum-based measures usually indicate that it performs appropriately, with much of recent research focused on predictive validity and clinical utility (Reschly, Busch, Betts, Deno, & Long, 2009). For the current sample, ORF scores correlated .45

with a measure of phonological awareness (Watkins & Edwards, 2004) among kindergarten-equivalent students and .73 with a cloze comprehension measure (Deno, Mirkin, & Chiang, 1982; McKenna & Layton, 1990) among students in standard 1 and standard 2 grades.

Procedure

A list of all government and assisted elementary schools—but excluding special schools—was used to identify a representative sample of students from 79 elementary schools in Trinidad and Tobago, stratified by the regional enrollment of the school-aged population. Thus, St. George West, with the largest school enrollment, was represented in the sample by 19 schools, whereas the Nariva/Mayaro region was represented in the sample by 3 schools. One classroom at each grade level at each school site was randomly selected to participate, and two students (one male and one female) were randomly selected from each identified classroom. All 700 students from the national normative sample were used for scale calibration and the validity oversample of 200 students was used to supplement structural and validity analyses. In total, 524 teachers at 75 schools completed ASCA forms.

Data were collected over one academic year by Guidance and Special Education Officers (GSEOs) from the Ministry of Education as part of a project between a consulting team based at Pennsylvania State University and the Ministry of Education (Watkins et al., 2014). All Trinidad officers possessed a university degree and received training from the consulting team, and most were assigned to gather data in the educational division in which they already worked. GSEOs were paid an honorarium for each school for which they gathered complete data, and teachers and parents also received an honorarium for completing the rating scales.

As ASCA includes positive behavior items even though its central focus is behavioral problems, 29 positive behavior items were identified and excluded from subsequent analysis, reducing the number of items in the analysis from 156 to 127. This process also averted the likelihood of difficulty, valence, and bipolar factors that would tend to emerge in the presence of disparately distributed dichotomous items (Bernstein & Teng, 1989).

Exploratory analysis

The full sample was randomly partitioned into an exploratory subsample ($n = 500$) and confirmatory subsample ($n = 400$). The EFA subsample was 50.4% female and 49.6% male, with 40.61% of African descent, 36.9% East Indian descent, and 22.4% mixed descent. Demographic characteristics of the CFA subsample were similar with 48.8% female, 51.3% male, 39.1% African, 40.1% East Indian, and 20.8% Mixed. MicroFACT (Waller, 2001) software was used to generate a smoothed tetrachoric correlation matrix, applying two-stage maximum-likelihood estimation (Olsson, 1979) and least-squares approximation of the original matrix (Knol & Berger, 1991). The matrix of 127 problem behavior items was smoothed for positive semidefiniteness. Due to the large number of items and their dichotomous nature and expected extreme positive skews, it was not feasible to produce a nonsingular matrix; thus, iterated components analysis was applied as recommended by Debelak and Tran (2013). Specifically, structures produced by components analysis with large numbers of items will tend to approximate those produced in common factoring (Snook & Gorsuch, 1989) and correlation matrices for components analysis need not be nonsingular. Minimum average partialing (MAP; Velicer, 1976) was used with the smoothed matrix to suggest the maximum number of retained components. Adhering to the recommendation by Snook and Gorsuch (1989) for scales with 50 or more items, principal components solutions were rotated toward simple structure via varimax, equamax, and promax criteria. Solution criteria included: (a) approximate simple structure as reflected by a maximized hyperplane count (Yates, 1987) and item coverage, (b) at least four salient items per component with loadings $\geq .40$ defined as salient, (c) reliable components (i.e., $\alpha \geq .70$), and (d) a theoretically sensible structure with parsimonious coverage of the data and concordance with leading area research (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Confirmatory analysis

It was anticipated that confirmatory factor analysis (CFA) involving a very large number of items would make it difficult, if not impossible, to successfully apply structural equations modeling (SEM) with item-level data. Parameters would need to be estimated simultaneously for a large number of highly skewed binary items, exceeding the capacity of current SEM estimation procedures (McDermott, Watkins, Rovine, & Rikoon, 2013). Additionally, the many small cross-loadings would degrade model fit and inflate factor intercorrelations (Hsu, Skidmore, Li, & Thompson, 2014). Researchers have noted the difficulty that SEM faces when attempting to minimize the discrepancies between the observed and predicted covariance matrix, the low reliability of correlations, correlated errors, and binary and highly-skewed item data (Hall, Snell, & Foust, 1999; Hau & Marsh, 2004; Kishton & Widaman, 1994; Nasser & Wisenbaker, 2003).

Alternatively, researchers have recommended the creation of item parcels (Bandalos, 2002; Hall et al., 1999; Sass & Smith, 2006; Thompson & Melancon, 1996; Wilkinson, 2007). A few items were assigned to each parcel and the parcels analyzed in CFA. Problem behavior data focuses on rare behaviors and these item data are inherently positively skewed and leptokurtic. The distributional

balance parceling method is advantageous in this situation and results in more normally distributed variables (Bandalos, 2002; Hau & Marsh, 2004; Nasser & Wisenbaker, 2003; Thompson & Melancon, 1996). Unsuitable item parcels can confound sources of error variance and result in misspecification of CFA models (Bandalos, 2002), whereas properly applied item parcels serve to increase the likelihood of normal distributions and reduce sampling error (Wilkinson, 2007) and enable otherwise infeasible estimation processes. Accordingly, each parcel contained items with high prevalences, items with moderate prevalences, and items with lower prevalences. Parcels based on the salient items from the exploratory components analytic solution were submitted to maximum-likelihood estimation under the Satorra–Bentler scaled difference chi-square for nonnormal data (Satorra & Bentler, 2001), with acceptable fit indicated by a Root Mean Squared Error of Approximation (RMSEA) $\leq$ .08 and Comparative Fit Index (CFI) $\geq$ .90 (Marsh, Liem, Martin, Morin, & Nagengast, 2011).

Scaling

The items associated with each respective dimension were scaled through multiple-group IRT, applying the one- and two-parameter logistic models based on the national normative sample ($N = 700$). The national normative sample was used for calibration purposes in order to yield representative parameters. Parameters were thereafter used to generate scores for the validation oversample ($n = 200$). Scores were estimated using the Bayesian Expected a Posteriori (EAP) method, with the population mean and standard deviation for the normative sample centered at $M = 50$, $SD = 10$. Reliability was assessed for those marker items comprising each dimension using Cronbach's alpha coefficient. Reliability was further examined through overplots displaying the distribution of test information (i.e., the inverse of measurement error) and measurement error for each derived ASCA scale.

External validity

Product–moment correlations were assessed to determine the direction and magnitude of relations between scores on each ASCA dimension and external criterion variables. Because the data were nested within teachers, following the recommendations by Waterman, McDermott, Fantuzzo, and Gadsden (2012), relationships were also estimated using hierarchical linear modeling (HLM), where each ASCA dimension served as the group-mean centered predictor in a two-level conditional model, indicating the percentage of between-children within-teacher variance accounted for by respective ASCA dimension scores.

RESULTS

Dimensionality

MAP for 127 problem behavior items suggested that up to 8 components might be extracted from the smoothed tetrachoric matrix. The 1- through 8-component models were tested against the stated criteria. The 2-component, promax-rotated ($k = 3$) model emerged as the optimal solution, where Waller's (2001) Goodness-of-Fit Index = .88 and the Root Mean Squared Residual = .09. Models extracting more than two components contained underidentified and unreliable dimensions and the 1-component model compressed the 2-component model into an uninterpretable composite dimension. Per Comrey's (1988) recommendations, 5 items providing multiple salient loadings (thus refining simple structure and reducing interfactor correlations) and 16 items yielding item-total scale correlations $<$ .20 (thus suppressing internal consistency and discrimination) were excluded from further analysis. The remaining 69 items were retained. Table 1 presents rotated pattern loadings, final communalities, item-scale correlations, and prevalence (in the national normative sample). Coefficient alpha for each scale is also posted (see the centered headings). Based on item content and patterns of descending loadings, the scales were named Overactivity (50 items; M behavioral prevalence = 9.4%) and Underactivity (19 items; M prevalence = 7.0%). Underactivity corresponds to internalizing problems and Overactivity to externalizing problems, as previously discussed. There was a weak correlation between Overactivity and Underactivity scores, $r = .13$, $p < .001$.

The two-dimensional structure was replicated for the confirmatory subsample, as represented by 15 quadruple and 3 triplet item parcels. Model fit was adequate, where Satorra–Bentler $\chi 2$ (134) = 234.13, CFI = .912, and RMSEA = .043 (90% CI = .034/.052).

Scaling and Reliability

The log-likelihood deviance test indicated that the two-parameter logistic was a superior fit to the one-parameter logistic model ($p < .0001$) for each scale. For Overactivity, the threshold parameters ranged 0.75– 3.23 ($M = 1.93$, $SD = 0.54$), slopes ranged 0.46– 2.68 ($M = 1.22$, $SD = 0.42$), average information = 11.36, and the approximate maximum information = 51.54 at $\theta = 1.75$, while for Underactivity the thresholds ranged from 1.61 to 3.21 ($M = 2.27$, $SD = .42$), slopes from 0.70 to 1.38 ($M = 1.02$, $SD = 0.23$), average information = 2.67, and maximum information = 13.17 at $\theta = 2.25$.

EAP (Thissen, Pommerich, Billeaud, & Williams, 1995) scaled scores (*SS*s) were estimated for members of the normative sample. Based on normative *SS*s and measurement error, the IRT reliability index for Overactivity was .92

TABLE 1
Dimensional Structure and Properties of the Adjustment Scales for Children and Adolescents

Item description[a]	*Scale pattern loadings*[b]				
	I	*II*	*Communality*	*Item/scale r*[c]	*%Prevalence*[d]
Scale I: Overactivity (coefficient α = .92[e])					
Disrupts team games by fooling around	*.84*	−.09	.66	.63	6.7
Does things in front of teacher	*.78*	−.10	.57	.57	9.1
Much too talkative with teacher	*.76*	−.25	.52	.53	13.4
Often cause of trouble in line	*.75*	.03	.59	.58	4.9
Talks, gazes, plays during schoolwork	*.72*	−.07	.49	.57	26.6
Disturbs others' fun at play	*.72*	.01	.52	.51	4.9
Snatches objects away from others	*.71*	−.04	.49	.51	5.4
Constantly restless, shifts, raps	*.70*	−.17	.44	.51	20.4
Wants to dominate and have own way at play	*.69*	−.19	.43	.50	13.0
Attacks others viciously if provoked	*.69*	.02	.49	.53	5.0
Doesn't stay in seat when should	*.66*	−.19	.39	.45	15.3
Poor loser, causes disturbances	*.66*	−.04	.42	.48	6.1
Quarrels, provokes others	*.66*	.00	.44	.51	6.7
Takes things from desks or lockers	*.64*	.02	.42	.48	8.9
Makes sexually offensive gestures/remarks	*.64*	.05	.43	.42	3.1
Associates with troublesome youth	*.63*	.00	.40	.52	7.3
Misbehaves when teacher attends others	*.62*	−.18	.34	.45	30.1
Answers before taking time to think	*.61*	−.28	.35	.38	20.3
Tries to dominate agemates	*.61*	−.07	.35	.44	8.7
Loses temper if can't get way	*.60*	−.06	.34	.43	9.3
Uses devices to gain teacher's attention	*.59*	−.16	.31	.40	13.7
Starts fights and rough play	*.59*	.18	.45	.44	3.9
Uses bad language that offends others	*.59*	.00	.35	.43	6.4
Doesn't hesitate to lie	*.58*	.17	.43	.39	3.4
Overly rough with smaller or weaker children	*.58*	.12	.40	.43	5.7
Helps teach unless in a bad mood	*.57*	.02	.33	.37	5.0
Clowns around, plays silly tricks	*.57*	−.14	.30	.40	8.4
Pushes in front of others in line	*.57*	−.16	.30	.41	14.9
Occasionally lies to avoid blame	*.56*	−.07	.29	.42	24.9
Takes correction badly, muttering	*.54*	.07	.31	.43	11.9
Destroys or defaces own books, etc.	*.54*	.10	.34	.45	10.7
Has stolen from other pupils	*.54*	.18	.38	.38	3.3
Deliberately destroyed others' belongings	*.53*	.18	.37	.33	2.4
Plays around when working with hands	*.52*	−.07	.25	.36	16.6
Made unprovoked attacks on other students	*.52*	.07	.30	.35	4.9
Inclined to cheat at play	*.50*	.05	.27	.35	6.9
Only works when watched	*.49*	.18	.32	.37	6.6
Tells tall tales about self/family	*.47*	.03	.23	.28	4.6
Answers aggressively to corrections	*.47*	.23	.34	.29	1.6
Throws or sweeps objects with no reason	*.47*	.16	.29	.29	3.9
Sometimes unfriendly to teacher	*.46*	.14	.27	.34	5.7
Charges in new learning tasks without thinking	*.46*	−.09	.20	.33	14.3
Asks for jobs but doesn't finish	*.45*	.17	.28	.33	6.3
Unusual sitting positions, climbs on desk	*.45*	.20	.29	.26	3.1
Unkind to weaker children	*.43*	.32	.37	.31	2.9
Improves after correction, doesn't last	*.42*	.07	.20	.37	23.9
Makes sudden inappropriate noises	*.42*	−.05	.17	.28	5.6
Rushes around shouting madly	*.42*	.31	.35	.25	1.4
Welcomes teacher loudly	*.41*	−.25	.17	.22	19.3
Can't keep a friend for long	*.41*	.20	.26	.26	2.0
Scale II: Underactivity (coefficient α = .74[e])					
Sits so quietly don't know if attending you	−.38	*.66*	.42	.45	11.1
Too unenergetic to be troublesome	.00	*.63*	.39	.23	1.3
Sits meekly, seems afraid to budge	−.35	*.63*	.38	.42	8.3
Too withdrawn to help teacher	−.32	*.62*	.36	.27	4.0

(*continued*)

TABLE 1 – (*Continued*)

Item description[a]	*Scale pattern loadings*[b]				
	I	*II*	*Communality*	*Item/scale r*[c]	*%Prevalence*[d]
Too timid to join in unorganized play	−.34	*.60*	.35	.36	5.7
Has a dejected look generally	.17	*.54*	.38	.34	3.0
Freezes up and doesn't answer questions	−.13	*.52*	.25	.28	5.4
Seems afraid to try work with hands	−.15	*.52*	.25	.25	4.9
Sits lifelessly most of time (at desk)	−.04	*.50*	.24	.31	5.1
Does not stand up for self	−.18	*.49*	.21	.30	10.4
Needs encouragement to join team games	−.27	*.49*	.23	.30	15.7
Rarely smiles	−.04	*.47*	.21	.27	4.3
Tends to have untalkative moods	−.06	*.45*	.19	.30	9.6
Not shy but rarely offers answer	.02	*.44*	.20	.32	10.3
Won't attempt if senses difficult new learning	.20	*.43*	.28	.39	9.6
Waits for teacher to greet first	−.13	*.43*	.17	.24	7.7
Distant, makes no relationship with teacher.	.24	*.42*	.29	.24	3.9
Slow and doesn't finish handwork	.10	*.42*	.21	.35	10.7
Used as scapegoat, object of ridicule	.10	*.41*	.20	.21	2.7

[a] Descriptions incorporate item content and relevant situational contexts. Item content and contexts are abbreviated for convenient presentation.

[b] Values are promaxian pattern loadings at $k = 3$, where hyperplane count is maximized. Salient pattern loadings ($\geq$ .40) are italicized. $N = 500$ comprising the random exploratory analysis subsample.

[c] Each correlation reflects the relationship between an item and the sum of the other items composing a given scale, where item distributions were standardized to unit-normal form.

[d] Entries indicate the percentage of children for whom the item behavior is scored present. Values are based on the normative calibration sample ($N = 700$), including the exploratory and confirmatory analyses subsamples.

[e] Reliability is based on the exploratory subsample ($N = 500$).

and for Underactivity was .73. Figure 1 illustrates for each dimension the overlap of total test information and measurement error. It is evident that *SS*s will have practical utility from ~ 2/3 *SD* below the population mean and throughout the highest *SS*s. This is particularly useful inasmuch as ASCA *SS*s are often used to discriminate between behavior that is adequately adjusted (< 60), versus at risk (≥ 60 and < 70) versus maladjusted (≥ 70). Overactivity and Underactivity scores were reliable (i.e., $\alpha \geq .70$) with coefficient α for Overactivity = .92 and Underactivity = .74. Coefficient α was based on the exploratory subsample ($n = 500$) for each scale.

External Validity

Table 2 displays concurrent relations between ASCA scores and independent criterion measures. All statistically significant correlations are in the expected direction with ASCA scores evincing moderate to strong relations with other teacher rating measures and anticipated lower relations with a parent rating measure and a direct assessment of reading achievement (Dinnebeil et al., 2013; Hartley, Zakriski, & Wright, 2011; van der Ende et al., 2012). Due to the nested nature of the data, the last column in the table lists the percentage of criterion

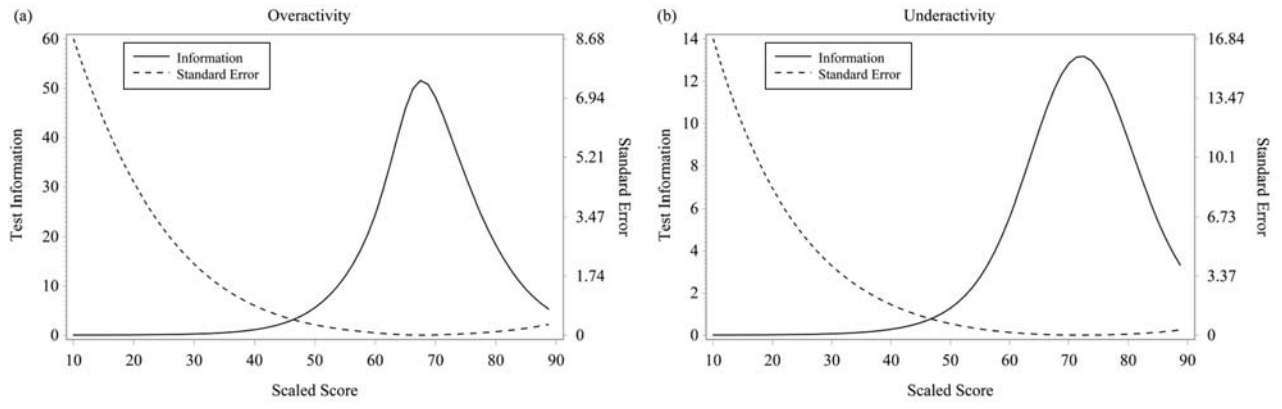

FIGURE 1 Distributions of estimated information functions and standard errors for ASCA scales.

TABLE 2
Relationships Between ASCA Scores and Concurrent Criterion Measures

Criterion measure	ASCA scale[a]		
	Overactivity	*Underactivity*	*% Explainable variance*[b]
Disruptive Behavior Disorder Rating Scale (teacher rating)			
Inattention ($n = 673$)	.55 (54.0)	.33 (27.1)	84.8
Oppositional/Defiant ($n = 630$)	.60 (59.7)	.13 (6.3)	90.9
Impulsivity/Overactivity ($n = 631$)	.65 (60.0)	−.02† (12.3)	80.0
Oral Reading Fluency (direct assessment)			
Fall Mean of A & B passages ($n = 678$)	−.26 (5.4)	−.20 (5.9)	60.5
Winter Mean of A & B passages ($n = 709$)	−.25 (8.3)	−.24 (6.0)	59.1
Spring Mean of A & B passages ($n = 678$)	−.26 (5.4)	−.20 (5.9)	60.5
Learning Behaviors Scale (teacher rating)			
Total score ($n = 755$)	−.52 (38.5)	−.44 (34.8)	83.7
Competence motivation ($n = 815$)	−.37 (17.2)	−.50 (35.0)	98.7
Attitude ($n = 811$)	−.43 (20.8)	−.48 (50.1)	87.7
Persistence ($n = 818$)	−.54 (33.3)	−.34 (19.2)	92.0
Strategy ($n = 797$)	−.55 (38.5)	−.17 (32.2)	72.3
Adjustment Scales for Children and Adolescents—Home Edition (parent rating)			
Attention-Deficit Hyperactivity ($n = 720$)	.25 (6.8)	−.01† (12.5)	88.7
Conduct Problems ($n = 719$)	.25 (7.1)	.12 (11.2)	97.9
Underactivity ($n = 719$)	−.03† (1.9)	.10 (5.8)	84.9

[a] Nonparenthetical entries are Pearson product moment correlations. Parenthetical entries indicate the percentage of variance in the respective criterion measure scores between children within classrooms that is accounted for by a given ASCA scale score. Values equal the proportional reduction in the residual variance (100) as estimated via hierarchical linear modeling. Each two-level random coefficients model entered a given ASCA scale as the covariate. All correlations and fixed effects associated with ASCA scales are significant statistically at $p < .01$ unless indicated † (nonsignificant). ASCA = Adjustment Scales for Children and Adolescents.

[b] Total percentage of potentially explainable variance between children within classrooms. Values equal 1 − intraclass correlation (100), where the intraclass correlation was estimated via hierarchical linear modeling. Each two-level, unconditional means model applied random intercepts for classrooms, where the random effect was significant at $p < .001$.

measure variance that reflects children's actual individual differences and parenthetical values indicate how much of that variance is accounted for by a given ASCA scale. For example, Table 2's last column entry for the DBDRS Inattention scale indicates that, whereas 84.8% of score variance stems from children's individual differences (rather than teacher characteristics), 54.0% of that variance is predictable from children's ASCA Overactivity scores and 27.1% predictable from ASCA Underactivity scores. Teachers seem to be more sensitive to reporting overactivity-type problems. ASCA Overactivity and Underactivity scores are similarly effective in accounting for individual differences in reading performance.

Demographic Trends

Table 3 displays the mean population distribution of Overactivity and Underactivity by gender and grade level in Trinidad and Tobago, while Table 4 shows the distribution by gender and ethnicity. There appears to have been no peaking or appreciable shifts in variance on either dimension for these characteristics.

DISCUSSION

This study sought to establish representative national norms and examine the psychometric properties for ASCA in Trinidad and Tobago. There has been extensive work on the normative development, dimensionality, and validity of ASCA for application in the United States (Canivez, 2004; Canivez & Bordenkircher, 2002; Canivez & Rains, 2002; Canivez & Sprouls, 2005; McDermott, 1993; McDermott et al., 2005), but little empirical evaluation regarding its use in international populations. This analysis of the properties of ASCA with a population from Trinidad and Tobago expands the literature on cross-cultural applications of ASCA, possibly laying the foundation for large-scale multicultural comparisons similar to those based on parent reports (Ivanova et al., 2010; Rescorla et al., 2007; Rescorla et al., 2011).

A nationally representative sampling of the distribution of behavior problems in Trinidad and Tobago uncovered two broad dimensions of childhood problem behavior, Overactivity and Underactivity, through exploratory and confirmatory analyses. The resulting dimensional structure of Overactivity and Underactivity differed from the hierarchical structure derived in the United States

TABLE 3
Mean Population Distribution of Overactivity and Underactivity by Gender and Grade Level in Trinidad and Tobago

		Overactivity		*Underactivity*	
Gender		*M*	*(SD)*	*M*	*(SD)*
		Infant 1			
Male	($n = 50$)	51.0	*(8.7)*	49.9	*(7.4)*
Female	($n = 50$)	51.7	*(9.0)*	51.1	*(8.3)*
		Infant 2			
Male	($n = 50$)	52.1	*(10.1)*	50.2	*(8.2)*
Female	($n = 50$)	49.0	*(8.1)*	51.4	*(8.8)*
		Standard 1			
Male	($n = 50$)	52.5	*(9.9)*	50.5	*(8.7)*
Female	($n = 50$)	49.3	*(8.7)*	49.2	*(8.1)*
		Standard 2			
Male	($n = 50$)	51.0	*(8.4)*	49.0	*(7.5)*
Female	($n = 50$)	46.4	*(8.1)*	49.5	*(7.7)*
		Standard 3			
Male	($n = 50$)	51.5	*(9.7)*	50.9	*(8.4)*
Female	($n = 50$)	48.7	*(9.0)*	50.0	*(8.2)*
		Standard 4			
Male	($n = 50$)	48.8	*(8.8)*	49.5	*(7.6)*
Female	($n = 50$)	46.1	*(8.0)*	48.6	*(6.8)*
		Standard 5			
Male	($n = 50$)	53.0	*(10.7)*	50.2	*(6.7)*
Female	($n = 50$)	48.9	*(8.5)*	50.1	*(8.2)*
		Total			
Male	($n = 350$)	51.4	*(9.5)*	50.0	*(7.8)*
Female	($n = 350$)	48.6	*(8.6)*	50.0	*(8.0)*

standardization that presented the same externalizing and internalizing dimensions, but with additional underlying dimensions of specific syndromes (McDermott, 1993). However, the Overactivity and Underactivity dimensional structure has been found in other populations including Hispanic/Latino, Native American, and Canadian youth (Canivez & Beran, 2009; Canivez & Bohan, 2006; Canivez, 2006; Canivez & Sprouls, 2010).

TABLE 4
Mean Population Distribution of Overactivity and Underactivity by Gender and Ethnicity in Trinidad and Tobago

		Overactivity		*Underactivity*	
Gender		*M*	*(SD)*	*M*	*(SD)*
		African descent			
Male	($n = 137$)	53.5	*(9.7)*	49.1	*(7.5)*
Female	($n = 132$)	49.8	*(9.4)*	49.8	*(8.0)*
		East Indian descent			
Male	($n = 126$)	49.4	*(9.0)*	50.4	*(7.8)*
Female	($n = 131$)	46.8	*(7.3)*	50.0	*(7.9)*
		Mixed descent			
Male	($n = 76$)	51.3	*(9.1)*	51.4	*(8.4)*
Female	($n = 73$)	49.2	*(8.3)*	49.7	*(7.6)*
		Total			
Male	($n = 350$)	51.4	*(9.5)*	50.0	*(7.8)*
Female	($n = 350$)	48.6	*(8.6)*	50.0	*(8.0)*

Emergence of the Overactivity and Underactivity factors provides further evidence on the cross-cultural universality of these two broadband problem behavior dimensions (Eysenck, 1953; Kohn, 1977; Peterson, 1961; Rutter, 1967). Our Underactivity dimension, similar to internalizing, is associated by problems with negative emotion (Markon, 2010) whereas Overactivity, similar to externalizing, is characterized by disinhibition problems. The two-dimensional finding in Trinidad and Tobago bolsters the robustness of the two-dimensional IE model, a model that has remained consistent across age, sex, ethnicity, culture, informant type, instrument, and DSM Axes (Achenbach, 1966; Eaton et al., 2011; Forbush & Watson, 2013; Kramer et al., 2008; Krueger et al., 1998; Krueger et al., 2003; Lahey et al., 2008; Slade & Watson, 2006; Wright et al., 2013).

Validity analyses indicated ASCA scores have moderate to strong relationships with other teacher rating measures and weaker relationships with a parent rating measure and reading achievement. Additionally, teachers seem to be more sensitive to overactive- than underactive-type problems. Overactivity and Underactivity scores accounted for similar levels of actual individual differences in reading fluency. The lower relationships between parent and teacher ratings were anticipated based on past discoveries of behavior rating differences among teachers and parents (Achenbach, McConaughy, & Howell, 1987). Informant discrepancies may be explained partially by cross-contextual variability in children's behavior, differences in the environments in which child behavior is observed, and informants' differing perspectives and tolerance for child behavior (De Los Reyes, Thomas, Goodman, & Kundey, 2013).

The islands' normative sample included only elementary school students, limiting the age range for comparisons between Trinidad and Tobago and the United States, which included secondary school children. Validity analyses of academic achievement were also constrained to a reading fluency measure because no other academic assessments were available. Further, this was a curriculum-based measure, as there are currently no mandated standardized achievement assessments in Trinidad and Tobago. Each student was rated by one teacher only, which precluded us from calculating interrater reliability. However, prior research in the U.S. found good agreement between teachers ($r \geq .80$) on the two broad ASCA classifications (Canivez & Watkins, 2002; Schaefer et al., 2001; Watkins & Canivez, 1997). Additionally, dependence on item-level binary data puts limits on standard estimation procedures (convergence problems, necessity for item parcels, etc.). One alternative we are exploring is item reduction by conducting context-level instead of item-level analysis, as per McDermott, Watkins, Rovine, and Rikoon (2014).

Given the popular ethnographic research conclusion that Trinidadian parents evince a distinct penchant for child rearing centered on order and discipline (Barrow, 2008; Cappa & Khan, 2011; Gopaul-McNicol, 1993, 1999), we

thought it important to contrast ASCA empirical trends in Trinidad and Tobago with those reported for the American standardization sample (McDermott, 1993, 1994). It has been hypothesized that a cultural standard more sensitive to indiscipline would likely generate higher levels of observed overactive behavior and show a greater tolerance and lower levels for underactive (albeit compliant) behavior (Roopnarine et al., 2014). Accordingly, the average prevalence for Overactivity problems in Trinidad and Tobago was 9.4%, whereas in the United States prevalence was only 6.7%. In turn, average prevalence for Underactivity behaviors in the islands was 7.0%, whereas in the United States it was 7.5%, thus lending empirical support to the ethnographic hypothesis.

Additional contrast is provided in ancillary analyses applying the U.S. national scoring parameters to the Trinidad and Tobago normative sample. Here, the U.S. Overactivity and Underactivity *SS* means were 51.5 and 49.7, respectively, for the same age group as the Trinidad and Tobago normative sample, with the corresponding means for the islands being 53.7 and 49.0, respectively, again lending support to the ethnographic viewpoint.

CONCLUSION

The standardization and validation of ASCA marks a major step in the establishment of culturally relevant standards for behavioral assessment in Trinidad and Tobago. In 1997, the Ministry of Education's Central Guidance and Special Education Units began the Continuous Assessment Progress project and set the goal of identifying low-performing students and providing them with supportive services (Watkins et al., 2014). The ministry later merged the Guidance and Special Education Units into the Student Support Services Unit while keeping the focus on preventative services. Though the Ministry of Education has conveyed interest in using assessment data to offer targeted services to students and to improve school performance, more development needs to be done. However, Trinidad and Tobago has invested far more in assessments and data collection than many other developing countries (Hu & Oakland, 1991; Oakland, 2004; Oakland & Hu, 1993; Oakland et al., 2013) and the information is ready and available for identifying at-risk students. The standardization and validation of ASCA in Trinidad and Tobago has created a bridge that can be used for intervention with struggling students. With proper interventions and supportive services, needy students are more likely to reach their full academic potential and become productive members of society.

FUNDING

This research was supported in part by the Ministry of Education, Trinidad and Tobago, and by the Institute of Education Sciences, U.S. Department of Education, Grant #R305B090015.

ABOUT THE AUTHORS

Dr. Paul A. McDermott, PhD, is a Professor of Quantitative Methods at the University of Pennsylvania, Graduate School of Education.

Dr. Marley W. Watkins, PhD, is a Non-Resident Scholar in the Department of Educational Psychology at Baylor University.

Anna M. Rhoad, MS, is a graduate student in Quantitative Methods at the University of Pennsylvania, Graduate School of Education.

Jessica L. Chao, MS, is a graduate student in Quantitative Methods at the University of Pennsylvania, Graduate School of Education.

Dr. Frank C. Worrell, PhD, is a Professor of School Psychology at the University of California- Berkeley, Graduate School of Education.

Dr. Tracey E. Hall, PhD, is a Senior Research Scientist and Instructional Designer at the Center for Applied Special Technology.

REFERENCES

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied*, *80*(7), 1–37. doi:10.1037/h0093906

Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin*, *85*, 1275–1301. doi:10.1037/0033-2909.85.6.1275

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232. doi:10.1037/0033-2909.101.2.213

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.). Washington, DC: Author.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, *9*, 78–102. doi:10.1207/S15328007SEM0901_5

Barrow, C. (2008). *Early childhood in the Caribbean. Working papers in early childhood development, No. 47*. The Hague, The Netherlands: Bernard van Leer Foundation.

Barry, C. T., Frick, P. J., & Kamphaus, R. W. (2013). Psychological assessment in child mental health settings. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology: Testing and assessment in clinical and counseling psychology* (Vol. 2, pp. 253–270). Washington, DC: American Psychological Association.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, *105*, 467–477. doi:10.1037/0033-2909.105.3.467

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., & Erklani, A. (1995). Children's mental health services use across services sectors. *Health Affairs*, *14*, 147–159. doi:10.1377/hlthaff.14.3.147

Canivez, G. L. (2004). Replication of the Adjustment Scales for Children and Adolescents core syndrome factor structure. *Psychology in the Schools*, *41*, 191–199. doi:10.1002/pits.10121

Canivez, G. L. (2006). Adjustment Scales for Children and Adolescents and Native American Indians: Factorial validity generalization for Ojibwe youths. *Psychology in the Schools*, *43*, 685–694. doi:10.1002/pits.20179

Canivez, G. L., & Beran, T. N. (2009). Adjustment Scales for Children and Adolescents: Factorial validity in a Canadian sample. *Canadian Journal of School Psychology*, *24*, 284–302. doi:10.1177/0829573509344344

Canivez, G. L., & Bohan, K. J. (2006). Adjustment Scales for Children and Adolescents and Native American Indians: Factorial validity generalization for Yavapai Apache youths. *Journal of Psychoeducational Assessment*, *24*, 329–341. doi:10.1177/0734282906291397

Canivez, G. L., & Bordenkircher, S. E. (2002). Convergent and divergent validity of the Adjustment Scales for Children and Adolescents and the Preschool and Kindergarten Behavior Scales. *Journal of Psychoeducational Assessment*, *20*, 30–45. doi:10.1177/073428290202000103

Canivez, G. L., Neitzel, R., & Martin, B. E. (2005). Construct validity of the Kaufman Brief Intelligence Test, Wechsler Intelligence Scale for Children–Third Edition, and Adjustment Scales for Children and Adolescents. *Journal of Psychoeducational Assessment*, *23*, 15–24. doi:10.1177/073428290502300102

Canivez, G. L., Perry, A. R., & Weller, E. M. (2001). Stability of the Adjustment Scales for Children and Adolescents. *Psychology in the Schools*, *36*, 217–227. doi:10.1002/pits.1012

Canivez, G. L., & Rains, J. D. (2002). Construct validity of the Adjustment Scales and Adolescents and the Preschool and Kindergarten Behavior Scales: Convergent and divergent evidence. *Psychology in the Schools*, *39*, 621–633. doi:10.1002/pits.10063

Canivez, G. L., & Sprouls, K. (2005). Assessing the construct validity of the Adjustment Scales for Children and Adolescents. *Journal of Psychoeducational Assessment*, *23*, 3–14. doi:10.1177/073428290502300101

Canivez, G. L., & Sprouls, K. (2010). Adjustment Scales for Children and Adolescents: Factorial validity generalization with Hispanic/Latino youths. *Journal of Psychoeducational Assessment*, *28*, 209–221. doi:10.1177/0734282909349213

Canivez, G. L., & Watkins, M. W. (2002). Interrater agreement for syndromic profile classifications on the Adjustment Scales for Children and Adolescents. *Assessment for Effective Intervention*, *28*, 38–46. doi:10.1177/073724770202800105

Canivez, G. L., Watkins, M. W., & Schaefer, B. A. (2002). Interrater agreement for discriminant classifications for the Adjustment Scales for Children and Adolescents. *Psychology in the Schools*, *39*, 375–384. doi:10.1002/pits.10043

Cappa, C., & Khan, S. M. (2011). Understanding caregivers' attitudes towards physical punishment of children: Evidence from 34 low- and middle-income countries. *Child Abuse & Neglect*, *35*, 1009–1021. doi:10.1016/j.chiabu.2011.10.003

Carlson, J. F., Benson, N., & Oakland, T. (2010). Implications of the international classification of functioning, disability and health (ICF) for test development and use. *School Psychology International*, *31*, 353–371. doi:10.1177/0143034310377149

Central Intelligence Agency. (2014). *The world factbook, 2013–2014, Trinidad and Tobago*. Retrieved July 15, 2014, from https://www.cia.gov/library/publications/the-world-factbook/index.html

Coffey, D. Y. (2006). *A structural validity study of the Adjustment Scales for Children and Adolescents—Home (ASCA-H)* (Doctoral dissertation). Pennsylvania State University, State College, PA. Retrieved from https://etda.libraries.psu.edu/paper/6971/2242

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, *56*, 754–761. doi:10.1037/0022-006X.56.5.754

Debelak, R., & Tran, U. S. (2013). Principal component analysis of smoothed tetrachoric correlation matrices as a measure of dimensionality. *Educational and Psychological Measurement*, *73*, 63–77. doi:10.1177/0013164412457366

De Clercq, B., De Fruyt, F., Van Leeuwen, K., & Mervielde, I. (2006). Structure of maladaptive personality traits in childhood: A step toward an integrative developmental perspective for DSM-V. *Journal of Abnormal Psychology*, *115*, 639–657. doi:10.1037/0021-843X.115.4.639

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, *9*, 123–149. doi:10.1146/annurev-clinpsy-050212-185617

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, *49*, 36–45.

Dinnebeil, L. A., Sawyer, B. E., Loganc, J., Dyniac, J. M., Cancioa, E., & Justice, L. M. (2013). Influences on the congruence between parents' and teachers' ratings of young children's social skills and problem behaviors. *Early Childhood Research Quarterly*, *28*, 144–152. doi:10.1016/j.ecresq.2012.03.001

Dowdy, E., Twyford, J., & Sharkey, J. D. (2013). Methods of assessing behavior: Observations and rating scales. In D. H. Saklofske, C. R. Reynolds, & V. Schwean (Eds.), *Oxford handbook of child psychological assessment* (pp. 646–718). New York, NY: Oxford University Press.

Eaton, N. R., Krueger, R. F., & Oltmanns, T. F. (2011). The structure and long-term stability of the internalizing spectrum of personality and psychopathology across the lifespan. *Psychology and Aging*, *26*, 987–993. doi:10.1037/a0024406

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, 6th ed). New York, NY: Wiley.

Elliott, C. D. (1990). *Differential Abilities Scale: Introductory and technical handbook*. San Antonio, TX: Psychological Corporation.

Eysenck, H. J. (1953). *The structure of human personality.* London, England: Methuen.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299. doi:10.1037/1082-989X.4.3.272

Forbush, K. T., & Watson, D. (2013). The structure of common and uncommon mental disorders. *Psychological Medicine*, *43*, 97–108. doi:10.1017/S0033291712001092

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*, 239–256. doi:10.1207/S1532799XSSR0503_3

George, S. M., McDermott, P. A., Watkins, M. W., Worrell, F. C., & Hall, T. E. (2012). The assessment of youth psychopathology in Trinidad and Tobago: A cross-cultural construct validity study of the Adjustment Scales for Children and Adolescents. *International Journal of Educational and Psychological Assessment*, *10*, 159–178.

Gomez, R., & Vance, A. (2014). Confirmatory factor analysis, latent profile analysis, and factor mixture modeling of the syndromes of the child behavior checklist and teacher report form. *Psychological Assessment*, *26*, 1307–1316. doi:10.1037/a0037431

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, *38*, 1179–1191. doi:10.1007/s10802-010-9434-x

Gopaul-McNicol, S.-A. (1993). *Working with West Indian families*. New York, NY: Guilford.

Gopaul-McNicol, S.-A. (1999). Ethnocultural perspectives on childrearing practices in the Caribbean. *International Social Work*, *42*, 69–86. doi:10.1177/002087289904200108

Greenberg, D. F., & Agozino, B. (2012). Executions, imprisonment and crime in Trinidad and Tobago. *British Journal of Criminology*, *52*, 113–140. doi:10.1093/bjc/azr072

Guerra, N. G., & Bradshaw, C. P. (2008). Linking the prevention of problem behaviors and positive youth development: Core competencies for positive youth development and risk prevention. *New Directions for Child and Adolescent Development*, *122*, 1–17. doi:10.1002/cd.225

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary

constructs. *Organizational Research Methods*, *2*, 232–256. doi:10.1177/109442819923002

Hambleton, R. K., & Patsula, L. (1998). Adapting tests for use in multiple languages and cultures. *Social Indicators Research*, *45*, 153–171. doi:10.1023/A:1006941729637

Hartley, A. G., Zakriski, A. L., & Wright, J. C. (2011). Probing the depths of informant discrepancies: Contextual influences on divergence and convergence. *Journal of Clinical Child & Adolescent Psychology*, *40*, 54–66. doi:10.1080/15374416.2011.533404

Hau, K.-T., & Marsh, H. W. (2004). The use of item parcels in structural equation modeling: Non-normal data and small sample sizes. *British Journal of Mathematical Statistical Psychology*, *57*, 327–351. doi:10.1111/j.2044-8317.2004.tb00142.x

Hsu, H.-Y., Skidmore, S. T., Li, Y., & Thompson, B. (2014). Forced zero cross-loading misspecifications in measurement component of structural equation models: Beware of even "small" misspecifications. *Methodology*, *10*, 138–152. doi:10.1027/1614-2241/a000084

Hu, S., & Oakland, T. (1991). Global and regional perspectives on testing children and youth: An empirical study. *International Journal of Psychology*, *26*, 329–344. doi:10.1080/00207599108246857

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Harder, V. S., Ang, R. P., Bilenberg, N., … Verhulst, F. C. (2010). Preschool psychopathology reported by parents in 23 societies: Testing the seven-syndrome model of the child behavior checklist for ages 1.5–5. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 1215–1224. doi:10.1016/j.jaac.2010.08.019

Johnstone, C. J. (2010). Inclusive education policy implementation: Implications for teacher workforce development in Trinidad and Tobago. *International Journal of Special Education*, *25*, 33–42.

Kishton, J. M., & Widaman, K. F. (1994). Unidimensional versus domain representatives parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement*, *54*, 757–765. doi:10.1177/0013164494054003022

Knol, D. L., & Berger, M. P. F. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research*, *26*, 457–477. doi:10.1207/s15327906mbr2603_5

Kohn, M. (1977). *Social competence, symptoms and underachievement in childhood: A longitudinal perspective*. Washington, DC: Winston-Wiley.

Kramer, M. D., Krueger, R. F., & Hicks, B. M. (2008). The role of internalizing and externalizing liability factors in accounting for gender differences in the prevalence of common psychopathological syndromes. *Psychological Medicine*, *38*, 51–61. doi:10.1017/S0033291707001572

Krishnakumar, A., Narine, L., Roopnarine, J. L., & Logie, C. (2014). Multilevel and cross-level effects of neighborhood and family influences on children's behavioral outcomes in Trinidad and Tobago: The intervening role of parental control. *Journal of Abnormal Child Psychology*, *42*, 1057–1068. doi:10.1007/s10802-014-9852-2

Krueger, R. F., Capsi, A., Moffitt, T. E., & Silva, P. A. (1998). The structure and stability of common mental disorders (DSM-III-R): A longitudinal-epidemiological study. *Journal of Abnormal Psychology*, *107*, 216–227. doi:10.1037/0021-843X.107.2.216

Krueger, R. F., Chentsova-Dutton, Y. E., Markon, K. E., Goldberg, D. P., & Ormel, J. (2003). A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *Journal of Abnormal Psychology*, *112*, 437–447. doi:10.1037/0021-843X.107.2.216

Lahey, B. B., Rathouz, P. J., van Hulle, C., Urbano, R. C., Krueger, R. F., Applegate, B., … Chapman, D. A. (2008). Testing structural models of DSM-IV symptoms of common forms of child and adolescent psychopathology. *Journal of Abnormal Child Psychology*, *36*, 187–206. doi:10.1007/s10802-007-9169-5

Markon, K. E. (2010). Modeling psychopathology structure: A symptom-level analysis of Axis I and II disorders. *Psychological Medicine*, *40*, 273–288. doi:10.1017/S0033291709990183

Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological measurement fruitfulness of exploratory structural equation modeling (ESEM): New approaches to key substantive issues in motivation and engagement. *Journal of Psychoeducational Assessment*, *29*, 322–346. doi:10.1177/0734282911406657

Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist*, *53*, 205, doi:10.1037/0003-066X.53.2.205

McConaughy, S. H., & Ritter, D. R. (2014). Best practices in multimethod assessment of emotional and behavioral disorders. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 367–390). Bethesda, MD: National Association of School Psychologists.

McDermott, P. A. (1993). National standardization of uniform multi-situational measures of child and adolescent behavior pathology. *Psychological Assessment*, *5*, 413–424. doi:10.1037/1040-3590.5.4.413

McDermott, P. A. (1994). *Manual of adjustment scales for children and adolescents*. Philadelphia, PA: Edumetric and Clinical Science.

McDermott, P. A. (1999). National scales of differential learning behaviors among American children and adolescents. *School Psychology Review*, *28*, 280–291.

McDermott, P. A., Steinberg, C. M., & Angelo, L. E. (2005). Situational specificity makes the difference in assessment of youth behavior disorders. *Psychology in the Schools*, *42*, 121–156. doi:10.1002/pits.20050

McDermott, P. A., Stott, D. H., & Marston, N. C. (1993). *Adjustment scales for children and adolescents*. Philadelphia, PA: Edumetric and Clinical Science.

McDermott, P. A., Watkins, M. W., Rovine, M. J., & Rikoon, S. H. (2013). Assessing changes in socioemotional adjustment across the early school transitions—New national scales for children at risk. *Journal of School Psychology*, *51*, 97–115. doi:10.1016/j.jsp.2012.10.002

McDermott, P. A., Watkins, M. W., Rovine, M. J., & Rikoon, S. H. (2014). Informing context and change in young children's sociobehavioral development: The national Adjustment Scales for Early Transition in Schooling (ASETS). *Early Childhood Research Quarterly*, *29*, 255–267. doi:10.1016/j.ecresq.2014.02.004

McDermott, P. A., Watkins, M. W., Sichel, A. F., Weber, E. M., Keenan, J. T., Holland, A. M., & Leigh, N. M. (1995). The accuracy of new national scales for detecting emotional disturbance in children and adolescents. *Journal of Special Education*, *29*, 337–354. doi:10.1177/002246699502900306

McKenna, M. C., & Layton, K. (1990). Concurrent validity of cloze as a measure of intersentential comprehension. *Journal of Educational Psychology*, *82*, 372–377. doi:10.1037/0022-0663.82.2.372

Merrell, K. W. (2008). *Behavioral, social, and emotional assessment of children and adolescents* (3rd ed.). New York, NY: Erlbaum.

Mordell, M. (2001). *Parent ratings of behavior: Contextually based assessment for children and adolescents* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA.

Morris, J., Belfer, M., Daniels, A., Flisher, A., Villé, L., Lora, A., & Saxena, S. (2011). Treated prevalence of and mental health services received by children and adolescents in 42 low-and-middle-income countries. *Journal of Child Psychology and Psychiatry*, *52*, 1239–1246. doi:10.1111/j.1469-7610.2011.02409.x

Mpofu, E., Oakland, T., Ntinda, K., Seeco, E., & Maree, J. G. (2014). Constructing a framework for the use of tests within a developing nation's school system. *International Perspectives in Psychology: Research, Practice, Consultation*, *3*, 106–122. doi:10.1037/ipp0000015

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*, 729–757. doi:10.1177/0013164403258228

Oakland, T. (2004). Use of educational and psychological tests internationally. *Applied Psychology: An International Review*, *53*, 157–172. doi:10.1111/j.1464-0597.2004.00166.x

Oakland, T., & Hu, S. (1993). International perspectives on tests used with children and youth. *Journal of School Psychology*, *31*, 501–517. doi:10.1016/0022-4405(93)90034-G

Oakland, T., Wechsler, S. M., & Maree, K. (2013). Test use with children across cultures: A view from three countries. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology: Testing and assessment in school psychology and education* (Vol. 3, pp. 231–257). Washington, DC: American Psychological Association.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, *44*, 443–460. doi:10.1007/BF02296207

Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *The Lancet*, *369*, 1302–1313. doi:10.1016/S0140-6736(07)60368-7

Pelham, W. E. Jr., Evans, S. W., Gnagy, E. N., & Greenslade, K. E. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders: Prevalence, factor analyses, and conditional probabilities in a special education sample. *School Psychology Review*, *21*, 285–299.

Pelham, W. E. Jr., Gnagy, E. N., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child & Adolescent Psychology*, *31*, 210–218. doi:10.1097/00004583-199203000-00006

Pendergast, L. L., Vandiver, B. J., Schaefer, B. A., Cole, P. M., Murray-Kolb, L. E., & Christian, P. (2014). Factor structure of scores from the Conners' Rating Scales–Revised among Nepali children. *International Journal of School & Educational Psychology*, *2*, 261–270. doi:I0.1080/21683603.2013.878678

Peterson, D. R. (1961). Behavior problems of middle childhood. *Journal of Consulting Psychology*, *25*, 205–209. doi:10.1037/h0038994

Psychological Corporation. (1983). *Basic Achievement Skills Individual Screener*. New York, NY: Author.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement of oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*, 427–469. doi:10.1016/j.jsp.2009.07.001

Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., … Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders*, *15*, 130–142. doi:10.1177/10634266070150030101

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Harder, V. S., Otten, L., Bilenberg, N., … Verhulst, F. C. (2011). International comparisons of behavioral and emotional problems in preschool children: Parents' reports from 24 societies. *Journal of Clinical Child & Adolescent Psychology*, *40*, 456–467. doi:10.1080/15374416.2011.563472

Roopnarine, J. L., Krishnakumar, A., Narine, L., Logie, C., & Lape, M. E. (2014). Relationships between parenting practices and preschoolers' social skills in African, Indo, and mixed-ethnic families in Trinidad and Tobago: The mediating role of ethnic socialization. *Journal of Cross-Cultural Psychology*, *45*, 362–380. doi:10.1177/0022022113509884

Rubin, K. H. (1998). Social and emotional development from a cultural perspective. *Developmental Psychology*, *34*, 611–615. doi:10.1037/0012-1649.34.4.611

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Journal of child Psychology and Psychiatry*, *8*, 1–11. doi:10.1111/j.1469-7610.1967.tb02175.x

Sass, D. A., & Smith, P. L. (2006). The effects of parceling unidimensional scales on structural parameter estimates in structural equation modeling. *Structural Equation Modeling*, *13*, 566–586. doi:10.1207/s15328007sem1304_4

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. doi:10.1007/BF02296192

Schaefer, B. A., Watkins, M. W., & Canivez, G. L. (2001). Cross-context agreement of the adjustment scales for children and adolescents. *Journal of Psychoeducational Assessment*, *19*, 123–136. doi:10.1177/073428290101900202

Slade, T., & Watson, D. (2006). The structure of common DSM-IV and ICD-10 mental disorder in the Australian general population. *Psychological Medicine*, *36*, 1593–1600. doi:10.1017/S0033291706008452

Slobodskaya, H. R. (2014). The hierarchical structure of personality and common psychopathology in childhood. *Journal of Research in Personality*, *53*, 36–46. doi:10.1016/j.jrp.2014.08.005

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, *106*, 148–154. doi:10.1037/0033-2909.106.1.148

Stott, D. H. (1979). The Bristol social adjustment guides. *Therapeutic Education*, *7*, 34–44.

Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. (1995). Item response theory for scores on tests including polytomous items with ordered responses. *Applied Psychological Measurement*, *19*, 39–49. doi:10.1177/014662169501900105

Thompson, B., & Melancon, J. (November, 1996). *Using item "testlets/parcels" in confirmatory factor analysis: An example using the PPDP-78*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Tuscaloosa, AL.

United Nations. (2014). *Trinidad and Tobago development assistance framework: 2014–2015 biennial rolling plan*. Retrieved on February 18, 2015, from http://www.tt.undp.org/content/dam/trinidad_tobago/docs/DemocraticGovernance/Publications/UNDAF%20%282014%20-%202015%29%20-%20Trinidad%20%20Tobago%20Final.pdf

United Nations Children's Fund. (2013). *The state of the world's children 2013*. New York, NY: Author.

van der Ende, J., Verhulst, F. C., & Tiemeier, H. (2012). Agreement of informants on emotional and behavioral problems from childhood to adulthood. *Psychological Assessment*, *24*, 293–300. doi:10.1037/a0025500

Van Meter, A., Youngstrom, E., Youngstrom, J. K., Ollendick, T., Demeter, C., & Findling, R. L. (2014). Clinical decision making about child and adolescent anxiety disorders using the Achenbach system of empirically based assessment. *Journal of Clinical Child & Adolescent Psychology*, *43*, 552–565. doi:10.1080/15374416.2014.883930

van Widenfelt, B. M., Treffers, P. D. A., de Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review*, *8*, 135–147. doi:10.1007/s10567-005-4752-1

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321–327. doi:10.1007/BF02293557

Waller, N. G. (2001). *MicroFACT 2.0: A microcomputer factor analysis program for ordered polytomous data and mainframe size problems*. St. Paul, MN: Assessment Systems Corporation.

Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2012). The matter of assessor variance in early childhood education—Or whose score is it anyway? *Early Childhood Research Quarterly*, *27*, 46–54. doi:10.1016/j.ecresq.2011.06.003

Watkins, M. W., & Canivez, G. L. (1997). Interrater agreement of the Adjustment Scales for Children and Adolescents. *Diagnostique*, *22*, 205–213. doi:10.1177/073724779702200402

Watkins, M. W., & Edwards, V. A. (2004). Assessing early literacy skills with the mountain shadows phonemic awareness scale (MS-PAS). *Journal of Psychoeducational Assessment*, *22*, 3–14. doi:10.1177/073428290402200101

Watkins, M. W., Hall, T. E., & Worrell, F. C. (2014). From Central Guidance Unit to Student Support Services Unit: The outcome of a

consultation process in Trinidad and Tobago. *Journal of Educational and Psychological Consultation*, *24*, 283–306. doi:10.1080/10474412.2014.929962

Watkins, M., & McDermott, P. A. (2002). *The adjustment scales for children and adolescents—Home*. Unpublished.

Weissberg, R. P., & Greenberg, M. T. (1998). School and community competence-enhancement and prevention programs. In I. E. Siegel & K. A. Renninger (Eds.), *Handbook of child psychology: Child psychology in practice* (Vol. 4, 5th ed., pp. 877–954). New York, NY: Wiley.

Wilkinson, W. W. (2007). The structure of the Lawrence locus of control scale in young adults: Comparing item and parcel indicator models. *Personality and Individual Differences*, *43*, 416–425. doi:10.1016/j.paid.2007.04.018

Williams, H. M. A. (2013). Postcolonial structural violence: A study of school violence in Trinidad & Tobago. *International Journal of Peace Studies*, *18*(2), 39–64.

World Health Organization. (2005). *Atlas: Child and adolescent mental health resources: Global concerns, implications for the future*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2007). *International classification of functioning, disability and health*. Geneva, Switzerland: World Health Organization.

Worrell, F. C., Vandiver, B. J., & Watkins, M. W. (2001). Construct validity of the Learning Behaviors Scale with an independent sample of students. *Psychology in the Schools*, *38*, 207–215. doi:10.1002/pits.1011

Wright, A. G., Krueger, R. F., Hobbs, M. J., Markon, K. E., Eaton, N. R., & Slade, T. (2013). The structure of psychopathology: Toward an expanded quantitative empirical model. *Journal of Abnormal Psychology*, *122*, 281. doi:10.1037/a0030133

Yates, A. (1987). *Multivariate exploratory data analysis: A prospective on exploratory factor analysis*. Albany, NY: State University of New York Press.