



HHS Public Access

Author manuscript

Psychol Assess. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

Psychol Assess. 2016 May ; 28(5): 613–625. doi:10.1037/pas0000214.

Examining the Validity of the Adapted Alabama Parenting Questionnaire Parent Global Report Version

Eugene Maguin,

School of Social Work, The State University of New York at Buffalo, New York, United States

Thomas Nochajski,

School of Social Work, The State University of New York at Buffalo, New York, United States

David Dewit, and

Centre for Addiction and Mental Health, London, Ontario, Canada

Andrew Safyer

Department of Social Work, Adelphi University, New York, United States

Abstract

The purpose of the present study was to comprehensively examine the validity of an adapted version of the parent global report form of the Alabama Parenting Questionnaire (APQ) with respect to its factor structure, relationships with demographic and response style covariates, and differential item functioning (DIF). The APQ was adapted by omitting the Corporal Punishment and the other discipline items. The sample consisted of 674 Canadian and United States families having a 9–12 year old child and at least one parent-figure who had received treatment within the past five years for alcohol problems or met criteria for alcohol abuse or dependence. The primary parent in each family completed the APQ. The four factor CFA model of the four published scales used and the three factor CFA model of those scales from prior research were rejected.

Exploratory structural equation modeling was then used. The final three factor model combined the author-defined Involvement and Positive Parenting scales and retained the original Poor Monitoring/Supervision and Inconsistent Discipline scales. However, there were substantial numbers of moderate magnitude cross-loadings and large magnitude residual covariances.

Differential item functioning (DIF) was observed for a number of APQ items. Controlling for DIF, response style and demographic variables were related significantly to the factors.

Keywords

Alabama Parenting Questionnaire; parent global report; confirmatory factor analysis; differential item functioning

Correspondence concerning this article should be addressed to Eugene Maguin, 219 Parker Hall, Buffalo Center for Social Research, University at Buffalo, Buffalo, NY 14214-8004. emaguin@buffalo.edu.
Eugene Maguin, School of Social Work, The State University of New York at Buffalo, New York; Thomas Nochajski, School of Social Work, The State University of New York at Buffalo, New York; David Dewit, Centre for Addiction and Mental Health, University of Western Ontario, London, Ontario, Canada; Andrew Safyer, Department of Social Work, Adelphi University, New York, New York.

Parenting behavior, meaning the relatively specific behaviors that parents perform in relation to their children—parenting practices as opposed to parenting styles (Locke & Prinz, 2002), long has occupied a key risk factor role in the developmental course of child and adolescent oppositional and conduct problems (e.g., Simons-Morton, Chen, Hand, & Haynie, 2008), delinquency (e.g., Hoeve et al., 2009; Loeber & Stouthamer-Loeber, 1986), and substance use (e.g., Feldstein & Miller, 2006). Changing parenting behavior has, therefore, been a key target in intervention programs for these same issues (e.g., August, Lee, Bloomquist, Realmuto, & Hektner, 2003; Lundahl, Risser, & Lovejoy, 2006) as well as for other problems such as ADHD (e.g., Wells et al., 2000) where the parent intervention components are intended to have a mediated effect on child behavior through improvements in parenting behavior.

Early work with delinquent youth (e.g., Glueck & Glueck, 1950) identified lack of involvement, poor supervision, rejection or lack of warmth, harsh and/or inconsistent discipline practices as important correlates of delinquency (see Loeber & Stouthamer-Loeber, 1986 for a meta-analysis of this topic). The notion of poor supervision was later refined and expanded into “monitoring” by Patterson and colleagues (Patterson, 1982). These five dimensions, often conceptualized as components of the superordinate constructs of nurturance and discipline (Locke & Prinz, 2002), form the core of parenting practices.

Of the different measures of parenting practices (see, for example, Hurley, Huscroft-D’Angelo, Trout, Griffith, & Epstein, 2014; Locke & Prinz, 2002 for reviews), the subject of this study, the Alabama Parenting Questionnaire (APQ; Shelton, Frick, & Wootton, 1996), seems to be frequently cited and frequently used, with Google Scholar reporting 438 citations to it as of July 23, 2013. In addition, the APQ has also been translated into at least nine languages, including Chinese and Persian (<http://www.psyc.uno.edu/Frick%20Lab/APQ/apq-translations9-16-2011.pdf>, accessed: 7/23/2013).

The Alabama Parenting Questionnaire

The design of the APQ, Frick (Frick, Christian, & Wootton, 1999; Shelton et al., 1996) wrote, was informed by the reviews of the parenting and child disruptive behavior and delinquency literature, particularly the meta-analysis by Loeber and Stouthamer-Loeber (1986), and was adapted from measures developed at the Oregon Social Learning Center (Capaldi & Patterson, 1989). The APQ consists of a global version and brief report version with each version having a parent form and a child form. The parent global report version is the focus of this article and it consists of 42 items grouped into five scales: Involvement, 10 items; Positive Parenting, 6 items; Poor Monitoring/Supervision, hereafter, Poor Monitoring, 10 items; Inconsistent Discipline, 6 items; and Corporal Punishment, 3 items; plus seven heterogeneous items describing other discipline practices such as reasoning, ignoring, loss of privileges, time-out, and extra work. The Corporal Punishment items range in severity from spanking to hitting with an object. The items are behaviorally focused and developmentally appropriate to the intended 6–13 year old age range. Alpha reliabilities were reported by both Shelton, Frick, and Wootton (1996) and Frick, Christian, and Wootton (1999) and the ranges for these values are .77 to .82 for mother-reported Involvement, .77 to .80 for Positive

Parenting, .49 to .67 for Poor Monitoring, .55 to .67 for Inconsistent Discipline, and .25 to .46 for Corporal Punishment.

Both Shelton et al. (1996) and Frick et al. (1999) investigated the association between the APQ scales and demographic and response style covariates. Shelton et al. found significant correlations of $-.16$, $-.26$, and $.25$ between child age and, respectively, mother-reported Involvement, Positive Parenting and Poor Monitoring and a non-significant correlation of $.02$ for Inconsistent Discipline for a sample of clinic-referred and community children aged 6–13 years. Using a sample of 6–17 year old clinic-referred youth grouped into three age categories, Frick et al. found that mother-reported Involvement and Positive Parenting decreased significantly and Poor Monitoring increased significantly from the 6–8 year old age group to the 13–17 year old age group. Inconsistent Discipline increased but not significantly. Shelton et al. also found that parent SES had a significant correlation of $.16$ with mother-reported Involvement and correlations of $-.05$ to $-.08$ with Positive Parenting, Poor Monitoring, and Inconsistent Discipline. Finally, Shelton et al. also found that the MMPI *K* scale, a measure of defensiveness, had correlations of $.19$ with mother-reported Involvement, $.15$ with Positive Parenting, $.07$ with Poor Monitoring and $-.24$ with Inconsistent Discipline, which was the only significant correlation. Taken together, the correlations with defensiveness indicate that some respondents were presenting themselves favorably.

Measurement Structure Results

A search of the PsychInfo database for articles citing Shelton et al. (1996) or naming the APQ in the record identified eight English language articles referencing seven studies that reported either factor analysis or confirmatory factor analysis results for the parent global report measure.¹ Table 1 summarizes sample characteristics, analytic methods, and key findings of the studies. Five of the seven datasets were normal population samples, one was a sample of ADHD-identified children and one was a sample of aggressive and nonaggressive children. Although the children's ages spanned by the seven datasets were 4 years through 18 years, 6 through 9 years were the most frequently included ages. Although five of the studies administered the 42 item APQ, only Zlomke, Lamport, Bauman, Garland, and Talbot (2014) and, perhaps, Wells et al. (2000) factor analyzed the 42 item set. The remaining studies either did not administer the seven other discipline items or omitted them from their analysis. Randolph and Radey (2009) deleted 13 items of the 42 items, including nine of the ten Poor Monitoring items, for high skewness and kurtosis, prior to their factor analysis. Zlomke et al. also found this problem as more than 50% of respondents endorsed the same extreme response for 13 items, including six Poor Monitoring items and all three Corporal Punishment items.

¹Two studies (Clerkin, Marks, Policaro, & Halperin, 2007 and de la Osa, Granero, Penelo, Domènech, & Ezpeleta, 2013) administered a parent report version that had been modified to be more appropriate for the age range of the children being rated, which was 3–6 years for Clerkin et al. and 3 years for Osa et al. Essau, Sasagawa, and Frick (2006) administered the child version only. Lastly, two studies (Wells, Epstein, Hinshaw et al., 2000 and Molinuevo, Pardo, & Torrubia, 2011) administered the parent report and child report versions. Factor analyses of the child report version are not reviewed because, in our view, those results do not directly bear on the factor structure of the parent version. Whether children and their parents structure parenting behavior similarly is an important question bearing on the validity of the APQ. This study cannot address that question; Wells et al. and Molinuevo et al. do.

The modal analysis sequence was a principal components extraction of the number of factors indicated by a scree plot of eigenvalues followed by an orthogonal, Varimax, rotation. Variations included parallel analysis and a minimum average partial test (Elgar, Waschbusch, Dadds, & Sigvaldason, 2007); principal axis (common factor model) extraction (Randolph & Radey, 2009; Zlomke, Lamport, Bauman, Garland, & Talbot, 2014); and oblique rotation (Molinuevo, Pardo, & Torrubia, 2011; Randolph & Radey, 2009; Zlomke et al. (2014). The uniform result, except for Robert (2009) who extracted five factors apriori and Zlomke et al. (2014) who found four factors, was three factors.

Broadly, these analyses found that the Involvement and Positive Parenting scales combined and the Poor Monitoring scale and Inconsistent Discipline scale remained as separate factors. Robert (2009), the exception, found that the Involvement items divided into two factors, one of which combined with the Positive Parenting items. However, the location of the Corporal Punishment (CP) items varied considerably. Hinshaw et al. (2000) found that they combined with the Inconsistent Discipline items; Zlomke et al. (2014) found they combined with the other discipline items as the fourth factor; and Robert (2009) found they formed a distinct (fifth) factor. However, Elgar et al. (2007), who alone provided complete factor pattern data, found that the CP items did not have above-threshold (.40) loadings on any factor. The spanking item, the highest loading CP item, loaded .15 on their Inconsistent Discipline factor and the slap and hit with object items had loadings of .07 or less on all factors.

Perhaps because of the age of their sample, children aged 4–9 years, Elgar et al. (2007) found that about one-third of the 35 items, particularly items from the Involvement and Poor Monitoring scales, did not have above-threshold loadings on any factor. Although most apparent for Elgar et al., even studies providing partial factor pattern data showed some evidence of cross-loadings or items, particularly items 28 and 29 (Hinshaw et al., 2000; Robert, 2009), changing factors; however, interpreting these results is complicated by a differing number of items analyzed (Zlomke et al., 2014) or the number of factors specified (Robert, 2009).

Children of Alcoholics in Two Countries

The sample to be used here differs from the two studies that used clinical samples: August et al. (2003) who selected children with aggression problems and Hinshaw et al. (2000) who selected children with ADHD problems. This study selected on adults with diagnosed or treated levels of alcohol problems and who were the parents or guardians of age-range children. It is well known that the alcoholic parents or guardians are at increased risk for multiple classes of psychiatric disorders (e.g., Goldstein, Dawson, Chou, & Grant, 2012; Sher, 1997; Schuckit et al., 2003). The partners of the alcoholic parents or guardians are themselves more likely to have an anxiety or major depression disorder (e.g., Maes et al., 1998, who did not include personality disorders). Lastly, the children are at increased risk for wide range of disorders (Johnson, Cohen, Kasen, & Brook, 2006; Morgan, Desai, Potenza, & Marc, 2010) and, more broadly, externalizing behavior problems and attention problems (Barnow, Schuckit, Smith, Preuss, & Danko, 2002; Loukas, Zucker, Fitzgerald, & Krull, 2003) as well as internalizing problems (Preuss, Schuckit, Smith, Barnow, & Danko,

2002). In addition, the children are at increased risk for the proximal effects of parental intoxication, including possibly witnessing alcohol-related, within-family violence (Jennison & Johnson, 1998). Certainly, some of the children in the August et al. and Hinshaw et al. studies had parents with one or multiple psychiatric disorders. However, we think that this sample gives a more diverse, complex, and inter-related comorbidity topography than do either of the two aforementioned studies. Because this sample selects on parents, the parents' ratings include variation due to possible parent disorders, variation due to child behavior as well as nonspecific parent variation. Thus we would expect this sample to better represent the ranges of scores for parent practices items.

Second, this sample is bi-national, with families residing in either Ontario, Canada or New York State and, therefore, offers the opportunity for a small-scale test of the generalizability of the APQ. Although Canada and the U.S. share a border, similar national origins of immigrant-settlers, and, except for Quebec, a common language, each country has evolved its own political, legal, and social culture within which family life and child development take place.

Summary

Rather than the five factor design structure, a three factor solution that combined the Involvement plus Positive Parenting scales and left the Poor Monitoring and Inconsistent Discipline scales as separate factors was consistently found, indicating a high level of configural consistency across studies. Placement of the Corporal Punishment items was inconsistent however, and this was likely due to between-study heterogeneity with respect to sample characteristics. Although the lack of detailed results do not afford a complete picture, instances of items cross-loading or changing factors were observed. Two of the normal sample studies noted that certain items had high percentages of response scale endpoints in their frequency distributions, indicating floor or ceiling effects. Whether these effects were also present for the clinical samples is not known; however, the presence of these effects underscores the need for analytical methods better matched to the categorical, sometimes skewed nature of item responses and to the plausible assumption of correlated factors.

Even though the analytical methods used served the aims of each study, the analytical heterogeneity, combined with the lack of result details, cannot contribute to assessments of the invariance of the APQ across samples, an important consideration given the widespread distribution of the APQ. Thus bi- or multi-cultural/national studies of the APQ are needed to assess invariance. Likewise, assessments of invariance under demographic variation are also needed. The finding that a response style variable was related to scale scores needs replication, as it suggests that some respondents may bias their ratings. Although a considerable amount of work has been done, there is a need, we believe, to apply more powerful analytical methods, combined with data-appropriate assumptions, to a large, bi-national dataset of families likely having wider range of clinical issues to address questions about both structure and invariance.

Present Study

The purpose of the present study is to examine the validity of an adapted version of the Alabama Parenting Questionnaire with respect to its factor structure, relationships with demographic and response style covariates, and differential item functioning (DIF). The corporal punishment and other discipline practices items were omitted. The dataset to be used is a multi-racial, bi-national sample of children having a parent/guardian who had received treatment within the past five years for alcohol problems or met criteria for alcohol abuse or dependence. The multi-problem nature of COA families offers a perspective on the validity of the APQ not found in prior studies. In addition the analytical methods will be selected to better match the frequency distributions of the items and to better characterize the structure of the APQ. Such efforts are important because parenting practices occupies a central role in children's development and is a central mediational target in interventions. This effort is particularly important because the APQ is a well-known and widely used measure of parenting practices.

The following questions will be addressed to further this goal.

1. Are the original scale definitions for the subscales included in the adapted version supported by a confirmatory factor analysis?
2. Does modeling the Involvement and Positive Parenting scales as a single factor provide a better, more interpretable fit to the observed data than modeling those scales as separate factors?
3. What are the relationships between the demographic and response style variables and factors and is there evidence of differential item functioning with the demographic and response style variables?

Methods

Sample

The sample for this study was the 674 families who participated in the Families Working Together (FWT) study. FWT was a randomized clinical trial of the Strengthening Families Program (Kumpfer, DeMarsh, & Child, 1989) that was conducted in the southern Ontario, Canada cities of Toronto, Barrie, Aurora, Kitchener, and London and the Buffalo, New York (Western New York) area. Families were eligible to participate in FWT if (a) the family had a child between 9 and 12 years of age; (b) the child had a parent figure who had received treatment for alcohol problems within the last five years or met the criteria for alcohol abuse or dependence based on the Research Diagnostic Criteria (Spitzer, Endicott, & Robbins, 1978); and (c) the family had not participated in a family skills building program in the past year. The parent/guardian with alcohol problems was not required to be currently living with the child or involved in their care, just that the person had been responsible at some time for the child's care. In cases where the parent/guardian with alcohol problems was no longer in the family, the screening for alcohol treatment or problems was based on the report of the parent in the family. Age-eligible children were excluded if the child, based on the parent's description, was judged to be unable to attend productively to the session content or that

their behavior was so disruptive that it would interfere with other children's involvement and/or prevent the group leaders from presenting the scheduled program content. If a family had more than one child in the 9–12 year age range, one of the age-eligible children was selected at random to be the target child. Although all of the target child's parents/guardians were invited to participate in the study, one parent/guardian, usually the one who called to inquire about the study, was designated as the primary parent/guardian. This study used pretest data from the primary parent/guardian.

As Table 2 shows, compared to U.S. families, Canadian families were significantly more likely to be of white, European ethnic background, married or married but now divorced or separated, better educated and in a much better economic situation as judged by the percentages who reported receiving direct cash and/or non-cash welfare benefits or reported incomes under 15,000 dollars per year, equal purchasing power assumed.

The primary parents had significant levels of multiple problems. In the total sample, just over three-fifths reported alcohol treatment or alcohol problems at recruitment and on the Short Michigan Alcohol Screening Test (Selzer, Vinokur, & Rovijen, 1975. Two-thirds reported a score at or above the clinical cutoff of 16, which is indicative of probable depression, on the Center for Epidemiological Studies-Depression scale (Radloff, 1977). The mean total score on the Aggression Questionnaire (Buss & Perry, 1992) was 2.8 ± 0.7 . Although Aggression Questionnaire norms have been published, they are proprietary. However, Buss and Perry (1992) reported means of 2.35 ± 0.59 and 2.68 ± 0.57 for first and second year of college females and males, respectively, and Webster (2014) reported a mean of 2.65 ± 0.62 for female and male college students combined. Significantly larger percentages of U.S. primary parents reported alcohol problems and U.S. primary parents had significantly higher Aggression Questionnaire scores.

Measures

Alabama Parenting Questionnaire (APQ)—The published, 42-item parent global report version (Shelton et al., 1996) was modified by deleting the three Corporal Punishment scale and the seven other disciplinary practices items.² The replacement items assessed non-abusive or non-aggressive discipline practices, verbally abusive behavior and spanking. The word *take* was added to item 15, “You drive your child to a special activity”, after the study started because a number of families reported using buses or taxis rather than cars.³ The published five-point response format, 1 (*never*), 2 (*almost never*), 3 (*sometimes*), 4 (*often*), 5

²The Principal Investigators judged that two of the three Corporal Punishment items (“You slap your child ...” and “You hit your child with a belt, switch, or other object ...” constituted events that mandated reporting to Child Protective Services in New York State for New York families or to its Ontario counterpart for Ontario families. Further, it was judged that such reporting at the time of recruitment might well lead a parent to withdraw from an intervention intended to improve their parenting. The items that replaced both the Corporal Punishment items and the other discipline practices items included items assessing use of timeout and loss of privileges, verbal abuse such as yelling or swearing at the child or name-calling, and a physical punishment item, bare-handed spanking of the bottom.

³The change was made at approximately the 48th recruited family. The effect of the change was investigated for all analyses by a) adding a dummy covariate for the wording change and by b) deleting those cases using the published wording. The dummy covariate was significant in all analyses for (a). Overall fit measures improved by 0.0% to 6.2%. Changes in unstandardized factor and residual covariances and regression coefficients ranged from 0.0 to 0.068 (absolute value) with the majority being less than 0.01. Although the wording change yielded a different response distribution resulting in changed values for model coefficients, the fundamental conclusions of the analyses were not negated.

(*always*), was used. When the study was designed, the APQ was among the few published parenting practices measures that assessed intervention targeted parental behaviors.

Response styles—Two scales from the FAM-III General scale (Skinner, Steinhauer, & Santa-Barbara, 1983; Skinner, Steinhauer, & Sitarenios, 1999) were used to assess a social desirability response style and a defensive response style. Social Desirability ($\alpha = .78$) was assessed by seven items and Defensiveness ($\alpha = .71$) was assessed by eight items. The FAM-III uses a four-point response format, 1 (*strongly agree*), 2 (*agree*), 3 (*disagree*), 4 (*strongly disagree*). Scale scores were the mean of item scores. A large value means *more* of the construct.

Psychological functioning and alcohol problems—Parents completed the Center for Epidemiological Studies-Depression scale (CES-D: Radloff, 1977), Aggression Questionnaire (AQ: Buss & Perry, 1992), and Short Michigan Alcoholism Screening Test (SMAST: Selzer, Vinokur, & Rovijen, 1975). The CES-D is a widely-used, self-report measure of depression for general population use. It is a 20 item measure with a four-point response format, 0 (*not at all*), 1 (*a little*), 2 (*somewhat*), 3 (*a lot*). A score of 16 or more is indicative of significant depressive symptoms. Total score alpha was .80. The AQ is a 29 item measure with four scales (physical aggression, verbal aggression, hostility, and anger, plus a total score) and using a five-point response format, 1 (*strongly agree*), 2 (*agree*), 3 (*neither agree nor disagree*), 4 (*disagree*), 5 (*strongly disagree*). A total score ($\alpha = .91$) was computed as the mean of items. Items were reversed prior to use so that a larger score indicated more aggression. The SMAST is a 13-item assessment of a range of alcohol problems. Items are scored as *Yes* or *No* and scores of three or more are indicative of abuse or dependence.

Procedures

Families were recruited from reviews of treatment agency records, fliers in social service agencies, presentations to treatment agency staff and treatment groups, and local newspapers advertisements. Interested families were screened for eligibility and read a description of the study. At the pretest assessment session, parents and children were read the informed consent and assent forms, questions were answered, and the forms were signed. Each person was interviewed separately and individually by trained interviewers. Interviews lasted from 45 to 75 minutes. Primary parents received 15 CAD/25 USD for their assessment and children received 10 CAD/10 USD for their assessments. Transportation or reimbursement for the same was provided if needed. This study was reviewed and approved by the University of Buffalo Institutional Review Board, the Centre for Addiction and Mental Health Research Ethics Board, and the Tri-Hospital Research Ethics Board in Kitchener, Ontario.

Nonresponse

Nonresponse on the APQ items was relatively minor, as only 8.8% of parents ($n = 59$) omitted a response to one or more items. Five percent omitted exactly one item and 1.9% omitted exactly two items. Although significantly more Canadian parents omitted responses

to one or more items than did U.S. parents, 11.7% vs. 5.9%; $\chi^2 = 6.50$ (1, $N = 674$), $p = .011$, nonresponse on an item-by-item basis was unrelated to sample demographic variables.

Plan of Analysis

In overview, confirmatory factor analysis (CFA) was used to test two specific factor structure models in the adapted version: the original four factor model of the four scales and the three factor model obtained by combining items in the Involvement and Positive Parenting scales as indicated from prior work. Exploratory structural equation modeling (ESEM: see Asparouhov & Muthén, 2009 for the technical development and Marsh et al., 2009, and Marsh et al., 2010, for applied work) then was used to investigate alternative factor solutions suggested by the scree plot of eigenvalues and, for the selected factor structure, to identify using a specification search residual covariances and investigate their effect on factor loadings. This final model was then used to investigate differential item functioning for selected demographic and response style covariates and to estimate relationships between the covariates and factors.

All CFA and ESEM analyses were conducted using Mplus 7.0 (Muthen & Muthen, 1998–2012). Cases with missing data for observed dependent variables were included by default. Except as noted, default settings were used in all analyses. Items were modeled as categorical and the WLSMV estimator was used as recommended by simulation work examining parameter bias in CFA models (Flora & Curran, 2004; Moshagen & Musch, 2014; Rhemtulla, Brosseau-Liard, & Savalei, 2012). As suggested by Hu and Bentler (1998), model fit was summarized by the chi-square test statistic and the multiple approximate fit indices provided in Mplus: Confirmatory Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Weighted Root Mean Square Residual (WRMR). Following Yu (2002), the following thresholds were used to indicate a good fitting model: CFI and TLI, .95, RMSEA, .05, WRMR, 1.00. Comparisons between nested models were made using the Mplus DiffTest option because the differences in chi-square values for the WLSMV estimator do not follow a chi-square distribution. Thus the difference test p-values from the model comparisons will be biased to an unknown degree because the difference test presumes the data fit the less restrictive model. Modification indices (MIs) and standardized expected parameter change (SEPC) values were examined to identify sources of model misfit.

ESEM was used to investigate alternative factor structures because it offers a number of advantages compared to alternatives. Compared to exploratory factor analysis (EFA), ESEM allows items to load on all factors but provides the same range of fit statistics and detailed residuals information including modification indices as does CFA. Like CFA, ESEM also allows specification of residual covariances and permits nested model comparisons via a chi-square difference test. Although an EFA analysis can be done within a CFA, the resulting factor loadings cannot be rotated to simple structure, which prevents comparison of alternative structures. However, ESEM does not permit, as does CFA, the a priori specification of non-zero values of loading coefficients for specific items on specific factors.

A forward specification search method using the chi square equivalent of a Bonferroni adjusted p-value (Green, Thompson, & Babyak, 1998) was used to identify possible residual

covariance modifications in the ESEM models and to identify items exhibiting DIF. The critical chi-square value was computed as $p = .05$ divided by the possible number of modifications for the element under study. In addition to a large MI/SEPC, covarying items had to have a plausible shared semantic or logical meaning beyond that implied by the common factor (Gerbing & Anderson, 1984). Modification of residual covariances proceeded iteratively: the best single residual covariance was identified and freed and the model re-estimated. DIF testing also proceeded iteratively. Covariates were individually regressed on the factors and the item with the largest above threshold p value was freed and the model re-estimated.

Results

Univariate Analysis

The item frequency distributions of all ten items on the Poor Monitoring scale and five of the six items on the Positive parenting scale, as well as one item each on the Involvement and Inconsistent Discipline scales, 17 of the 32 total items, were right-triangular, J , in shape with large endorsement percentages for the most positive response category (see Table 3). Poor Monitoring scale items had the most extreme values of both skewness, 0.70 to 2.15, and kurtosis, -0.51 to 4.41.

Confirmatory factor analysis

The four factor CFA model was that implied by the measure's authors: all items in a scale loaded only on the factor of the same name and no residual error covariances were allowed. The data did not fit the model ($\chi^2 = 1378.81$, $df = 458$, $p < 0.001$, CFI = 0.925, TLI = 0.919, RMSEA = 0.055, and WRMR = 1.559). Factor loadings were all significant at the .001 level and standardized absolute values ranged between .38 and .75 for Involvement, .44 and .85 for Positive Parenting, .39 and .81 for Poor Monitoring, and .45 and .70 for Inconsistent Discipline. Involvement correlated .86 with Positive Parenting, $-.36$ with Poor Monitoring, and $-.32$ with Inconsistent Discipline. Positive Parenting correlated $-.29$ with Poor Monitoring and $-.21$ with Inconsistent Discipline. Lastly, Poor Monitoring correlated .42 with Inconsistent Discipline. All factor correlations were significant at $p < .001$.

The second model combined the Involvement and Positive Parenting scales into a new factor, Involvement+Positive Parenting, as per the several reviewed factor analyses, and left the Poor Monitoring and Inconsistent Discipline factors unchanged. The data also did not fit this model ($\chi^2 = 1463.05$, $df = 461$, $p < 0.001$, CFI = 0.918, TLI = 0.912, RMSEA = 0.057, WRMR = 1.628). All factor loadings were significant at the .001 level. Standardized factor loadings for Involvement+Positive Parenting ranged between .37 and .85 while those for Poor Monitoring and Inconsistent Discipline were the same, $\pm .005$, as for the four factor model. Involvement+Positive Parenting correlated $-.34$ with Poor Monitoring and $-.27$ with Inconsistent Discipline. Poor Monitoring correlated .42 with Inconsistent Discipline. All factor correlations were significant at $p < .001$.

Compared to the four factor model, unstandardized factor loadings for Involvement scale items on the Involvement+Positive Parenting factor increased by less than 1.4 percent while

loadings for Positive Parenting scale items increased by 10 to 15 percent. The unstandardized variance of the Involvement+Positive Parenting factor was 0.523 compared to 0.569 for Involvement and 0.715 for Positive Parenting. In R-squared terms, combining the factors reduced the R-squares of items on those two factors by six to eight percent while leaving the R-squares for Poor Monitoring and Inconsistent Discipline items virtually unchanged. Relative to the four factor model, there was a greater under-estimation of covariances among both Involvement items and Positive Parenting items but an over-estimation of covariances between Involvement and Positive Parenting items.

Of the 32 cross-loadings with MIs greater than 10, approximately the .001 level, for the four factor model, 14 involved the six Inconsistent Discipline items and 11 involved the ten Poor Monitoring items. In both cases, the cross-loadings were with the Involvement and Positive Parenting factors. Of the 26 residual covariances with MIs greater than 10, equal numbers involved items on the same factor, mainly Poor Monitoring, or on different factors, mainly Involvement and Inconsistent Discipline.

Exploratory Structural Equation Modeling

Factor Models—To better investigate the APQ measurement structure, we turned to ESEM. Although a scree plot of eigenvalues suggested the knee to be at the fourth factor, we examined three, four, and five factor models. As would be expected, fit statistics improved as the number of factors extracted increased (Three factor: $\chi^2 = 1240.79$, $df = 403$, $p = 0.0000$, RMSEA = 0.056, CFI = 0.932, TLI = 0.916, WRMR = 1.191. Four factor: $\chi^2 = 964.98$, $df = 374$, $p = 0.0000$, RMSEA = 0.048, CFI = 0.952, TLI = 0.936, WRMR = 0.988. Five factor: $\chi^2 = 761.06$, $df = 346$, $p < 0.0000$, RMSEA = 0.042, CFI = 0.966, TLI = 0.952, WRMR = 0.827). However, a side-by-side comparison of the factor loadings for the three models yielded the rather stunning conclusion that the three factor model was the most interpretable model—despite its having the worst fit.

Why? In the four factor model, two Involvement scale items, 15 (Drive/take to special activity) and 26 (Attend meetings at school), and one Poor Monitoring scale item, 29 (Don't tell child where you are going), all with loading absolute values of 0.4–0.5 coalesced to form the new fourth factor. Six other items, Involvement scale items 4, 14, 23, Positive Parenting item 13, and Poor Monitoring items 24 and 28, had loadings in the 0.2 absolute value range on this factor. Factor loadings are standardized because factor variances and categorical item variances are fixed at 1.00. In the five factor model, items 15 and 26 were joined by Involvement scale item 4 (Volunteer to help with special activities), on the fourth factor and item 29 was joined by Poor Monitoring scale item 28 (Do not check child comes home when supposed to), both with loadings of 0.5–0.6, to form the fifth factor. In addition five other Involvement scale items, 7, 9, 11, 14 and 23, had cross-loadings in the 0.2–0.3 range on the fourth factor and a Poor Monitoring scale item, 24, had a cross-loading in the same range on the fifth factor.

In the three factor model, the Involvement and Positive Parenting scale items combined to form a single factor, Involved_Parenting, the Poor Monitoring scale items loaded on a single factor and the Inconsistent Discipline scale items loaded on a single factor. Primary factor

loadings, except for item 29, were in the 0.3 to 0.8 range. There were, however, a substantial number of small cross-loadings: 35 in the 0.1–0.2 range and 19 in the 0.2–.03 range.

Residual covariances—Using the three factor model, we next examined the MIs for the residual covariances and found that 27 had MI values of 10 or greater with the largest being 130.7. Fourteen of the 27 involved items on the Involved_Parenting factor, six involved items on the Poor Monitoring factor and another six were cross-factor residual covariances between Involved_Parenting items and either Poor Monitoring or Inconsistent Discipline items. To examine the impact of these residual covariances on model fit and factor loading, we systematically freed residual covariances provided they exceeded the Bonferroni threshold of .0001 (.05/496 possible covariances), equivalent to a chi square value of 15.14, and seemed to share a common logical or semantic meaning beyond that of the factor(s). A total of 11 residual covariances were freed (see Table 4). Eight of the eleven involved Involved_Parenting items, two involved Poor Monitoring items and one involved an Involved_Parenting and a Poor Monitoring item. After doing so, the fit statistics for the final model was $\chi^2 = 820.66$, $df = 392$, $p = 0.0000$, RMSEA = 0.040, 90% CI [0.036, 0.044], CFI = 0.965, TLI = 0.956, WRMR = 0.932. The factor loadings are reported in Table 4. As shown there, the factor loadings (absolute values) of items on their primary factor were in the 0.6 and above range with only four items, 15, 2, 32, and 25, having loadings in the 0.3 to 0.4 range and one, 29, with a loading below 0.3. Of the 64 possible cross-loadings, none were above 0.20 in absolute value and 19 were in the 0.10 to 0.20 range. Involved_Parenting correlated -0.27 ($p < .001$) with Poor Monitoring and -0.13 ($p < .05$) with Inconsistent Discipline. Poor Monitoring correlated 0.32 ($p < .001$) with Inconsistent Discipline.

Although freeing the 11 residual covariances significantly improved the fit of the model from a chi-square of 1240.79 to one of 820.66 (difference test $\chi^2 = 425.57$, $df = 11$, $p = .001$), their cumulative effect on model parameters depended on the parameter type. Factor correlations changed by between .002 and .011 in absolute terms. Loadings for items with large loadings decreased from the initial to the final models, most by less than 5%. Item 29, which decreased 51%, and items 2 and 13, which increased 13% and 11%, respectively, were exceptions. Loadings for items with small loadings both increased and decreased, sometimes by very large percentages if the initial value was very small. Freeing identified residual covariances, although markedly improving fit, did not change our conclusions about the preferred model or the salience of cross-loading items.

Differential Item Functioning

The latent variables were regressed on five parent covariates, U.S. resident, white ethnicity, parent education, marital status, received cash or non-cash benefits, two child covariates, female child and child age, and two response style covariates, social desirability and defensiveness. Each covariate was analyzed separately. Parent education and marital status were categorized as shown in Table 2. White was defined as white only versus all else. A Bonferroni corrected threshold of .00156 (.05/32, $\chi^2 = 10.01$) was set for continuous and dichotomous covariates and one of .00078 (.05/64, $\chi^2 = 11.29$) was set for trichotomous covariates, e.g., parent education. If one contrast was significant for a polychotomous covariate, both contrast terms were entered in the regression.

Except for female child, all covariates exhibited DIF. The most involved covariates were U.S. resident, education, and white. The most commonly identified items were 32 (Child at home without adult supervision), 15 (Drive/take to special activity), and 29 (Don't tell child where you are going). A positive covariate regression coefficient denotes an upward shifted response distribution, i.e., an increased likelihood of an *Often* or *Always* response, relative to that predicted by the factor-on-covariate regression coefficient. Thus, for item 32, for example, parents who had attended college or had children older in the study age range had upward shifted distributions relative to parents who had not graduated from high school or who had younger children. On the other hand, U.S. resident parents or parents receiving benefits had downward shifted response distributions relative to Canadians or to persons not receiving benefits, i.e., less frequently leaving a child without supervision.

Response Style and Covariate Regressions

Table 6 shows the results of the final regressions of the factors, along with any identified DIF items, on each demographic or response style covariate. As shown there, greater social desirability and defensiveness were significantly associated with increased Involved_Parenting and Inconsistent Discipline scores and with decreased Poor Monitoring scores. Persons with higher levels of social desirability and defensiveness, therefore, tended to present a more favorable view of these dimensions of their parenting behavior. In addition, each parenting behavior factor was associated significantly with one or more demographic characteristics. Decreased involved_parenting was associated with married/common-law parents relative to never married parents, white ethnicity parents, and older children. Worse monitoring was associated with Canadian residency, having been or currently married relative to never married, white ethnicity parents, and the target child being male or older. Lastly, greater discipline inconsistency was associated with U.S. residency.

For a given covariate, comparing coefficients from models with DIF included and without DIF included reveals that unmodeled DIF had a variable but sometimes large impact on covariate-factor relationships. Twenty-two of the 30 covariate-factor regression coefficients decreased in magnitude and eight of those coefficients either decreased in significance or became not significant. Only eight coefficients increased in magnitude and just two of those became significant but none increased in significance.

Discussion

The purpose of the present study was to conduct a comprehensive study of the validity with respect to factor structure and invariance of an adapted version of the Alabama Parenting Questionnaire that omitted the Corporal Punishment items. We found that a number of the items had either floor or ceiling effects as moderate to large percentages of respondents endorsed positive endpoint responses, as did both Randolph and Radey (2009) and Zlomke et al. (2014). To better estimate correlations, given such distributions, we used current, state-of-the-art methods. We found that the data did not fit a four factor or a three factor confirmatory model of the four published scale definitions. Furthermore, we found the three factor model fit worse than the four factor model. Both models showed substantial numbers

of small to moderate cross-loadings and moderate to large residual covariances, indicating complex inter-item and item-factor relationships.

We examined three, four, and five factor exploratory structural equation models (ESEMs) and selected the three factor model, which combined the original Involvement and Positive Parenting scales but left the Poor Monitoring scale and the Inconsistent Discipline scale intact, as the preferred model, even though it fit more poorly than either of the alternative models. The three factor model was preferred because the alternative models split the Involvement scale items into two factors, one consisting of in-home activities with the child and the other consisting of mainly out-of-home activities such as taking the child to activities or attending school-related functions, and also defined a doubleton factor of two Poor Monitoring scale items, as was also noted by both Wells et al. (2000) and Robert (2009), thereby yielding a less obviously meaningful factor definition. Our results also document the complexity of the APQ in terms of the numerous instances of both significant cross-loadings, which were especially apparent in Elgar et al.'s (2007) results, and residual covariances.

We found evidence of differential item functioning (DIF) principally between the demographic covariates of U.S. resident, White, and parent education and several Involved Parenting factor items. The specific pattern indicates that out-of-home activities were differentially reported by parents who had more education, were married, were Canadian, were white or were not receiving cash benefits. Thus a clear effect of higher income and social stability. Modeling DIF resulted in reductions 19% to 31% in the covariate-factor regression coefficients, indicating that the DIF-implicated items wielded a measure of influence. After controlling for DIF, the demographic covariates had generally small and often nonsignificant relationships with the three factors. U.S. resident, white, and child age most often had significant relationships with the factors and the Poor Monitoring factor most often had significant relationships with demographic covariates. Child age, which had a medium magnitude relationship with Poor Monitoring, was the exception to the generally small regression coefficient magnitudes. As such, our results for child age match both Shelton et al. (1996) and Frick et al. (1999). Other demographic covariates examined here have not been previously studied.

We found little evidence of DIF involving the response style variables. However, after modeling DIF, we found that the response style variables, social desirability and defensiveness, had significant, small to moderate sized relationships with all three factors, positive for Involved Parenting and negative for both Poor Monitoring and Inconsistent Discipline. Shelton et al. (1996) reported a similar pattern of results, although their values were typically numerically smaller, particularly for the Poor Monitoring scale.

One possible explanation of the relationships with the response style variables is acquiescence bias (Weijters, Geuens, & Schillewaert, 2010) in which parents, some more than others, presented themselves in a favorable light. Our results as well as those of Zlomke et al. (2014), which indicated a high prevalence of positive endpoint responses, particularly on the Poor Monitoring scale items might be an example of this effect. An alternative explanation is a methods factor in which, perhaps, some parents simply answered each

question more or less similarly. Although we did not test this possibility, it could be done by means of a bifactor model.

Participation in the samples used by prior analyses was defined by the clinical status of the children. Participation in this study was defined by the clinical status of the children's parents. Although this sample offers a different perspective on the APQ, the breadth and magnitude of the difference relative to the two clinical sample studies reviewed (August et al., 2003; Wells et al., 2000) is difficult to determine. August et al. (2003) did not report on the types or levels of symptoms of their parents. Epstein et al. (2000) compared parents participating in the study described by Hinshaw et al. (2000) and Wells et al. (2000) to matching community families, none of whose children met ADHD criteria, and found that the parents of children diagnosed with ADHD had significantly elevated levels of ADHD symptoms.

Limitations

That we analyzed a version of the APQ that did not include the Corporal Punishment items makes comparisons with prior analyses difficult because we do not know how the results would have been different had those items been included. A useful framework for comparing our results with prior work is that of measurement consistency rather than invariance to deemphasize the quantitative element and with the focus restricted to the configural and metric levels. Our results cannot speak to the number of factors that would have resulted. That would depend on the specific patterns of set-wise correlations among the corporal punishment items and between those items and the other factor-defined sets of items. Our results do show strong evidence of configural consistency for the combining of the Involvement and Positive Parenting scales and for the Poor Monitoring scale. Although our results show strong evidence of configural consistency for the Inconsistent Discipline scale items, the unknown effect of the corporal punishment items, i.e., result like Elgar et al. (2007) or like Wells et al. (2000), diminishes the strength of the consistency characterization for Inconsistent Discipline scale.

Given the differences in analysis methods, a useful measure of metric consistency is simply the similarity in items ranked by their loading on their factor as assessed by the correlation between analyses. Doing this, we found that our rankings correlated between .73 and .87 with those from prior analyses (excluding Robert, 2009) for the involvement plus positive parenting items, between .75 and .90 for the monitoring items, and .77 to .81 for the inconsistent discipline items only.

Our results with respect to differential item functioning, response style, and covariate relationships, particularly those for the Involved Parenting and Poor Monitoring factors, offer important new information that, we believe, would have a strong probability of replication in a dataset that included the Corporal Punishment items. However, because the placement of the Corporal Punishment is ambiguous in prior studies, the results for the Inconsistent Discipline factor may have less replicability.

In our differential item functioning analyses we chose to test only whether the item intercepts differed as a function of the covariate, e.g., Canada versus the U.S. These tests

assumed that configural and metric invariance had been satisfied. We tested only the intercepts because the other tests of invariance required a multiple group model resulting in within group Ns of 337, at best, and too small given the number of parameters to be estimated. Thus we believe that our findings of differential functioning must be regarded as provisional. However, given an adequate sample size, ESEM could have been employed to test the required multiple group models as illustrated by Marsh et al. (2009) and Marsh et al. (2010).

The ESEM analyses and the subsequent specification searches used to identify residual covariances and differential item functioning capitalized on the specific pattern of correlations found for this sample. Although we used a Bonferroni adjusted procedure to limit the extent of sets of searches, a close replication of our results, even with a similarly selected sample and this adapted version, may be unlikely. Thus it is important that other investigators completely analyze the APQ to determine the extent to which our results replicate and that these results offer a starting point.

Conclusions

Investigators considering using the APQ in etiological or prevention/intervention research programs need to consider their research questions, child age and presenting problems, and the planned analysis sophistication. The APQ's defined scales and their associated items have strong content validity for externalizing and delinquent behavior but less is known about their validities for internalizing behavior or other dependent variables. To this point we surveyed studies citing Shelton et al. (1996) as to the dependent variable set used. We found very few studies that only assessed internalizing behavior. A very large majority assessed some form of externalizing behavior alone or in combination with internalizing behavior as behavior problems. Likewise, we found relatively few studies that used only normal child samples. Like Randolph and Radey (2009) and Zlomke et al. (2014) found, we suspect other investigators may find a high positive endpoint endorsement percentages, particularly for the Poor Monitoring items. Possibly because of an acquiescence bias or a methods factor but also possibly because the true prevalence is low. The same is likely to be true as child age decreases. In addition, it is not known whether the APQ adequately samples the parenting behavior domain for increasingly younger children (see Verhoeven, Junger, Van Aken, Dekovic, & Van Aken, 2007, for a conceptualization of parenting dimensions during toddlerhood). That an investigator modified and omitted APQ items before administering it to parents of 3 to 6 year old children underscores this point (see Footnote 1). We believe that the adapted version of the APQ might well be useful for work with parents having neglect or abuse issues. Our review of articles using the APQ in substantive investigations revealed that the APQ is most often analyzed as observed scale scores rather than as latent variables with items as indicators and that the scales are analyzed separately. Even when latent variable models are used, an APQ scale is used as an indicator, along with scales from one or more other relevant measures. When a latent variable analysis is desired, it may be useful to consider forming parcels of items rather than working with items directly.

Acknowledgments

This research funded by National Institute of Alcohol Abuse and Alcoholism grant R01 AA1164701A1

References

- Asparouhov T, Muthén B. Exploratory structural equation modeling. *Structural Equation Modeling*. 2009; 16:397–438.10.1080/10705510903008204
- August GJ, Lee SS, Bloomquist ML, Realmuto GM, Hektner JM. Dissemination of an evidence-based prevention innovation for aggressive children living in culturally diverse, urban neighborhoods: The Early Risers Effectiveness Study. *Prevention Science*. 2003; 4:271–286.10.1023/A:1026072316380 [PubMed: 14598999]
- Barnow S, Schuckit M, Smith TL, Preuss U, Danko G. The real relationship between the family density of alcoholism and externalizing symptoms among 146 children. *Alcohol and Alcoholism*. 2002; 37:383–387.10.1093/alcalc/37.4.383 [PubMed: 12107042]
- Buss AH, Perry M. The Aggression Questionnaire. *Journal of Personality and Social Psychology*. 1992; 63:452–459.10.1037/0022-3514.63.3.452 [PubMed: 1403624]
- Capaldi, DM.; Patterson, GR. Psychometric properties of fourteen latent constructs from the Oregon Youth Study. New York, NY: Springer-Verlag Publishing; 1989.
- Chassin L, Pillow DR, Curran PJ, Molina BS, Barrera M. Relation of parental alcoholism to early adolescent substance use: A test of three mediating mechanisms. *Journal of Abnormal Psychology*. 1993; 102(4):3–19.10.1037/0021-843X.102.1.3 [PubMed: 8436697]
- Elgar FJ, Waschbusch DA, Dadds MR, Sigvaldason N. Development and validation of a short form of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies*. 2007; 16:243–259.10.1007/s10826-006-9082-5
- Epstein JN, Conners CK, Erhardt D, Arnold LE, Hechtman L, Hinshaw SP, Vitiello B. Familial aggregation of ADHD characteristics. *Journal of Abnormal Child Psychology*. 2000; 28:585–594.10.1023/A:1005187216138 [PubMed: 11104319]
- Essau CA, Sasagawa S, Frick PJ. Psychometric properties of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies*. 2006; 15:597–616.10.1007/s10826-006-9036-y
- Feldstein SW, Miller WR. Substance use and risk-taking among adolescents. *Journal of Mental Health*. 2006; 15:633–643.10.1080/09638230600998896
- Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*. 2004; 9:466–491.10.1037/1082-989X.9.4.466 [PubMed: 15598100]
- Frick PJ, Christian RE, Wootton JM. Age trends in association between parenting practices and conduct problems. *Behavior Modification*. 1999; 23:106–128.10.1177/0145445599231005
- Gerbing DW, Anderson JC. On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*. 1984; 11:572–580.10.1086/208993
- Glueck, S.; Gluck, ET. *Unraveling juvenile delinquency*. Cambridge, MA: Harvard University Press; 1950.
- Goldstein RB, Dawson DA, Chou SP, Grant BF. Sex differences in prevalence and comorbidity of alcohol and drug use disorders: Results from wave 2 of the national epidemiologic survey on alcohol and related conditions. *Journal of Studies on Alcohol and Drugs*. 2012; 73:938–950.10.1037/t04807-000 [PubMed: 23036212]
- Green SB, Thompson MS, Babyak MA. A Monte Carlo investigation of methods for controlling Type I errors with specification searches in structural equation modeling. *Multivariate Behavioral Research*. 1998; 33:365–383.10.1207/s15327906mbr3303_3 [PubMed: 26782719]
- Hinshaw SP, Owens EB, Wells KC, Kraemer HC, Abikoff HB, Arnold LE, Wigal T. Family processes and treatment outcome in the MTA: Negative/ineffective parenting practices in relation to multimodal treatment. *Journal of Abnormal Child Psychology*. 2000; 28:555–568.10.1023/A:1005183115230 [PubMed: 11104317]
- Hoeve M, Dubas JS, Eichelsheim VI, van der Laan PH, Smeenk W, Gerris JRM. The relationship between parenting and delinquency: A meta-analysis. *Journal of Abnormal Child Psychology*. 2009; 37:749–775.10.1007/s10802-009-9310-8 [PubMed: 19263213]
- Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*. 1998; 3:424–453.10.1037/1082-989X.3.4.424

- Hurley KD, Huscroft-D' Angelo J, Trout A, Griffith A, Epstein M. Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures. *Journal of Child and Family Studies*. 2014; 23:812–823.10.1007/s10826-013-9733-2
- Jennison KM, Johnson KA. Alcohol dependence in adult children of alcoholics: Longitudinal evidence of early risk. *Journal of Drug Education*. 1998; 28:19–37.10.2190/BRRQ-W96E-UGJN-GA9R [PubMed: 9567578]
- Johnson JG, Cohen P, Kasen S, Brook JS. A multiwave multi-informant study of the specificity of the association between parental and offspring psychiatric disorders. *Comprehensive Psychiatry*. 2006; 47:169–177.10.1016/j.comppsy.2005.05.005 [PubMed: 16635644]
- Kumpfer, K.; DeMarsh, J.; Child, W. Strengthening families program: Children's skills training curriculum manual, parent training manual, children's skill training manual, and family skills training manual (Prevention Services to Children of Substance-abusing Parents). Salt Lake City: Department of Health, Alcohol and Drug Research Center; 1989.
- Locke LM, Prinz RJ. Measurement of parental discipline and nurturance. *Clinical Psychology Review*. 2002; 22:895–929.10.1016/S0272-7358(02)00133-2 [PubMed: 12214330]
- Loeber, R.; Stouthamer-Loeber, M. Family factors as correlates and predictors of juvenile conduct problems and delinquency. In: Tonry, M.; Morris, N., editors. *Crime and Justice*. Vol. 7. University of Chicago Press; 1986. p. 29-149.
- Loukas A, Zucker RA, Fitzgerald HE, Krull JL. Developmental trajectories of disruptive behaviour problems among sons of alcoholics: Effects of parent psychopathology, family conflict, and child undercontrol. *Journal of Abnormal Psychology*. 2003; 112:119–131.10.1037/0021-843X.112.1.119 [PubMed: 12653420]
- Lundahl B, Risser HJ, Lovejoy MC. A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*. 2006; 26:86–104.10.1016/j.cpr.2005.07.004 [PubMed: 16280191]
- Maes HHM, Neale MC, Kendler KS, Hewitt JK, Silberg JL, Foley DL, Meyer JM, Rutter M, Simonoff E, Pickles A, Eaves LJ. Assortive mating for major psychiatric diagnoses in two population-based samples. *Psychological Medicine*. 1998; 28:1389–1401. [PubMed: 9854280]
- Marsh HW, Muthén B, Asparouhov T, Lüdtke O, Robitzsch A, Morin AJS, Trautwein U. Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*. 2009; 16:439–476.10.1080/10705510903008220
- Marsh HW, Lüdtke O, Muthén B, Asparouhov T, Morin AJS, Trautwein U, Nagengast B. A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*. 2010; 22:471–491.10.1037/a0019227 [PubMed: 20822261]
- Molina BSG, Chassin L. The parent-adolescent relationship at puberty: Hispanic ethnicity and parent alcoholism as moderators. *Developmental Psychology*. 1996; 32(4):675–686. doi.org/10.1037/0012-1649.32.4.675.
- Molinuevo B, Pardo Y, Torrubia R. Psychometric Analysis of the Catalan Version of the Alabama Parenting Questionnaire (APQ) in a Community Sample. *Spanish Journal of Psychology*. 2011; 14:944–955.10.5209/rev_SJOP.2011.v14.n2.40 [PubMed: 22059338]
- Morgan PT, Desai RA, Potenza, Marc N. Gender-related influences of parental alcoholism on the prevalence of psychiatric illnesses: Analysis of the National Epidemiologic Survey on Alcohol and Related Conditions. *Alcoholism: Clinical and Experimental Research*. 2010; 34:1759–1767.10.1111/j.1530-0277.2010.01263.x
- Moshagen M, Musch J. Sample size requirements of the robust weighted least squares estimator. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*. 2014; 10:60–70.10.1027/1614-2241/a000068
- Muthen, LK.; Muthen, BO. *Mplus User's Guide*. 7. Los Angeles, CA: Muthén & Muthén; 1998–2012.
- Patterson, GR. *Coercive family process*. Eugene, OR: Castalia; 1982.
- de la Osa N, Granero R, Penelo E, Domènech JM, Ezpeleta L. Psychometric properties of the Alabama Parenting Questionnaire-Preschool Revision (APQ-Pr) in 3 year-old Spanish preschoolers. *Journal of Child and Family Studies*. 2014; 23:776–784.10.1007/s10826-013-9730-5

- Preuss UW, Schuckit MA, Smith TL, Barnow S, Danko GP. Mood and anxiety symptoms among 140 children from alcoholic and control families. *Drug and Alcohol Dependence*. 2002; 67:235–242.10.1016/S0376-8716(02)00076-5 [PubMed: 12127194]
- Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977; 1:385–401.10.1177/014662167700100306
- Randolph KA, Radey M. Measuring parenting practices among parents of elementary school-age youth. *Research on Social Work Practice*. 2011; 21:88–97.10.1177/1049731509353048
- Rhemtulla M, Brosseau-Liard PÉ, Savalei V. When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*. 2012; 17:354–373.10.1037/a0029315 [PubMed: 22799625]
- Robert, CJ. Doctoral dissertation. 2009. Parenting Practices and Child Behavior in Mexico: A Validation Study of the Alabama Parenting Questionnaire. (Retrieved from ProQuest Dissertations & Theses Full Text). UMI Number: 3352810
- Schuckit MA, Smith TL, Barnow S, Preuss U, Luczak S, Radzinski S. Correlates of externalizing symptoms in children from families of alcoholics and controls. *Alcohol and Alcoholism: International Journal of the Medical Council on Alcoholism*. 2003; 38:559. doi:org/10.1093/alcalc/agg116.
- Selzer ML, Vinokur A, Rovijen L. A self-administered short Michigan alcoholism screening test (SMAST). *Journal of Studies on Alcohol*. 1975; 36:117–126. [PubMed: 238068]
- Sher KJ. Psychological characteristics of children of alcoholics. *Alcohol Health and Research World*. 1997; 21:247–254. [PubMed: 15706777]
- Shelton KK, Frick PJ, Wootton J. Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology*. 1996; 25:317–329.10.1207/s15374424jccp2503_8
- Simons-Morton B, Chen R, Hand LS, Haynie DL. Parenting behavior and adolescent conduct problems: Reciprocal and mediational effects. *Journal of School Violence*. 2008; 7:3–25.10.1300/J202v07n01_02 [PubMed: 19777074]
- Skinner HA, Santa-Barbara J, Steinhauer PD. The Family Assessment Measure. *Canadian Journal of Community Mental Health*. 1983; 2:91–105.
- Skinner HA, Steinhauer PD, Sitarenios G. Family Assessment Measure (FAM) and Process model of family functioning. *Journal of Family Therapy*. 2000; 22:190–210.10.1111/1467-6427.00146
- Smith M. Measures for assessing parenting in research and practice. *Child and Adolescent Mental Health*. 2011; 16:158–166.10.1111/j.1475-3588.2010.00585.x
- Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry*. 1978; 35:773–782.10.1001/archpsyc.1978.01770300115013 [PubMed: 655775]
- Verhoeven M, Junger M, Van Aken C, Dekovic M, Van Aken MAG. Parenting during toddlerhood - Contributions of parental, contextual, and child characteristics. *Journal of Family Issues*. 2007; 12:1663–1691.10.1177/0192513X07302098
- Webster GD, DeWall CN, Pond RS Jr, Deckman T, Jonason PK, Le BM, Bator RJ. The Brief Aggression Questionnaire: Psychometric and behavioral evidence for an efficient measure of trait aggression. *Aggressive Behavior*. 2014; 40:120–139.10.1002/ab.21507 [PubMed: 24115185]
- Weijters B, Geuens M, Schillewaert N. The stability of individual response styles. *Psychological Methods*. 2010; 15:96–110.10.1037/a0018721 [PubMed: 20230106]
- Wells KC, Epstein JN, Hinshaw SP, Conners CK, Klaric J, Abikoff HB, Wigal T. Parenting and family stress treatment outcomes in attention deficit hyperactivity disorder (ADHD): An empirical analysis in the MTA study. *Journal of Abnormal Child Psychology*. 2000; 28:543–553.10.1023/A:1005131131159 [PubMed: 11104316]
- Yu, CY. Doctoral dissertation. 2002. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. (Retrieved from ProQuest Dissertations & Theses Full Text). UMI Number: 3066425
- Zlomke KR, Lampion D, Bauman S, Garland B, Talbot B. Parenting adolescents: Examining the factor structure of the Alabama Parenting Questionnaire for adolescents. *Journal of Child and Family Studies*. 2014; 23:1484–1490.10.1007/s10826-013-9803-5

Table 1
 Studies Reporting Measurement Structure Results for Alabama Parenting Questionnaire—Parent Global Version

Study	Location	Child age (yrs)	N	Sample definition	Items administered/Items analyzed	Analysis method	Loading cutoff	Number of factors
August et al. (2003)	U.S.	5–6	448	Aggressive + Controls	42/35 ^b	PC- Varimax	NS	3
Elgar et al. (2007)	Australia	4–9	1,402	Normal Population	35 ^b /35	PC- Varimax	.40	3
Hinshaw et al. (2000); Wells et al. (2000)	U.S.	7–9	579	ADHD ^a	42/NS	PC- Varimax	.30	3
Molinuevo, Pardo, & Torrubia (2011)	Spain	10–15	364	Normal Population	35 ^b /f/35	PC- Oblique	.30	3
Randolph & Radey (2009)	U.S.	6–10	790	Normal Population	42/29 ^c	PA- Promax	.30	4
Robert (2009)	Mexico	11–12	862	Normal Population	42 ^d /e/35 ^b	PC- Varimax	NS	5, a priori
Zlomke et al. (2013)	U.S.	11–18	311	Normal Population	42/42	PA- Oblimin	NS	4

Note. NS = Not Stated. PC = Principal Components. PA = Principal Axis (common factor).

^a ADHD classification based on checklist score.

^b Other discipline practices deleted from administered or analyzed item set.

^c 13 items deleted for non-normality.

^d Four point response format used.

^e Items 28 and 29 reworded to remove double negative.

^f Item 28 reworded to remove double negative.

Table 2

Demographic Characteristics of the Total Sample and Canadian and U.S. Subsamples

Demographic Characteristic	Canada (<i>n</i> = 334)	U.S. (<i>n</i> = 340)	Total sample (<i>N</i> = 674)
Female parent	88.6%	87.6%	88.1%
Parent age: <i>M</i> (<i>SD</i>)	39.2 (5.7)	38.7 (6.7)	38.9 (6.4)
Female child	47.6%	45.9%	46.7%
Child age: <i>M</i> (<i>SD</i>)	10.7 (1.2)	11.0 (1.3)	10.9 (1.3)
Ethnicity ***			
White	80.5%	29.2%	54.5%
Black	3.3%	63.9%	32.0%
Aboriginal, Hispanic, or Asian	7.6%	5.6%	6.6%
Multiple	8.5%	5.3%	6.9%
Marital status ***			
Never married	9.5%	34.5%	22.2%
Separated/Divorced/Widowed	45.5%	38.7%	42.1%
Married/Common-law	44.9%	26.8%	35.7%
Education *			
Not a high school graduate	26.2%	25.7%	26.0%
High school graduate	14.9%	24.0%	19.5%
Any college	58.8%	50.3%	54.5%
Not employed ***	36.2%	56.8%	46.7%
Receive noncash benefits ***	35.7%	75.6%	56.0%
Income under \$15,000 ***	28.1%	60.5%	44.4%
Alcohol treatment or problems ***	50.6%	74.7%	62.8%
3+ problems on SMAST *	58.3%	66.3%	62.4%
CES-D score of 16 or more	69.6%	64.1%	66.8%
Aggression Questionnaire Total **	2.7 (0.8)	2.9 (0.6)	2.8 (0.7)

Note. SMAST = Short Michigan Alcohol Screening Test. CES-D = Center for Epidemiological Studies-Depression scale.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 3

Item Statistics for Alabama Parenting Questionnaire (Parent Report) (N = 674)

Item	Positive endpoint	Mean	SD	Skew	Kurtosis
Scale: Involvement					
1) Have friendly talk with child	31.0%	4.01	0.824	-0.472	-0.092
4) Volunteer to help with special activities	16.9%	3.21	1.177	-0.182	-0.575
7) Play games/do fun things with child	16.5%	3.57	0.882	-0.054	-0.174
9) Ask about day in school	58.8%	4.45	0.739	-1.110	0.590
11) Help child with homework	27.9%	3.74	0.999	-0.317	-0.448
14) You ask what plans are for day	16.4%	3.42	1.006	-0.238	0.024
15) Drive/take to special activity	28.2%	3.61	1.191	-0.626	-0.242
20) You talk to child about friends	29.4%	3.93	0.864	-0.389	-0.146
23) Child helps plan family activities	9.7%	3.29	0.902	-0.147	0.399
26) Attend meetings at school	26.9%	3.46	1.290	-0.500	-0.688
Scale: Positive parenting					
2) Let child know when does good job	49.6%	4.32	0.767	-0.826	-0.077
5) Reward or give extra for obeying	18.6%	3.54	0.958	-0.184	-0.063
13) Compliment when child does things well	52.5%	4.38	0.746	-0.979	0.511
16) Praise if behaves well	40.0%	4.15	0.831	-0.621	-0.196
18) Hug or kiss child when he/she does well	50.0%	4.24	0.898	-1.045	0.627
27) Tell child you like when he/she helps out	43.2%	4.15	0.895	-0.857	0.451
Scale: Poor monitoring/supervision					
6) Child fails to leave note when going	43.2%	2.04	1.123	0.825	-0.143
10) Child stays out past time to be home	58.2%	1.70	0.952	1.196	0.609
17) Goes out with friends you do not know	61.8%	1.58	0.847	1.326	1.094
19) Goes out without set time to be home	70.7%	1.52	0.996	2.154	4.097
21) Out after dark without adult	67.6%	1.52	0.863	1.698	2.512
24) Forget where child is and what he/she is doing	65.4%	1.49	0.748	1.345	0.915
28) Do not check that child comes home when supposed to	66.4%	1.75	1.281	1.608	1.253
29) Don't tell child where you are going	51.4%	1.96	1.203	1.071	0.203
30) Child home more than hour past time expected	75.3%	1.39	0.768	2.133	4.412

Item	Positive endpoint	Mean	SD	Skew	Kurtosis
32) Child at home without adult supervision	48.0%	1.87	0.960	0.697	-0.513
Scale: Inconsistent discipline					
3) Threaten to punish then do not	4.0%	3.29	0.952	-0.044	0.057
8) Child talks you out of being punished	16.3%	2.74	1.070	0.034	-0.354
12) Obeying more trouble than worth	40.9%	2.23	1.215	0.501	-0.818
22) Let out of punishment early	8.8%	3.09	1.012	-0.158	0.079
25) Not punished when does wrong	18.9%	2.49	0.983	0.172	-0.152
31) Punishment depends on your mood	20.0%	2.59	1.057	0.130	-0.234

Note. Item text has been edited for length. The positive endpoint is *Always* for Involvement and Positive Parenting scale items and *Never* for Poor Monitoring/Supervision and Inconsistent Discipline items. SE for skewness is approximately 0.094; SE for kurtosis is approximately 0.188.

Table 4

Standardized Factor Loadings and Standard Errors for Final Exploratory Structural Equation Model Three
Factor Solution

Item	Involved Parenting	Poor Monitoring	Inconsistent Discipline
1	0.684±0.025***	0.001±0.029	-0.127±0.045**
4	0.482±0.034***	-0.085±0.046	0.028±0.045
7	0.602±0.029***	-0.091±0.041*	-0.114±0.045*
9	0.715±0.029***	-0.005±0.041	0.023±0.040
11	0.483±0.034***	-0.097±0.046*	0.070±0.046
14	0.568±0.034***	0.182±0.042***	-0.155±0.043***
15	0.357±0.040***	0.039±0.050	-0.131±0.047**
20	0.645±0.028***	-0.015±0.039	0.008±0.036
23	0.544±0.034***	0.029±0.047	-0.143±0.044***
26	0.347±0.041***	0.053±0.050	-0.046±0.046
2	0.715±0.027***	-0.033±0.040	-0.027±0.038
5	0.476±0.034***	-0.006±0.037	0.190±0.042***
13	0.771±0.025***	0.025±0.038	-0.014±0.028
16	0.750±0.026***	-0.019±0.037	0.030±0.032
18	0.703±0.029***	-0.029±0.045	0.023±0.038
27	0.688±0.031***	0.081±0.051	0.058±0.045
6	0.041±0.045	0.632±0.039***	0.032±0.047
10	0.017±0.032	0.775±0.034***	0.052±0.053
17	0.015±0.034	0.730±0.036***	0.040±0.056
19	-0.044±0.049	0.521±0.045***	0.030±0.052
21	-0.061±0.053	0.814±0.037***	-0.042±0.061
24	-0.170±0.048***	0.419±0.049***	0.197±0.055***
28	-0.080±0.052	0.441±0.047***	0.037±0.053
29	-0.192±0.045***	0.119±0.052*	0.069±0.048
30	0.063±0.059	0.730±0.036***	-0.020±0.027
32	-0.141±0.051**	0.370±0.052***	-0.103±0.054
3	0.003±0.012	-0.144±0.062*	0.765±0.034***
8	0.143±0.042***	0.055±0.054	0.567±0.039***
12	-0.207±0.043***	0.133±0.052**	0.449±0.038***
22	0.011±0.035	-0.017±0.052	0.756±0.034***
25	-0.072±0.042	0.115±0.052*	0.384±0.042***
31	-0.181±0.043***	0.003±0.030	0.423±0.036***

Note. Factor variances fixed to 1.00; factor loadings, therefore, are standardized.

Residual covariances freed, all significant at $p < .001$ and listed in order with modification index and residual correlation values: (29,28): 130.70, .42; (26,15): 66.40, .27; (13,2): 38.61, .21; (2,1): 33.40, .17; (16,13): 27.57, .17; (15,4): 30.16, .18; (18,16): 19.44, .14; (16,2): 22.52, .12; (29,15): 22.01, -.17; (26,4): 20.65, .15; (29,24): 18.58, .17.

*
 $p < .05$.

**
 $p < .01$.

 $p < .001$.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5
Standardized Regression Coefficients for Differential Item Functioning Analysis for Parent and Child Demographic Covariates

Factor	Item	Education ^a				Marital status ^b	
		U.S. resident	High school graduate	Any college	No current partner ^c	Married/Common-law	
Involved Parenting	4	0.152 ***					
Involved Parenting	5	0.144 ***					
Involved Parenting	7	0.200 ***					
Involved Parenting	15	-0.238 ***	0.049	0.312 ***	0.281 ***	0.359 ***	
Involved Parenting	26		0.063	0.167 ***	0.149 **	0.179 ***	
Poor Monitoring	29		-0.097 *	-0.159 ***	-0.199 ***	-0.087	
Poor Monitoring	32	-0.177 ***	0.030	0.284 ***			
Inconsistent Discipline	12	-0.237 ***					

Factor	Item	White	Cash/Non-cash benefits	Child age	Social desirability	Defensiveness
Involved Parenting	15	0.297 ***	-0.343 ***			
Involved Parenting	23				0.139 ***	
Poor Monitoring	29	-0.171 ***				
Poor Monitoring	32	0.203 ***	-0.201 ***	0.335 ***		-0.137 **
Inconsistent Discipline	12	0.216 ***				

Note. A separate model was computed for each covariate.

^aReference category is high school non-graduates.

^bReference category is single, never married respondents.

^cNo current partner includes divorced, separated, and widowed.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Regressions of Factors on Demographic Covariates and Validity Scales with Differential Item Functioning Included

Table 6

IV	Involved Parenting <i>b</i> ± <i>SE</i> (<i>b</i> *)	Poor Monitoring <i>b</i> ± <i>SE</i> (<i>b</i> *)	Inconsistent Discipline <i>b</i> ± <i>SE</i> (<i>b</i> *)
U.S. resident	0.067±0.087 (0.034)	-0.198±0.092 (-0.099*)	0.228±0.096 (0.113*)
Education ^d			
High school graduate	0.065±0.122 (0.026)	0.182±0.135 (0.072)	-0.068±0.130 (-0.027)
Any college	0.156±0.099 (0.077)	-0.158±0.104 (-0.078)	-0.171±0.106 (-0.085)
Marital status ^b			
No current partner ^c	-0.016±0.105 (-0.008)	0.232±0.114 (0.114*)	-0.197±0.112 (-0.097)
Married/Common law	-0.225±0.110 (-0.107*)	0.294±0.120 (0.140*)	-0.158±0.118 (-0.076)
White	-0.174±0.085 (-0.087*)	0.309±0.091 (0.152***)	-0.033±0.093 (-0.016)
Cash/Non-cash benefits	0.138±0.087 (0.068)	-0.164±0.094 (-0.081)	0.147±0.093 (0.073)
Female child ^d	0.012±0.084 (0.006)	-0.285±0.091 (-0.141**)	0.128±0.090 (0.064)
Child age	-0.071±0.033 (-0.091*)	0.249±0.036 (0.305***)	-0.049±0.039 (-0.063)
Social desirability	0.677±0.113 (0.270***)	-0.588±0.121 (-0.237***)	-0.862±0.112 (-0.337***)
Defensiveness	0.777±0.126 (0.264***)	-0.592±0.131 (-0.204***)	-1.086±0.133 (-0.358***)

Note. A separate model was computed for each covariate.

^aReference category is high school non-graduates.

^bReference category is single, never married respondents.

^cNo current partner includes divorced, separated, and widowed respondents.

^dNo DIF was observed for female child.

* *p* < .05.

** *p* < .01.

*** *p* < .001.