Psychology in the Schools, Vol. 53(6), 2016
View this article online at wileyonlinelibrary.com/journal/pits

DOI: 10.1002/pits.21930

CLASSROOM CONTEXTS AS THE FRAMEWORK FOR ASSESSING SOCIAL–EMOTIONAL ADJUSTMENT: A NATIONAL STUDY IN TRINIDAD AND TOBAGO

PAUL A. MCDERMOTT

University of Pennsylvania

MARLEY W. WATKINS

Baylor University

ANNA RHOAD DROGALIS AND JESSICA L. CHAO

University of Pennsylvania

FRANK C. WORRELL

University of California, Berkeley

TRACEY E. HALL

Center for Applied Special Technology, Wakefield

Contextually based assessments reveal the circumstances accompanying maladjustment (the when, where, and with whom) and supply clues to the motivations underpinning problem behaviors. The Adjustment Scales for Children and Adolescents (ASCA) is a teacher rating scale composed of indicators describing behavior in 24 classroom situational contexts. This study examines the Trinidad and Tobago national normative process for the ASCA contextual dimensions with a representative sample of elementary school children ($N = 900$). Exploratory and confirmatory factor analyses yielded the same three dimensions (peer context problems, teacher context problems, and learning context problems) observed in U.S. national samples. Dimensions were scaled using item response theory (IRT) and Bayesian scoring methods, with peer and learning context problems scores relating more strongly to clinical behavior disturbances and learning context problems showing stronger association with classroom learning styles. Implications for future research and practice are discussed. © 2016 Wiley Periodicals, Inc.

Many commonly used social–emotional rating scales for children and youth require teacher observers to indicate the presence or severity of problem behaviors while neglecting the contextual frameworks within which those behaviors emerge (McDermott, Watkins, Rovine, & Rikoon, 2014). Popular devices, such as the Behavior Assessment System for Children-Second Edition (Kamphaus & Reynolds, 2007), Devereux Student Strengths Assessment (LeBuffe, Shapiro, & Naglieri, 2009), and Teacher Report Form (Achenback & Edelbrock, 1986), infer psychological difficulties purely from the number and frequency of problematic behaviors regardless of classroom context. This approach fails to supply useful information on whether problem behaviors are situationally specific, and thus perhaps reactive and limited, or whether they are pervasive across many diverse situations, and thus more indicative of serious pathology. Absent information about the situational circumstances surrounding problem behavior (the *when*, *where*, and *with whom*), it is rather implausible to infer behavioral motivation as would be fundamental to any useful intervention program.

Leading measurement theorists have emphasized the importance of contextualized specificity in instrument design. Individuals behave differently across situations so that behavior in any given situation is shaped both by the situation and the individual's unique propensities (Mischel, 1973;

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

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 no. R305B090015.

Shaffer & Postelthwaite, 2012). Contextualized scales arguably offer more validity potential than noncontextualized scales because they provide respondents with a reference point for describing children's behavior, such as behavioral withdrawal within the context of peer interactions versus within the context of teacher demands. When observers respond to decontextualized scales, they may describe the child of interest based on different (and unknown) situations for each item (Shaffer & Postelthwaite, 2012).

Alternatively, specifically contextualized behavioral measures should be more accurate representations of reality and more accurate predictors of actual outcomes than broad, decontextualized measures (Robie, Schmit, Ryan, & Zickar, 2000; Schmit, Ryan, Stierwalt, & Powell, 1995). For example, a contextual measure of reading motivation was superior to a noncontextual measure in predicting subsequent reading performance (Neugebauer, 2014), and contextualized measures of personality were superior to noncontextualized measures with validity coefficients of .24 versus .11, respectively (Shaffer & Postlethwaite, 2012). Schmit et al. (1995) explained the strength of contextualized measures using frame-of-reference (FOR) effects. FOR effects occur when ratings on scales vary idiosyncratically based on the situation that given *respondents* select as a referent when answering scale items. The FOR biasing effect can be averted by *scale developers* specifying a priori the target situations within the scale (Shaffer & Postelthwaite, 2012).

The importance of context has been supported by empirical work on classroom situations. For example, the literature suggests that (a) classroom transitions and free time noticeably influence student behavior (Joosten, Bundy, & Einfeld, 2012); (b) emotional regulation and withdrawal vary based on classroom situation (Buss, 2011); (c) the level of aggressiveness in children changes after transferring from special education to regular education classrooms (Visser, Kunnen, & van Geert, 2010); (d) preschool language skill acquisition varies as a function of peer-group context (Justice, Petscher, Schatschneider, & Mashburn, 2011); (e) modifying seating arrangements and restructuring classroom environments can improve behavior (Kern & Clemens, 2007); (f) the level of teacher involvement and individual or group nature of the activity relates to student engagement (Powell, Burchinal, File, & Kontos, 2008); and (g) emotionally supportive classrooms reduce the negative effects of child misbehavior (Dominguez, Vitiello, Fuccillo, Greenfield, & Bulotsky-Shearer, 2011).

Contextually based assessments mark a shift away from measuring the frequency of problem behaviors toward examining when and where these behaviors occur. Such assessments reveal the circumstances accompanying emergent maladjustment and thereby provide clues to the motivations underpinning the problems. Notable among these instruments is the Adjustment Scales for Children and Adolescents (ASCA; McDermott, Stott, & Marston, 1993), which presents a framework of many classroom situations (e.g., reacting to correction by the teacher, behaving with smaller or weaker children, responding to new learning tasks, helping the teacher, engaging in organized play, etc.) and within each situation asks the teacher to select from alternative healthy and problematic indicators to describe the child's typical behaviors. ASCA was standardized on a large representative sample of American children spanning aged 5–17 years (McDermott, 1993, 1994) and yielded three core dimensions reflecting peer context problems, learning context problems, and teacher context problems (McDermott, Steinberg, & Angelo, 2005). Moreover, the same contextual dimensions were discovered with modified ASCA classroom situations and behavioral contexts for Head Start (Bulotsky-Shearer, Fantuzzo, & McDermott, 2008) and for children enrolled nationwide in various types of preschools through first grade (McDermott, Watkins, Rovine, & Rikoon, 2014).

Although there is now substantial support for the universality of the peer, learning, and teacher contextual dimensions in the U.S. school population, recent events have invited the possibility of testing the generality of this type of contextually based behavioral assessment in an international setting. Specifically, the island nation of Trinidad and Tobago elected through its Ministry of Education to standardize ASCA scores on a nationally representative sample (see Watkins, Hall,

& Worrell, 2014). This initiative was part of a broader national effort to provide avenues for early identification and intervention for distressed youth. Trinidad and Tobago, given its unique colonial history, eventual independence, cultural diversity, and prevailing traditions for child rearing, provided a special opportunity to assess the universality of our understanding of how classroom context can inform youth behavioral adjustment and maladjustment. In this study, we report on the national normative process for ASCA in Trinidad and Tobago, including sample construction, data collection, structural analyses, scaling, and validity support for resulting contextual dimensions. It was hypothesized that representative assessments within the islands would generate the same reliable dimensions of context problems discovered in the United States and that those dimensions would find external support with independent measures of children's school and home adjustment, learning behaviors, and academic achievement.

Method

Setting

Trinidad and Tobago is a twin island country located northeast of Venezuela in the Caribbean. Originally colonized by several European nations, the country obtained its independence from Great Britain in 1962, but remains an English-speaking, Commonwealth country. Education is free and obligatory for all children aged 5 through 16 years. Parent–child interactions tend to be characterized by discipline and order, and parents generally expect children to help with domestic tasks starting around age 5 (Barrow, 2008). Physical punishment by parents in the home is not uncommon and respect for authority is highly valued in children (Barrow, 2008; Cappa & Khan, 2011; Gopaul-McNicol, 1999). The emphasis on discipline and respect grows as children reach school age, with disobedience considered a marker of parental leniency (Barrow, 2008; Gopaul-McNicol, 1993).

Sample and Participants

Participants included children aged 4- to 15-years old ($M = 8$, $SD = 2$) attending government and assisted elementary schools across the islands. Prior to selection, schools were stratified by regional enrollment and regions with lower percentages of pupils in the population were less represented than regions with higher percentages. The full sample of 900 was composed of a national normative sample ($n = 700$) representatively blocked by grade, gender, and region plus a supplemental validity oversample ($n = 200$) enrolled in 75 elementary schools with 524 teachers. The 700 students from the normative sample were used for scale calibration and the oversample was included in structural and validity analyses. The sample was 49.7% male, with 39.9% African, 38.3% East Indian, and 21.7% mixed race/ethnicity. This is similar to national ethnic distributions in Trinidad and Tobago (i.e., 34.2% of African descent, 35.4% of East Indian descent, and 24.3% of mixed descent; Central Intelligence Agency, 2014).

Instruments

Classroom Social–Emotional Adjustment. ASCA is a teacher rating scale composed of behavioral indicators describing typical reactions in 24 classroom situational contexts. The classroom contexts encompass a range of situations, from responses to demanding learning tasks, to interactions with teacher and with other students, to conduct during play. The teacher records a student's observed behavior over a 2-month period by selecting any behavioral descriptions relevant to each context. Multiple descriptions can be endorsed to represent the child's behavior within a play, social, or learning context. For example, within the context where the student is reacting to teacher correction, the teacher may describe typical behavior as "improves for the moment but it does not last long," "accepts correction without fuss," "takes correction badly (sulky muttering, expression, etc.)," and/or "answers back aggressively, makes threats or creates a disturbance." In this way, each

context provides three to eight different indicators of behavior. As teachers find it easier to respond to descriptions that allow them to identify positive behaviors if appropriate (McDermott, 1993), 29 positive behavioral indicators (with prevalence greater than 50%) were included in the ASCA in an effort to avoid response bias associated with scales that present only negative behavior options (LeBoeuf, Fantuzzo, & Lopez, 2010). In U.S. normative research, three reliable dimensions (peer, learning, and teacher context problems) were found to define the contextual nature of problem behaviors in classrooms (McDermott et al., 2005). Appreciable validity evidence has been provided on relationships with external measures, classification accuracy, and structural invariance across age, gender, and ethnicity (McDermott, 1993, 1994).

Classroom Learning Behavior. The Learning Behaviors Scale (LBS; McDermott, 1999) is a teacher rating scale containing 29 items, reflecting variations of patterns with which children approach learning tasks. It was standardized on 1,500 students aged 5–17 years, and stratified by age, gender, and grade level according to the U.S. Census. The scale requires that teachers observe the child for no less than 50 days, whereafter behaviors are assessed on a 3-point scale. The instrument includes four subscales measuring competence motivation, attitude toward learning, attention–persistence, and strategy/flexibility. Convergent and divergent validity evidence were provided through consistent relationships between LBS subscales and criterion measures of intellectual functioning, classroom adjustment, and academic achievement. Internal consistency and interrater reliability were substantiated (McDermott, 1999; Worrell, Vandiver, & Watkins, 2001), and structural generality has been replicated in multiple contexts (Canivez & Beran, 2011; Canivez, Willenborg, & Kearney, 2006; Worrell et al., 2001).

Classroom Clinical Behavior. The Disruptive Behavior Disorders Rating Scale (DBDRS; Pelham, Gnagy, Greenslade, & Milich, 1992) is a teacher rating scale that uses criteria from the three disruptive behavior categories from the Diagnostic and Statistical Manual of Mental Disorders, revised Version III (DSM-III-R; American Psychiatric Association, 1987) to help classify clinical disorders. Evidence that this scale demonstrates adequate score stability and validity was documented with research with males in regular classrooms (Pelham et al., 1992) and special education settings (Pelham, Evans, Gnagy, & Greenslade, 1992).

Home Social–Emotional Behavior. The ASCA-Home Edition (ASCA-H; Watkins & McDermott, 2002) is a parent rating scale that, similar to ASCA, presents behaviors in situational contexts. The 202 behaviors, in 34 situations, are related to behaviors observable by parents in the home, with the parent marking any behaviors that they observe over a 2-month period (Watkins & McDermott, 2002). Evidence of structural validity and scale reliability was substantiated by Coffey (2006).

Academic Achievement. Oral reading fluency (ORF; Fuchs, Fuchs, Hosp, & Jenkins, 2001) is an individually administered curriculum-based academic measure that aggregates reading comprehension using reading fluency, or the number of words that a child is able to read accurately in 1 minute. Primary grades tend to experience a steeper growth curve for fluency, where the curve is negatively accelerated in later grades (Fuchs et al., 2001). Research has supported the predictive validity and clinical utility of ORF scores (Petscher & Kim, 2011), and the ORF has been found to correlate appropriately with other curriculum-based measures (Eckert et al., 2013)

Procedure

Data were collected as part of a consultation project agreement between a team based at the Pennsylvania State University and the Trinidad and Tobago Ministry of Education (Watkins et al., 2014). Guidance and Special Education Officers from the Ministry of Education received training

from the consulting team and were assigned to gather data over the course of one academic year. They were generally assigned to the educational divisions in which they already worked, and were paid an honorarium for each school with complete data. Additionally, teachers and parents received an honorarium for participating by completing the rating scales.

Exploratory Analysis. The sample was randomly divided into two subsamples, an exploratory subsample ($n = 500$) and confirmatory subsample ($n = 400$). The number of endorsed problem behaviors was totaled for each of the situational contexts. The resulting correlation matrix of the total scores for each of the 24 situational contexts was smoothed for positive semidefiniteness. MicroFACT (Waller, 2001) software produced a smoothed polychoric correlation matrix, employing two-stage maximum-likelihood estimation and least-squares approximation of the original matrix. Minimum average partialling (MAP; Velicer, 1976) was applied to suggest the number of factors to be retained. As per the recommendation of Snook and Gorsuch (1989) for scales featuring fewer than 40 variables, iterated common factor solutions were rotated toward simple structure according to varimax, equamax, and promax criteria. Each structure was evaluated using solution criteria that included (a) approximate simple structure as indicated by a maximized hyperplane count and context coverage; (b) at least four salient contexts per factor with loadings $\geq .40$ defined as salient; (c) adequately reliability (i.e., $\alpha \geq .70$); and (d) a structure with parsimonious and comprehensive coverage of the data and in accordance with leading research and theory (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Confirmatory Analysis. Based on the pattern of salient contexts marking each exploratory factor, factor solutions were submitted to maximum-likelihood estimation. The Satorra–Bentler scaled difference chi-square for nonnormal data was applied for the confirmatory subsample (Satorra & Bentler, 2001). Acceptable fit was defined by a root mean squared error of approximation (RMSEA) $\leq .08$ and comparative fit index (CFI) $\geq .90$ (Weston & Gore, 2006).

Scaling. The contexts associated with each respective factor were scaled through multiple-group IRT, applying generalized partial credit logistic models based on the normative sample ($N = 700$) to optimize representativeness of parameters. The resultant parameters were thereafter applied to estimate scores for the validity oversample ($N = 200$). Scores were produced via the Bayesian expected a posteriori (EAP) method, with the normative sample $M = 50$ and $SD = 10$. . In this case Bayesian EAP estimation sets the mean of the posterior distribution to 50, given the observed response pattern and the assumption of a normal population distribution, and has a smaller average error in the population than alternative estimators (Bock & Mislevy, 1982). Reliability was examined for factors using Cronbach's α, and assessment through overplots illustrated the distribution of test information (i.e., the inverse of measurement error) and measurement error.

External Validity. Product–moment correlations were used to evaluate the strength and direction of relations between scores on each ASCA scale and external criterion variables. Per the recommendations of Waterman, McDermott, Fantuzzo, and Gadsden (2012) for data nested within teachers, relationships were also assessed with hierarchical linear modeling (HLM), where ASCA scores were the group-mean centered predictors in two-level conditional models, denoting the percentage of between-children within-classroom variance accounted for by respective ASCA scales.

Results

Dimensionality

MAP for the 24 classroom contexts suggested that a minimum of two factors might be extracted from the smoothed polychoric matrix. The 1- through 5-factor models were assessed against the stated

Table 1
Dimensional Structure and Properties of Problem Behavior Contexts in the Adjustment Scales for Children and Adolescents

Component Situation Description[a]	Context Pattern Loadings[b] I	II	III	Communality	Context/ Scale r^c	Context/Scale Polyserial r^d
Peer contexts (coefficient $\alpha = .85^e$)						
Getting along with agemates	*.97*	.01	−.08	.85	.72	.82
Playing fairly	*.80*	−.06	.17	.79	.68	.78
Informal/unorganized play	*.76*	−.02	.06	.63	.64	.81
Respecting others' belongings	*.74*	−.02	.13	.68	.59	.72
Standing in line	*.63*	−.06	.18	.53	.53	.67
Telling the truth	*.60*	−.14	.32	.60	.59	.75
Standing one's own ground	*.55*	.37	−.10	.53	.52	.65
Has companions	*.46*	.27	.06	.47	.50	.64
Greeting teacher	*.41*	.19	−.06	.24	.33	.51
Teacher contexts (coefficient $\alpha = .70^e$)						
General manner with teacher	−.09	*.82*	−.04	.57	.52	.81
Seeking teacher help	−.37	*.79*	.33	.71	.50	.78
Talking to teacher	.11	*.61*	.05	.50	.48	.76
Valuing teacher attention	.33	*.49*	−.16	.37	.43	.74
Taking part in team games	.08	*.49*	.15	.42	.33	.64
Learning contexts (coefficient $\alpha = .86^e$)						
Working by self	.00	.09	*.84*	.81	.74	.87
Paying attention in class	−.02	.28	*.68*	.75	.68	.82
Caring for books, etc.	.21	−.11	*.66*	.57	.56	.72
Sitting at desk	.29	.01	*.58*	.66	.66	.80
Behaving in classroom	.39	−.03	*.52*	.67	.63	.78
Working with hands (art, shop)	.07	.17	*.49*	.44	.52	.70
Reaction to correction	.39	.07	*.45*	.66	.61	.77

[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 $k = 4$, where hyperplane count is maximized. Salient pattern loadings (≥.40) are italicized. $N = 500$ comprising the random exploratory analysis subsample.
[c]Each correlation reflects the relationship between the sum of observed behaviors for a given context and the sum of behaviors across contexts comprising a scale, where distributions were standardized to unit-normal form.
[d]Values are correlations between ordered categorical sums of behaviors within a given context and the continuous sum of behaviors across all contexts comprising a scale.
[e]Reliability is based on the exploratory subsample ($N = 500$).

criteria. The 3-factor, promax-rotated ($k = 4$) model emerged as the optimal solution having met all criteria, where Waller's (2001) goodness-of-fit index = .99 and root mean squared residual = .05. Models extracting more than three factors contained underidentified and unreliable dimensions and the 1- and 2-factor models compressed the 3-factor model into less meaningful composite factors.

Following Comrey's (1988) recommendations, one context with multiple salient loadings was excluded from subsequent analyses, thus retaining 21 contexts. Table 1 displays rotated pattern loadings, final communalities, product–moment context-scale correlations, and polyserial context-scale correlations. Coefficient α for each scale is also presented (see the centered headings). Based on contextual content and patterns of descending loadings, the scales were named peer context problems (nine contexts), teacher context problems (five contexts), and learning context problems

(seven contexts). Peer context problems correspond to behavior with agemates, teacher context problems to interactions with teachers, and learning context problems to learning and general classroom expectations. Scale intercorrelations are .56 between teacher and learning context problems, .69 between peer and learning context problems, and .45 between peer and teacher context problems.

The three-dimensional structure was evaluated with the confirmatory subsample. Model fit was adequate, where Satorra–Bentler χ^2 (186) = 402.89, CFI = .94, and RMSEA = .054 (90% CI = .047/.061).

Scaling and Reliability

The generalized partial credit threshold parameters for peer context problems ranged .73–3.07 ($M = 2.07$, $SD = .53$), slopes ranged .47–1.93 ($M = 1.14$, $SD = .50$), average information = .62, and the approximate maximum information = 4.62 at $\theta = 2.07$; the teacher context problems thresholds ranged 1.25–2.92 ($M = 2.10$, $SD = .60$), slopes .67–1.29 ($M = 1.01$, $SD = .27$), average information = .47, and maximum information = 1.67 at $\theta = 3.17$; and the learning context problems thresholds ranged 1.24–2.05 ($M = 1.54$, $SD = .26$), slopes ranged .77–1.51 ($M = 1.23$, $SD = .24$), average information = .64, and approximate maximum information = 4.15 at $\theta = 2.07$.

EAP (Thissen, Pommerich, Billeaud, & Williams, 1995) scaled scores (*SS*s) were generated for all students, with the normative sample centered at $M = 50$ and $SD = 10$. Figure 1 displays the overlap of total test information and measurement error for each scale. The illustrations indicate *SS*s will have practical utility from ~ 1/2 *SD* below the population mean for peer context problems and ~ 1 *SD* below the population mean for teacher and learning context problems, and continuing throughout the highest *SS*s on all scales. Peer, teacher, and learning contexts were internally consistent, yielding α coefficients of .85, .70, and .85, respectively.

External Validity

Table 2 displays relationships between the context scales and independent criterion measures. Under the columns for peer, teacher and learning contexts, it lists the conventional Pearson product–moment correlations (the nonparenthetical values) between respective ASCA context scores and scores from the various external criterion variables. The magnitude and directionality of these values generally comport with theoretical expectations. Peer and learning context problems had stronger relationships with DBDRS dimensions than did teacher context problems, indicating that peer and learning dimensions detect clinical disturbance in the classroom. Attention deficit and hyperactivity disorder and conduct problems dimensions on the ASCA-H were more correlated with peer and learning problems, suggesting that parents are more sensitive to reporting problems in these two contexts relative to the teacher context. Learning context problems had the expected stronger negative relationship with LBS dimensions, confirming that a higher level of problems in learning contexts corresponds to poor learning behavior. As expected, the learning context problems dimension was more related to ORF performance than were peer and teacher context problems.

Because the participant students were nested within specific classrooms, it was conceivable that the Pearson correlations could under- or overestimate the true nature of score relationships (Waterman et al., 2014). As a result, each relationship was reassessed through HLM modeling that first established the actual percentage of score variance that each external criterion variable conveyed that was between-students (the last column in Table 2) rather than between-teachers/classrooms and then determined the percentage of that between-students score variation that was accounted for by a given ASCA context scale (the parenthetical values of the first three columns). For example, viewing Table 2, the ordinary Pearson correlation between ASCA peer context problems and the external Inattention scores was .55. Ordinarily that would indicate that peer scores account for 30.3% of the

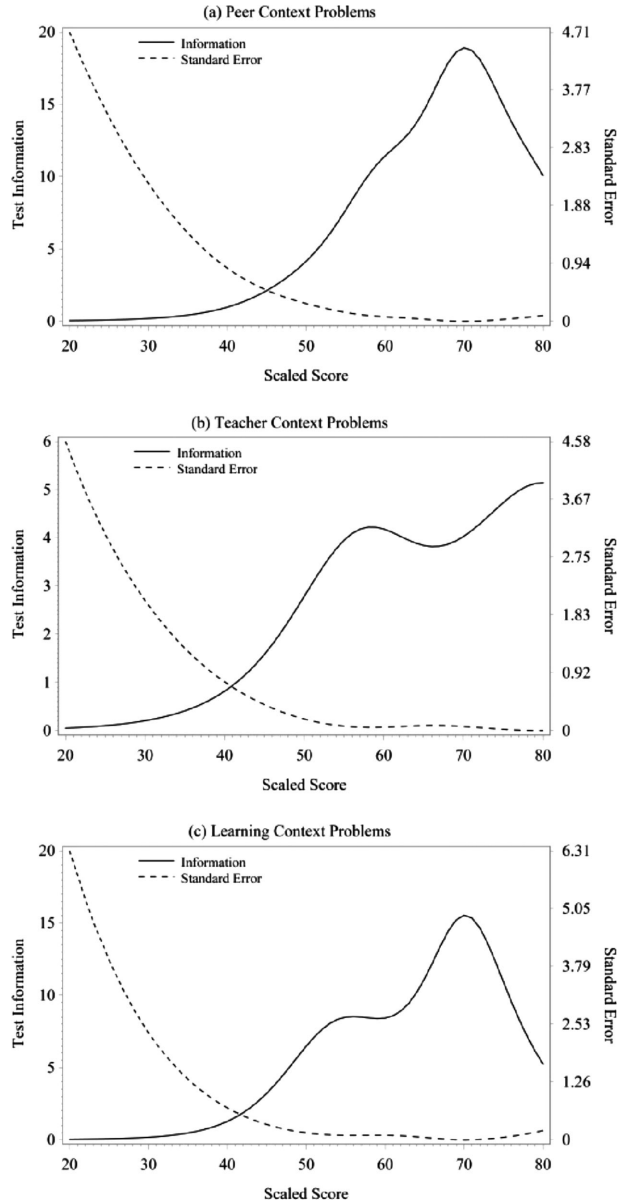

FIGURE 1. Distributions of estimated information functions and standard errors for ASCA problem contexts.

variance in inattention scores (coefficient of determination = .55 × .55 = .303). But HLM revealed that only 84.8% (not 100%) of the variance in inattention was legitimately score variation between students, meaning that 15.2% of inattention score variance was between teachers/classrooms and not indicative of individual student differences. Consequently, HLM further revealed that, when peer scores are applied as a predictor, they account for 49.9% (the tabled parenthetical value for inattention) of individual differences in inattention, not the lesser 30.3% value suggested by the Pearson correlation. In summary, overall Table 2 results show that ASCA scale scores are quite

effective in predicting true between-student differences on the criterion variables. Teachers appear to be more sensitive to recording peer and learning problems than teacher problems. Peer, teacher, and learning context problems account for relatively equal portions of variation in individual differences for children's reading achievement.

Demographic Trends

Table 3 presents the mean population distributions and standard deviations of peer, teacher, and learning context problems by gender and grade level in Trinidad and Tobago and Table 4 displays the distributions and standard deviations by gender and ethnicity. These show that normatively, no marked peaks or shifts in variation were discovered.

Discussion

The hypothesis that a representative sample from Trinidad and Tobago would generate the same three reliable dimensions of context problems discovered in U.S. national samples of the ASCA was confirmed. The core dimensions of peer context problems, learning context problems, and teacher context problems have emerged in a large representative sample of American children spanning aged 5–17 years, among Head Start children, and for children enrolled nationwide in various types of preschools through first grade (Bulotsky-Shearer et al., 2008; McDermott et al., 2005, 2014). These observations point to some universality of the dimensions in the United States, and now an important international extension is added to those observations.

Given the discovery of the same context dimensions in this sample and U.S. standardization samples, we thought it informative to make a cross-cultural comparison in that respect. Ethnographic research has indicated that adult authorities in Trinidad and Tobago are distinctly sensitive to child misbehavior, seeing it as a reflection of leniency or poor management (Barrow, 2008; Cappa & Khan, 2011; Gopaul-McNicol, 1993, 1999). This interpretation might suggest that Trinidad and Tobago teachers would be more sensitive to problem behaviors than U.S. teachers. To test this supposition, we generated a direct comparison by applying the U.S. scoring parameters for the context dimensions to the children in the Trinidad and Tobago normative sample. In this manner, we were able to contrast average levels of problem behavior across the United States and island children of the same age ranges.

Cross-nationally, the teacher context problems *SS* mean of 49.7 for Trinidad and Tobago was comparable to the U.S. mean of 50.0 ($p = .32$). On the other hand, the islands' mean *SS*s for both learning and peer context problems, 52.5 and 54.3, respectively, were higher than their U.S. counterparts of 51.3 ($p = .01$) and 52.1 ($p < .0001$). Thus, whereas Trinidad and Tobago teachers observe no more or less problem behavior than U.S. teachers in situations involving direct personal relationships with children; they are, in contrast, either more sensitive to problem behavior in situations involving peers or demands for learning, or else possibly such behaviors actually do manifest more often in Trinidad and Tobago.

The relationships between the ASCA context scores (as produced by teachers) and the ASCA-H scores (produced by parents) are overall low, with underactivity correlating near zero for all contexts. There is often a low-to-moderate correlation between teacher and parent ratings on behavior scales, with more agreement on externalizing behaviors than on internalizing ones (Achenbach, McConaughy, & Howell, 1987; van der Ende, Verhulst, & Tiemeier, 2012). This pattern is attributed to the more observable nature of externalizing behaviors, where they typically manifest as disruptive rather than passive or covert. The relationship roles (parent/child vs. teacher/student) and environmental circumstances (home vs. school) will further generate discrepancies between parents and teachers. There is also evidence that distortion of perceptions is particularly evident in cases of

Table 2
Relationships Between ASCA Context Scores and Concurrent Criterion Measures

Criterion Measure	Peer Contexts	Teacher Contexts	Learning Contexts	% Explainable Variance[a]
		Disruptive Behavior Disorder Rating Scale (teacher rating)		
Inattention ($n = 673$)	.55 (49.9)	.43 (24.2)	.69 (60.4)	84.8
Oppositional/defiant ($n = 630$)	.62 (67.5)	.29 (14.4)	.52 (49.3)	90.9
Impulsivity/overactivity ($n = 631$)	.61 (56.9)	.20 (2.0)	.51 (44.6)	80.0
		Oral reading fluency (direct assessment)		
Fall ORF mean of A and B passages ($n = 678$)	−.24 (4.5)	−.22 (26.9)	−.36 (21.6)	60.5
Winter ORF mean of A and B passages ($n = 709$)	−.26 (5.8)	−.24 (31.5)	−.36 (21.5)	59.1
Spring ORF mean of A and B passages ($n = 678$)	−.24 (4.5)	−.22 (26.5)	−.36 (21.6)	60.5
		Learning Behaviors Scale (teacher rating)		
Total score ($n = 755$)	−.52 (41.1)	−.51 (32.3)	−.68 (58.3)	83.7
Competence motivation ($n = 815$)	−.37 (14.5)	−.50 (23.9)	−.62 (51.4)	98.7
Attitude ($n = 811$)	−.43 (29.8)	−.52 (41.5)	−.61 (55.3)	87.7
Persistence ($n = 818$)	−.49 (37.5)	−.43 (27.1)	−.70 (57.5)	92.0
Strategy ($n = 797$)	−.53 (43.0)	−.31 (8.2)	−.53 (43.3)	72.3
		Adjustment Scales for Children and Adolescents—Home (parent rating)		
ADH ($n = 720$)	.26 (4.6)	.07† (1.3)	.25 (9.9)	88.7
Conduct problems ($n = 719$)	.29 (13.2)	.12 (10.9)	.25 (28.7)	97.9
Underactivity ($n = 719$)	.03† (0.3)	.06† (7.3)	.04† (5.2)	84.9

Note. 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 1—reduction in the intraclass correlation (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, ORF = oral reading fluency, ADH = attention deficit and hyperactivity disorder.

[a]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 $p < .001$.

Table 3
Mean Population Distribution of Contexts by Gender and Grade Level in Trinidad and Tobago

		Peer Contexts		Teacher Contexts		Learning Contexts	
Gender		*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
			Infant 1				
Male	($n = 50$)	51.0	*(9.3)*	49.2	*(10.2)*	52.0	*(8.8)*
Female	($n = 50$)	52.5	*(9.9)*	51.5	*(11.6)*	51.7	*(10.7)*
			Infant 2				
Male	($n = 50$)	53.8	*(10.3)*	50.8	*(10.7)*	51.8	*(10.6)*
Female	($n = 50$)	49.6	*(8.3)*	49.3	*(8.7)*	48.4	*(8.5)*
			Standard 1				
Male	($n = 50$)	52.7	*(10.3)*	51.6	*(9.6)*	53.2	*(10.2)*
Female	($n = 50$)	48.4	*(10.1)*	48.5	*(8.8)*	49.0	*(9.9)*
			Standard 2				
Male	($n = 50$)	49.6	*(9.5)*	50.5	*(9.5)*	50.4	*(9.0)*
Female	($n = 50$)	46.5	*(9.3)*	48.0	*(10.7)*	46.5	*(10.4)*
			Standard 3				
Male	($n = 50$)	51.4	*(11.9)*	51.9	*(11.0)*	52.3	*(10.2)*
Female	($n = 50$)	49.3	*(10.7)*	49.2	*(10.4)*	49.3	*(9.0)*
			Standard 4				
Male	($n = 50$)	49.4	*(9.1)*	48.6	*(8.9)*	48.6	*(9.4)*
Female	($n = 50$)	45.7	*(8.7)*	47.0	*(9.9)*	45.8	*(9.7)*
			Standard 5				
Male	($n = 50$)	52.8	*(11.9)*	50.6	*(9.4)*	51.8	*(10.8)*
Female	($n = 50$)	49.1	*(9.4)*	50.8	*(9.9)*	50.3	*(10.1)*
			Total				
Male	($n = 350$)	51.5	*(10.4)*	50.5	*(9.9)*	51.4	*(9.9)*
Female	($n = 350$)	48.7	*(9.7)*	49.1	*(10.0)*	48.7	*(9.9)*

Table 4
Mean Population Distribution of Contexts by Gender and Ethnicity in Trinidad and Tobago

		Peer Contexts		Teacher Contexts		Learning Contexts	
Gender		*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
			African descent				
Male	($n = 137$)	52.1	*(11.4)*	50.0	*(9.9)*	53.5	*(10.1)*
Female	($n = 132$)	49.8	*(10.1)*	49.4	*(10.1)*	49.2	*(10.7)*
			East Indian descent				
Male	($n = 126$)	50.0	*(9.6)*	50.7	*(9.6)*	49.3	*(9.0)*
Female	($n = 131$)	47.0	*(8.9)*	48.5	*(10.1)*	47.0	*(8.6)*
			Mixed descent				
Male	($n = 76$)	53.0	*(9.3)*	51.3	*(10.2)*	51.9	*(9.8)*
Female	($n = 73$)	49.5	*(9.6)*	49.0	*(9.4)*	49.7	*(8.9)*
			Total				
Male	($n = 350$)	51.5	*(10.4)*	50.5	*(9.9)*	51.4	*(9.9)*
Female	($n = 350$)	48.7	*(9.7)*	49.1	*(10.0)*	48.7	*(9.9)*

maternal depression and personal distress (De Los Reyes & Kazdin, 2005), such that ratings may reveal as much about a parent as a focal child. Situational specificity may also be a factor, where children truly behave quite differently in the home than in the school (Hoffenaar & Hoeksma, 2002).

The current research has certain limitations. Trinidad and Tobago's normative sample included elementary-age school children only, whereas the U.S. normative sample also included middle and secondary school students. Moreover, external validity analyses of school performance, although including various measures of clinical behavior and classroom approaches to learning, were limited in assessment of academic achievement because only a curriculum-based assessment of reading achievement was available. In this respect, it should be noted that standardized measures of achievement in Trinidad and Tobago do not presently exist. Further, although our ASCA work did not reach into upper levels of schooling, the larger project did yield preventive and intervention research that extended into high school and targeted various manifestations of clinical symptomatology (Watkins et al., 2014). Future research is intended that will mitigate the shortcomings by drawing wider age samples and developing more rigorous and representative assessments of school achievement.

Traditional child behavior rating scales, rather than yielding contextually sensitive dimensions, have focused on measurement of disturbances broadly represented across contexts with no reference to those contexts. This type of measurement typically produces two or more broad-brand behavioral dimensions borrowing nominally from Eysenck's (1953) extraversion–introversion versus neuroticism dichotomy. Subsequently, terminology describing this dichotomy shifted to an internalizing versus externalizing problems designation (Achenbach, 1966). The internalizing–externalizing model has shown robustness across age, sex, ethnicity, culture, informant type, instrument, and DSM diagnoses (Mash & Barkely, 2014; Lewis & Rudolph, 2014; Merrell, 2008).

Phenotypes Versus Situtypes

The ASCA also is capable of producing the more traditional broad-band dimensions. This outcome was first demonstrated by McDermott (1993) for the U.S. standardization sample. That research focused on ASCA's item-level rather than context-based aspects and established reliable dimensions of underactivity and overactivity. This two-dimensional structure has been generalized in other populations including Hispanic/Latino, Native American, and Canadian school children (Canivez & Beran, 2009; Canivez & Bohan, 2006; Canivez & Sprouls, 2010). The same overactivity and underactivity dimensions, as based on ASCA item-level data, were also confirmed for our Trinidad and Tobago national sample (McDermott et al., 2015, March). All of these dimensions are referred to as *phenotypes* because they describe the familiar surface syndromes and phenomenological forms of maladjustment, irrespective of the motivational contexts for that maladjustment.

In sharp contrast, the present study emphasizes the advantage of viewing problem behavior, not by its general pervasiveness as a phenotype, but rather as a function of its emergence within specific contexts. The dimensions derived through situation- or context-based assessments are termed *situtypes* (McDermott et al., 2005), from the Latin root, "situs," referring to situation or place. Traditional assessments tend to assume that children with a certain level on a phenotype dimension will react similarly in different situations. Such assessment is further inclined to treat contextual variation as random noise or measurement error (Mischel, 2004). This approach fails to acknowledge that any given type of behavior or pervasiveness of that behavior may be differentially motivated and thereby signal different pathways to intervention. For example, a student who manifests relatively high levels of phenotypically underactive and disengaged behavior may do so for a myriad of reasons. That same disengagement in the context of learning activities, as opposed to other contexts involving the teacher and peers, points toward intervention strategies that take advantage of teacher and/or peer assistance as exemplified by reciprocal peer tutoring and other peer-assisted learning interventions

(Leung, 2015). With information on both phenotypes and situtypes, practitioners can generate profiles that are more relevant, and provide a clearer understanding of underlying motivations for behavior disturbance and how best to mitigate them.

Conclusion

The standardization and validation of ASCA situtypes in Trinidad and Tobago has provided insight into circumstances that incentivize problem behavior and that may be used to target interventions. Scaled scores may be used to categorize performance levels, where *SS*s < 60 indicate *good* or *adequate adjustment*, 60–69 *at risk*, and ≥ 70 *maladjustment*. If applied, this information will allow the Ministry of Education to move beyond general interventions and on to interventions tailored to specific situational contexts. Through targeted interventions, more students will hopefully have the opportunity to reach their full academic potential and provide valuable contributions to the nation.

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. (1986). Teachers Report Form. Burlington, VT: Achenbach.

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.

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

Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. Applied Psychological Measurement, 6, 431–444.

Bulotsky-Shearer, R. J., Fantuzzo, J. W., & McDermott, P. A. (2008). An investigation of classroom situational dimensions of emotional and behavioral adjustment and cognitive and social outcomes for Head Start children. Developmental Psychology, 44, 139–154. doi:10.1037/0012-1649.44.1.139

Buss, K. A. (2011). Which fearful toddlers should we worry about? Context, fear regulation, and anxiety risk. Developmental Psychology, 47, 804–819. doi:10.1037/a0023227

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., & Beran, T. N. (2011). Learning Behaviors Scale and Canadian youths: Factorial validity generalization and comparisons to the U.S. standardization sample. Canadian Journal of School Psychology, 26, 193–208. doi:10.1177/0829573511416440

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., & 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., Willenborg, E., & Kearney, A. (2006). Replication of the Learning Behaviors Scale factor structure with an independent sample. Journal of Psychoeducational Assessment, 24, 97–111. doi:10.1177/0734282905285239

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

Central Intelligence Agency. (2014). The world factbook, 2013-2014, Trinidad and Tobago. Retrieved 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). The 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

De Los Reyes, A., & Kazdin, A. E. 2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study.Psychological Bulletin, 131, 483–509. doi:10.1037/0033-2909.131.4.483

Dominguez, X., Vitiello, V. E., Fuccillo, J. M., Greenfield, D. B., & Bulotsky-Shearer, R. J. (2011). The role of context in preschool learning: A multilevel examination of the contribution of context-specific problem behaviors and classroom process quality to low-income children's approaches to learning. Journal of School Psychology, 49, 175–195. doi:10.1016/j.jsp.2010.11.002

Eckert, T. L., Truckenmiller, A. J., Rymanowski, J. L., Koehler, J. L., Koenig, E. A., & Hier, B. O. (2013). Curricular assessment. In K. F. Geisinger (Ed.), APA handbook of testing and assessment in psychology: Testing and assessment in school psychology and education (Vol. 3, pp. 169–181). Washington, DC: American Psychological Association.

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

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

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

Hoffenaar, P. J., & Hoeksma, J. B. (2002). The structure of oppositionality: Response dispositions and situational aspects. Journal of Child Psychology and Psychiatry, 43, 375–385. doi:10.1111/1469-7610.00028

Joosten, A. V., Bundy, A. C., & Einfeld, S. L. (2012). Context influences the motivation for stereotypic and repetitive behaviour in children diagnosed with intellectual disability with and without autism. Journal of Applied Research in Intellectual Disabilities, 25, 262–270. doi:10.1111/j.1468-3148.2011.00663.x

Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates ' skills? Child Development, 82, 1768–1777. doi:10.1111/j.1467-8624.2011.01665.x

Kamphaus, R. W., & Reynolds, C. R. (2007). Behavior assessment system for children—Second edition (BASC-2): Behavioral and Emotional Screening System (BESS). Bloomington, MN: American Guidance Service.

Kern, L., & Clemens, N. H. (2007). Antecedent strategies to promote appropriate classroom behavior. Psychology in the Schools, 44, 65–75. doi:10.1002/pits.20206

LeBoeuf, W. A., Fantuzzo, J. W., & Lopez, M. L. (2010). Measurement and population miss-fits: A case study on the importance of using appropriate measures to evaluate early childhood interventions. Applied Developmental Science, 14, 45–53. doi:10.1080/10888690903510349

LeBuffe, P. A., Shapiro, V., & Naglieri, J. A. (2009). Devereux Student Strengths Assessment. Lewisville, NC: Kaplan.

Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. Journal of Educational Psychology, 107, 558–579. doi:10.1037/a0037698

Lewis, M., & Rudolph, K. D. (2014). Handbook of developmental psychopathology (3rd ed.). New York, NY: Springer.

Mash, E. J., & Barkely, R. A. (2014). Child psychopathology (3rd ed.). New York, NYGuilford.

McDermott, P. A. (1993). National standardization of uniform multisituational 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., Rhoad, A. M., Chao, J. L., Watkins, M. W., Worrell, F. C., & Hall, T. E. (2015, March). Trinidad and Tobago: National standardization of Adjustment Scales for Children and Adolescents. Poster presented at the 2015 Society for Research on Child Development Biennial Meeting, Philadelphia, PA.

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

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

Mischel W. (1973). Toward a cognitive social learning reconceptualization of personality. Psychological Review, 80, 252–283. doi:10.1037/h0035002

Mischel, W. (2004). Toward an integrative science of the person. Annual Review of Psychology, 55, 1–22. doi:10.1146/annurev.psych.55.042902.130709

Neugebauer, S. R. (2014). Context-specific motivations to read for adolescent struggling readers: Does the motivation for reading questionnaire tell the full story? Reading Psychology, 35, 160–194. doi:10.1080/02702711.2012.679171

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

Petscher, Y., & Kim, Y.-S. (2011). The utility and accuracy of oral reading fluency score types n predicting reading comprehension. Journal of School Psychology, 49, 107–129. doi:10.1016/j.jsp.2010.09.004

Powell, D. R., Burchinal, M. R., File, N., & Kontos, S. (2008). An eco-behavioral analysis of children's engagement in urban public school preschool classrooms. Early Childhood Research Quarterly, 23, 108–123. doi:10.1016/j.ecresq.2007.04.001

Robie, C., Schmit, M. J., Ryan, A. M., & Zickar, M. J. (2000). Effects of item context specificity on the measurement equivalence of a personality inventory. Organizational Research Methods, 3, 348–365. doi:10.1177/109442810034003

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

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. Journal of Applied Psychology, 80, 607–620. doi:10.1037/0021-9010.80.5.607

Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. Personnel Psychology, 65, 445–494. doi:10.1037/0021-9010.80.5.607

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

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

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

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

Visser, M., Kunnen, S. E., & van Geert, P. L. C. (2010). The impact of context on the development of aggressive behavior in special elementary school children. Mind, Brain, and Education, 4, 34–43. doi:10.1111/j.1751-228X.2009.01081.x

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., 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. W., & McDermott, P. A. (2002). The Adjustment Scales for Children and Adolescents—Home. Unpublished scale. The Pennsylvania State University.

Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. Counseling Psychologist, 34, 719–751. doi:10.1177/0011000006286345

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