Scientific Research Publishing

Alabama Parenting Questionnaire—Short Form (APQ-9): Evidencing Construct Validity with Factor Analysis, CFA MTMM and Measurement Invariance in a Greek Sample

Theodoros A. Kyriazos, Anastassios Stalikas

Department of Psychology, Panteion University, Athens, Greece
Email: th.kyriazos@gmail.com

Abstract

This study focused on the factor structure, measurement invariance, reliability, and validity of the Greek version of APQ-9 in a sample of 621 parents of children aged 7 - 13 years. The factor structure was examined first with EFA in the 30% subsample and CFA in the rest 70%. Power analysis indicated adequate CFA sample power at 80% probability of rejecting a false null hypothesis. The original structure of APQ-9 was verified. Full measurement invariance was also examined across child gender to a strict level. Convergent and discriminant validity of APQ-9 parenting practices were evaluated by the CFA MTMM framework with a model of three traits and three methods. Convergent and discriminant validity was also evaluated further with correlation analysis. A consistent pattern of correlations emerged by examining five parenting measures with 13 dimensions of parenting. APQ-9 has also adequate internal consistency and factor-based reliability and validity (α, ω, and AVE).

Keywords

Alabama Parenting Questionnaire, APQ-9, APQ, Parenting Scales, Greece, Normative Data, CFA, EFA, Measurement Invariance, Psychometrics, CFA MTMM, Sample Splitting

1. Introduction

Social and developmental psychology postulates a relationship between both the quality and consistency of parenting practices and psychological adjustment of offspring (Baumrind, 1967; Dadds, Maujean, & Fraser, 2003; Pickering & Sand-

ers, 2016). Parenting practices are specific patterns of actions during parent-child interactions in a given situation (Darling & Steinberg, 1993). Effective parenting practices contribute to psychological and behavioral developmental "outcomes" valuable in western societies (Belsky, 2015; Rasmussen, 2009).

Therefore, reliable and valid measures of parenting effectiveness are important both for clinical and non-clinical research settings (Święcicka et al., 2019). However, in the past, with few exceptions, the most popular measures of parenting examined a narrow range of risk factors related to child misconduct (Dadds et al., 2003). Reviews of parenting measures (Locke & Prinz, 2002) argue that most measures focus on ineffective discipline and parental neglect (Elgar, Waschbusch, Dadds, & Sigvaldason, 2007), or presented a rather questionable psychometric profile (Holden & Edwards, 1989; Locke & Prinz, 2002) as commented by Badahdah and Le (2015). To overcome this problem, the Alabama Parenting Questionnaire was developed (APQ; Frick, 1991; Shelton, Frick, & Wooton, 1996; Frick, Christian, & Wooton, 1999). The questionnaire is among the most frequently used self-report measures of parenting research. Specifically, Google Scholar resulted in more than 430 citations (July 2013; Maguin, Nochajski, De Wit, & Safyer, 2016).

1.1. The Alabama Parenting Questionnaire (APQ-42)

APQ is a multi-method, multi-informant assessment scheme with parallel forms, administered to both children and parents (global report) available also as a phone interview schedule (Essau et al, 2006; Adams, 2015). Parenting behaviors tap five theoretical constructs: Parental Involvement, Positive Parenting, Poor Monitoring/Supervision, Inconsistent Discipline, and Corporal Punishment (Frick et al., 1999). However, previous work suggested a variety of structures with either 3, 4 or 5 factors (Adams, 2015; Maguin, et al., 2016), using mostly EFA (Essau et al. 2006; Badahdah & Le, 2015), CFA (Święcicka et al., 2019) or ESEM (Maguin et al., 2016). More specifically, the APQ Child Global Report has a five-factor structure (Essau et al., 2006), whereas for the Parent Global Report a two, three or four-factor structure emerged (Hawes & Dadds, 2006; Hinshaw et al., 2000; Randolph & Radey, 2011; Molinuevo et al., 2011; Zlomke et al., 2014; Esposito et al., 2016; Maguin et al., 2016). Additionally, the APQ structure was also tested in single-parent family structures (Adams, 2015). However, direct comparisons of the results are challenging due to wide variations in the items used in each study and in child ages of the samples (see also Maguin et al., 2016). APQ has been translated into at least 11 languages (Seabridge, 2012), including German (Essau et al., 2006), Spanish (Molinuevo et al., 2011), Italian (Esposito et al., 2016) Chinese, Arabic (Badahdah & Le, 2015), Ukrainian (Burlaka et al., 2017) and Polish (Święcicka et al., 2019). The APQ-preschool version has been also tested in a sample of hyperactive-inattentive preschool children and controls and three factors emerged (Clerkin et al. 2007; de la Osa et al., 2014). Maguin et al. (2016) examined APQ parenting constructs specific to a special parent population with alcohol-related problems. Internal consistency for the APQ was

reported (Frick et al., 1999; Shelton et al., 1996) to range from α = 0.67 - 0.82, except Corporal Punishment (α = 0.37 - 0.46).

1.2. The Alabama Parenting Questionnaire, Short (APQ-9)

However, the need for faster assessment (Gross, Fleming, Mason, & Haggerty, 2015) leads to a 9-item version of the APQ-42 (Elgar et al., 2007). The factor structure of the APQ-42 was examined in a community sample of 1402 parents from Australia (90% mothers). PCA identified 5 factors, however Parallel Analysis (Horn, 1965) and Minimum Average Partial Correlations test (Velicer, 1976) failed to support 2 factors (Parental Involvement and Corporal Punishment), thus a shorter scale (APQ-9) emerged by retaining three factors (Positive Parenting, Inconsistent Discipline, and Poor Supervision) having three items each with the highest loading (Elgar et al., 2007). Factor loadings were 0.77, 0.76, and 0.79 for the Positive Parenting factor, 0.74, 0.63 and 0.74 for the Inconsistent discipline factor and 0.62, 0.75 and 0.65 for Poor supervision. The three factors (explaining 26.31% of the total variance) were highly correlated with their corresponding APQ-42 scale, $r = 0.89$ (Positive Parenting), $r = 0.90$ (Inconsistent Discipline) and $r = 0.76$ (Poor Supervision (ps < 0.01). The item reduction from 42 to 9 was 78.57% (Elgar et al., 2007). The test developers estimated that APQ-9 could be completed in one-fifth of the time in comparison to APQ-42 (<1 minute).

Subsequently, criterion validity and psychometric properties of this shortened version were examined in an independent sample of parents from Canada (1296 mothers and 745 fathers). In this study, the developers of APQ-9 evaluated the validity in differentiating parents of children with behavior disorders and parents of children without behavior disorders. The Conners Parent Rating Scale-Revised (CPRS-R; Conners, Sitarenios, Parker, & Epstein, 1998) was used to evaluate criterion validity. CPRS-R is an 80-item measure of behavioral problems in children of 3 to 17 years. The 3-factor structure emerging in the first study was confirmed with Confirmatory Factor Analysis separately for mothers and fathers with good model fit for mothers, (CFI) = 0.99, NFI = 0.98 and fathers CFI = 0.99, NFI = 0.98. Factor Loadings ranged from 0.52. - 0.82 for mothers and 0.46 - 0.90 for fathers. Factor intercorrelations ranged from −0.24 to 0.30 for mothers and −0.21 to 0.29 for fathers (Elgar et al., 2007). In a later study, the validity of the short-scale was further supported by correlations between parenting practices and child symptoms to a sample of 133 parents (90.98% mothers) of 5- to 18-year-old children (Elgar et al., 2007).

Internal consistency reliability of the APQ-9 factors ranged from 0.59 - 0.79 for mothers and 0.63 - 0.84 for fathers. The internal consistency of the APQ in the third sample was moderate, ranging from α = 0.57 (Positive Parenting) to α = 0.62 (Inconsistent Discipline). Reliability per age varied for children aged 4 to 9 years, mean α = 0.44; for children aged 5 to 12 years, α = 0.59 to 0.84 and for children aged 5 to 18 years, α = 0.57 to 0.61 (Elgar et al., 2007 as summarized by

Gross et al., 2015). Later, Gross et al. (2015) examined the longitudinal invariance of the APQ-9 for parents and youngsters, and the multigroup invariance between parents and adolescents during their transition from middle school to high school.

1.3. The Present Study

The purpose of this study is to examine the factor structure of APQ-9 using EFA and CFA in a Greek sample of parents of the general population with children from 7 - 13 years. To this end, the study had also the following goals: 1) to evaluate measurement invariance across child gender; 2) to build evidence of convergent and discriminant validity of APQ-9 based on the CFA Multitrait-Multimethod method (CFA MTMM); 3) to reinforce convergent and discriminant validity with correlation analysis; 4) to evaluate internal consistency reliability (with α), model-based reliability (with ω), model-based convergent validity (with AVE) and finally, 5) to calculate normative data for the mean factor scores.

2. Method

2.1. Participants

The sample comprised 621 Greek parents (75% females) with at least one child from 7 to 13 years (M = 10.23 years, SD = 2.11, 54% females). The parents (72% biological mothers, biological 24% fathers, 4% other) had one child (32%), two (48%), three (15%) or more children (5%). More than half of the parents (54%) were from 41 - 50 years old, 28% from 31 - 40 years, 10% from 51 - 60, 7% from 21 - 30 and 1% were over 60 years. Less than half of the participants (39%) had a B.A. or higher (20%), or they had finished high-school (36%) or lower (5%). Most participants (38%) had an annual income between 10,001€ and 20,000€ or lower (21%) while 25% had an income 20,001€ - 30,000€ or higher (16%).

2.2. Measures

Alabama Parenting Questionnaire—Short Form (APQ-9, Elgar et al., 2007)

This nine-item short form of the original APQ-42 (Frick, 1991; Shelton et al., 1996; Frick et al., 1999) is designed to assess parenting practices related to disruptive behaviors (Shelton et al., 1996). It was shortened for faster assessment (Gross et al., 2015). APQ-9 items (e.g. You threaten to punish your child and then do not actually punish him/her) are rated on a 5 point Likert Scale (1 = *never*; 2 = *almost never*; 3 = *sometimes*; 4 = *often*; 5 = *always*). Higher scores indicate higher ratings of the measured parenting practice (i.e. Positive Parenting, Inconsistent Discipline, Poor Supervision).

APQ-9 Translation procedure. APQ-9 was translated in Greek using the translation-back-translation method (Brislin, 1970). First, it was translated in Greek by the first author. Back-translation to English followed by a bilingual psychologist, not familiar with the English version. All items of the original English and the back-translated version went through an iterative process of transla-

tion/ back-translation (3 times) to eliminate differences or ambiguities before the final version.

Kansas Parental Satisfaction Scale (KPSS, James, Schumm, Kennedy, Grigsby, Shectman, & Nichols, 1985)

KPSS is a 3-item scale measuring parental satisfaction with the following: 1) children, 2) parenting role, and 3) parent-child relationship. Items are rated on a 7-point Likert scale (1 = extremely dissatisfied, 7 = extremely satisfied) and aggregated to a total score ranging from 3 (minimum satisfaction) to 21 (maximum satisfaction). An EFA was carried out in the current ample. Kaiser-Meyer-Olkin measure of sampling adequacy (Kaiser, 1970, 1974) was 0.71, and Bartlett's test of sphericity (Bartlett, 1954) was significant (χ^2(3) = 687.06, p < 0.001). A single parent satisfaction factor emerged (PAF extraction, Obilin rotation) explaining a total variance of 61.28%. Factor loadings for items 1 - 3 were 0.80, 0.69 and 0.85 and communalities 0.64, 0.48, 0.72 (Kyriazos & Stalikas, 2019e). The internal consistency reliability of the factor was α = 0.82. The KPSS has been reported having internal consistency reliability from 0.78 to 0.95 (Nitsch et al., 2015).

Parenting Behaviours and Dimensions Questionnaire (PBDQ; Reid, Roberts, Roberts, & Piek, 2015)

PBDC is a scale of parental behaviors containing 33 items on six factors (Emotional Warmth, Punitive Discipline, Autonomy Support, Permissive Discipline, Anxious Intrusiveness, Democratic Discipline). All items (e.g. I try to meet my child's desires immediately) rate the frequency of behaviors on a 6-point Likert scale, from 1 ("*never*") to 6 ("*always*"). The score is calculated based on factor means. The fit of this 6-factor model to this sample was adequate, χ^2(465) = 826.86, χ^2/df = 1.78, RMSEA = 0.042, CFI = 0.922, TLI = 0.912, SRMR = 0.071 (Kyriazos & Stalikas, 2019a). Internal consistency reliability per factor in this study was α = 0.85 (Emotional Warmth), α = 0.82 (Punitive Discipline), α = 0.77 (Anxious Intrusiveness), α = 0.79 (Autonomy Support), α = 0.69 (Permissive Discipline), α = 0.76 (Democratic Discipline). The PBDQ developers reported an alpha coefficient ranging from 0.66 to 0.83 (Reid et al., 2015).

Parent Behavior Inventory (PBI; Lovejoy, Weis, O'Hare, & Rubin, 1999)

PBI is a 20-item measure of parenting practices. Items (e.g. I threaten my child) are rated on a 5-point Likert scale ranging from 1 ("*not at all true*" or "*I do not do this*") to 5 ("*very true*" or "*I often do this*"). Higher scores indicate a higher frequency of the rated practice. Items are divided in two factors, the hostile/coercive factor and the and the supportive/engaged factor. This factor structure was tested in the current sample and showed an adequate fit, χ^2(159) = 322.77, χ^2/df = 2.03, RMSEA = 0.049, CFI = 0.925, TLI = 0.911, SRMR = 0.069 (Kyriazos & Stalikas, 2019b). In this study, internal consistency reliability for the supportive/engaged factor was α = 0.86, and for the hostile/coercive factor α= 0.81. Lovejoy et al., (1999) reported an alpha coefficient of 0.83 and 0.81 for the supportive/engaged parenting and hostile/coercive parenting factor respectively.

Parent Concerns Questionnaire (PCQ; Sheppard, 2010)

PCQ is a 37-item measure of child development or parental problems (Sheppard, 2010). PCQ has three domains (parenting capacity, child development, family/environmental factors). Each item (e.g. I/we are rather too critical of my children) is rated on a 3-point scale (0 = *not present*, 1 = *present*, and 2 = *severe*), producing an aggregated score. Problems perceived by the respondent as "severe" may suggest that professional intervention is required. In the current study this 3-dimensional theoretical model was verified with CFA, $\chi^2(30) = 57.76$, $\chi^2/df = 1.93$, RMSEA = 0.046, CFI = 0.965, TLI = 0.947, SRMR = 0.041 (Kyriazos & Stalikas, 2019c). Factor 1 (child development problems) contained items 24, 25, 29, Factor 2 (Parenting Capacity problems) items 34, 35, 36, and Factor 3 (family/environmental problems) contained items 4, 10, 11, 12 (Kyriazos & Stalikas, 2019c). The alphas per factor of this 10-item structure were 0.76, 0.71 and 0.77 for factors 1 - 3 respectively. Sheppard (2010) reported alpha coefficients of 0.89, 0.79 and 0.73 for the Child Development problems, Parenting Capacity problems and Family/Environmental problems respectively.

Parental Stress Scale (PSS; Berry & Jones, 1995)

PSS is a self-report questionnaire of perceived stress of the parental experience. All 20 items (e.g. The major source of stress in my life is my child) are rated on a 5-point Likert scale (from 1 = "*strongly disagree*" to 5 = "*strongly agree*"). Higher ratings suggest higher parental stress. Items can be arranged in two major domains (positive and stressful parenting themes). Berry and Jones (1995: p. 470) found a 4-factor structure to "support the dichotomy of the parenting experience and the theoretical bases of the Parental Stress Scale". This theoretical dichotomy of the PSS structure was confirmed with CFA, $\chi^2(72) = 148.86$, $\chi^2/df = 2.07$, RMSEA = 0.050, CFI = 0.951, TLI = 0.938, SRMR = 0.062 (Kyriazos & Stalikas, 2019d). Factor 1 (Positive Parenting Themes) comprised items 1, 5, 6, 7, 8, 17, 18 and Factor 2 (Stressful Parenting Themes) comprised items 3, 4, 10, 11, 12, 15, 16. The internal consistency reliability for these two factors was α = 0.87 for positive parenting themes (reversed scored) and α = 0.76 for stressful parenting themes. Berry & Jones (1995) reported a total alpha coefficient of 0.83.

2.3. Procedure

Data were collected with the assistance of psychology students. Specifically, about 100 students forwarded a link of the study to at least 5 parents in their social environment (*M* = 6.21), inviting them to participate in the study. During the data collection, all parents the students recruited, first read a digital description of the study, accepting an inform consent. Then they specified a personal code to ensure anonymity. Students received extra credit for carrying out the recruitment process.

2.4. Research Design

The sample was split in two (about 1/3 and 2/3, Guadagnoli & Velicer, 1988).

The EFA subsample was 30% and the CFA subsample was 70%. A CFA followed the EFA. After CFA, additional analyses were performed in the optimal CFA model: 1) full measurement invariance to the strict level (highest possible, Wang & Wang, 2012); 2) Internal consistency reliability using Cronbach's alpha coefficient (1951) and model-based reliability (Mair, 2018; Sha & Ackerman, 2018) using Bollen's Omega (Bollen, 1980; see also Raykov, 2001) Bentler's Omega, (Bentler, 1972, 2009), and McDonald's Omega (1999, 1970, ω_b) and 3) model-based convergent validity with Average Variance Extracted (AVE; Fornell & Larcker, 1981). To test convergent validity, discriminant validity related to facets of APQ perceived parenting practices a comparison of nested CFA models was carried out within the CFA Multitrait-Multimethod framework (CFA MTMM; Widaman, 1985; an original non-CFA method by Campbell & Fiske, 1959). Convergent and discriminant validity were examined further by correlation analysis using five parenting measures with 13 different scales. Finally, descriptive statistics and normative data were calculated based on factor means for easier comparisons of the scales to APQ scales of different length.

Data were collected electronically on Google Forms® and were analyzed with R software (R Development Core Team, 2019) with the following packages: "haven" V 2.1.1 (Wickham, 2019a), "psych" V1.8.12 (Revelle, 2019), "lavaan" V0.6-4 (see Rosseel, 2012), "MVN" 5.7 (Korkmaz, 2019), "caret" v6.0-84 (Kuhn, 2019), "knitr" V1.23 (Xie, 2019), "dplyr" v0.7.8 (Wickham, 2019a), "tidyr" v0.8.3 (Wickham, 2019b), semPlot v1.1.1 (Epskamp, 2019), "semTools" v0.5-1 (Jorgensen, 2019).

3. Results

Data contained no missing values because all the fields of the digital test-battery were set as "required" to eliminate non-response. Twenty-six out of 621 cases were identified as multivariate outliers, with scores exceeding the critical value χ^2 [9] = 27.88, $p < 0.001$ for Mahalanobis distance (Mahalanobis, 1936; Tabachnick & Fidell, 2013). However, outliers did not alter results so they were included in the dataset. The final sample was $N = 621$ cases. The sample was randomly split in two subsamples ($n_{EFA} = 187$ and $n_{CFA} = 434$). The cases to measured variables ratios for n_{EFA} and n_{CFA} (Costello & Osborne, 2005; Ullman, 2013) were 22.78 and 48.22 respectively. The cases to estimated parameters ratio (see Schumacker & Lomax, 2016) for the hypothesized CFA model (Elgar et al., 2007) was 9.64. Power analysis based on population RMSEA (MacCallum, Browne, & Sugawara, 1996) recommended a CFA sample size ≥ 375 cases (0 = 0.05, α = 0.08, df = 24, 1 − β = 0.80).

3.1. Univariate and Multivariate Normality

The assumption of univariate normality was examined in the whole data set ($N = 621$) with Kolmogorov-Smirnov, Shapiro-Wilk, Shapiro-Francia, and Anderson-Darlingall tests and they were statistically significant ($p < 0.001$) for all meas-

ured variables (Table 1). Multivariate normality was examined with Mardia's multivariate kurtosis test (Mardia, 1970), Mardia's multivariate skewness test (Mardia, 1970), Henze-Zirkler's consistent test (Henze & Zirkler, 1990), Doornik-Hansen omnibus test (Doornik & Hansen, 2008), E-statistic and Roston test. The multivariate normality tests were significant, $p < 0.001$ for all samples (Total, EFA and CFA) as presented in Table 1.

3.2. Exploratory Factor Analysis (n_{EFA} = 187)

Initially, the factorability of the correlation matrix was evaluated (Tabachnick & Fidell, 2013). All APQ items correlated ≥0.30 with at least a second item. Kaiser-Meyer-Olkin measure of sampling adequacy (Kaiser, 1970, 1974) was 0.69, and Bartlett's test of sphericity (Bartlett, 1954) was significant ($\chi^2(36) = 454.42$, $p < 0.01$). The anti-image correlation matrix diagonals were >0.50. Given the above factorability indications, EFA was carried out with all nine items.

Factors were extracted with Principal Axis Factoring and oblique rotation (Oblimin). The number of factors to retain was determined with the following methods: the scree plot (Cattell, 1966), Parallel Analysis (PA; Horn, 1965), Very Simple Structure (VSS; Revelle & Rocklin, 1979), Minimum Average Partial Correlations (MAP; Velicer, 1976), and the goodness of model fit. Model fit was evaluated with the Root Mean Square Error of Approximation (RMSEA;

Table 1. Descriptive Statistics and univariate normality tests for each APQ-9 measured variable along with Multivariate Normality Tests for the total sample and subsamples.

Measured Variables	Descriptive Statistics ($N = 621$)				Univariate Normality Tests ($N = 621$)			
	Mean (M)	St.Dev. (SD)	Skew	Kurtosis	Kolmogorov-Smirnov	Shapiro-Wilk	Shapiro-Francia	Anderson-Darling
APQ 1	4.44	1.01	−2.18	4.26	0.37	0.60	0.60	97.46
APQ 2	2.76	1.04	−0.16	−.68	0.23	0.90	0.90	27.09
APQ 3	1.77	1.18	1.53	1.32	0.35	0.68	0.68	81.05
APQ 4	2.61	1.05	0	−0.66	0.22	0.90	0.90	27.22
APQ 5	1.4	0.86	2.48	6.05	0.45	0.53	0.53	126.61
APQ 6	4.66	0.75	−2.86	9.06	0.44	0.50	0.50	124.66
APQ 7	4.34	1.04	−1.75	2.52	0.35	0.67	0.67	79.75
APQ 8	1.47	0.65	2.35	5.08	0.42	0.55	0.55	117.26
APQ 9	2.83	1.01	−0.08	−0.35	0.24	0.90	0.90	27.72

Multivariate Normality Tests

Sample	Mardia's Skew	Mardia's kurtosis	Henze-Zirkler's	Doornik-Hansen (df)	E-statistic	Royston
Total sample ($N = 621$)	3692.40	51.17	7.37	339.99 (18)	24.56	1187.64
EFA subsample (n_{EFA}= 187)	1360.83	21.57	2.63	69.12 (18)	7.12	617.61
CFA subsample (n_{CFA} = 434)	2764.76	42.80	6.16	149.21 (18)	18.24	1011.80

Note. All univariate and multivariate normality tests were significant at $p < 0.001$ level.

Steiger & Lind, 1980), Root Mean Square of Residuals (RMSR), Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973) and Bayesian information criterion (BIC; Schwartz, 1978). Fit criteria (Hu & Bentler, 1999; Browne & Cudeck, 1993) were RMSEA ≤ 0.06 [90% Confidence Intervals ≤ 0.06], RMSR ≤ 0.0448 (Kelley's criterion; Kelley, 1935; Harman, 1962; Lorezo-Seva & Ferrando, 2013) CFI and TLI ≥ 0.95, and lowest possible BIC

PA (see Figure 1) suggested three factors. VSS complexity 1 achieved a maximum of 0.72 with 2 factors and complexity 2 achieved a maximum of 0.81 with 4 factors. MAP achieved a minimum of 0.05 with 1 factor. BIC reached a minimum with 3 factors and Sample Size adjusted BIC achieved a minimum with 4 factors. Taking into account the joined findings of the above methods, 3 factors were extracted (total explained variance of 65.11%). The Extraction Sums of Squared Loadings suggested that the first factor explained 35.44% of the variance, the second 19.11% of the variance, and the third factor 10.56% of the variance with communalities > 0.30. The fit of this model was adequate, RMSR = 0.03, TLI = 0.923, RMSEA = 0.072 [90% CI 0.021, 0.112] and BIC = −40.09. Regarding item allocation to the extracted factors, items 1, 6 and 7 loaded on the first factor (Positive Parenting) with loadings ranging from 0.513 to 0.862, items 2, 4, and 9 loaded on the second factor (Inconsistent Disciple), with loadings from 0.465 to 0.767. Items 3, 5, 8 loaded on the third factor (Poor Supervision) with loadings ranging from 0.640 to 0.777. Table 2 contains the APQ-9 factor loadings above 0.30 and factor inter-correlations (also presented in Figure 2).

3.3. Confirmatory Factor Analysis (n_{CFA} = 434)

CFA was carried out with the Robust Maximum Likelihood estimator (MLR; see Yuan & Bentler, 2000). Goodness of model fit was evaluated by the RMSEA ≤ 0.06, RMSEA 90% CI ≤ 0.06, SRMR ≤ 0.08, CFI ≥ 0.95, TLI ≥ 0.95 (Hu & Bentler, 1999; Browne & Cudeck, 1993; Brown, 2015), and Chi-square/df ratio < 3
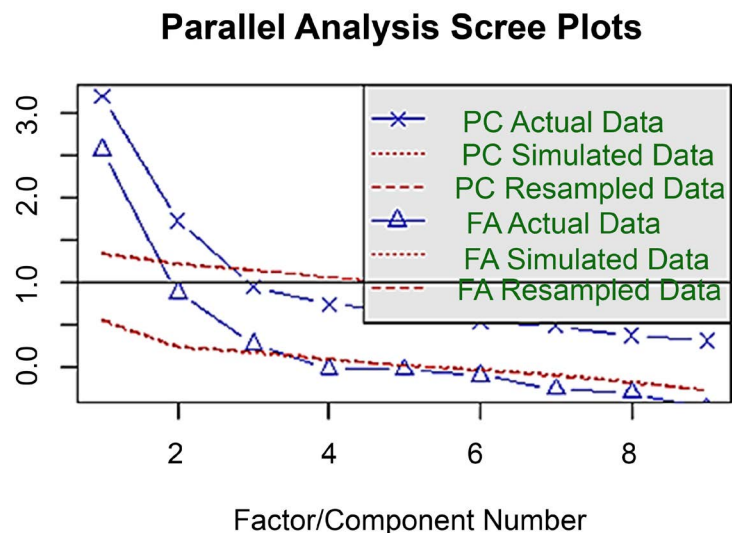

Parallel Analysis Scree Plots

Figure 1. Scree plots of actual and simulated data.

Table 2. EFA factor loadings, communalities and factor Inter-correlations for the APQ-9.

(n_{EFA} = 187) Measured Variables	Factors			Communalities
	Factor 1 Positive Parenting (PP)	Factor 2 Inconsistent Discipline (ID)	Factor 3 Poor Supervision (PS)	
APQ-1	0.704	0	0	0.53
APQ-2	0	0.767	0	0.56
APQ-3	0	0	0.658	0.39
APQ-4	0	0.465		0.31
APQ-5	0	0	0.777	0.66
APQ-6	0.862	0	0	0.73
APQ-7	0.513	0	0	0.29
APQ-8	0		0.640	0.46
APQ-9	0	0.636	0	0.44
Factor Inter-Correlations				
	PP	ID	PS	
PP	-	−0.06	−0.62	
ID	−0.06	-	0.32	
PS	−0.62	0.32	-	

Note. Extraction = PAF, Rotation = Oblimin. Loadings < 0.30 were excluded.

Factor Analysis
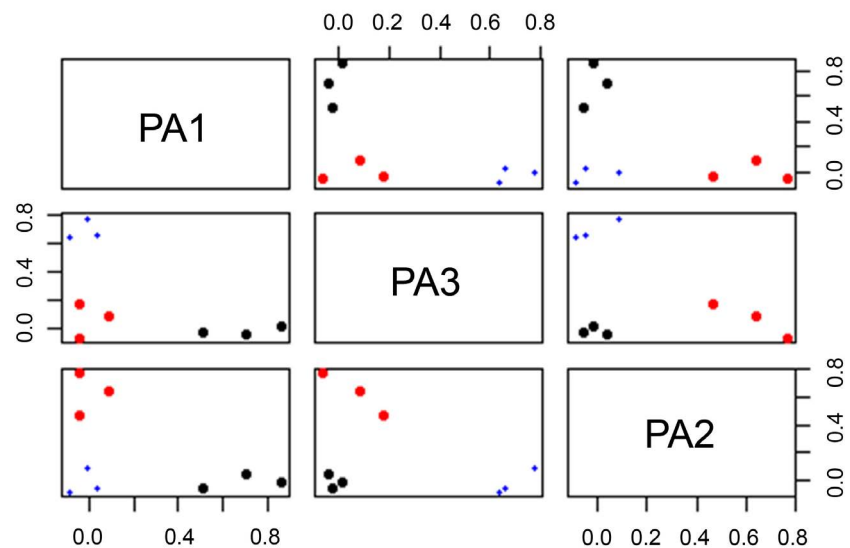
Figure 2. Factor Loadings of each factor.

(Kline, 2016). Models with smaller values of Akaike information criterion (AIC; Akaike, 1987) and BIC are preferable (Mair, 2018).

Three models were tested: (A) a single-factor model with all nine items in a

single factor to test the maximum parsimony hypothesis (Brown, 2015); (B) a first-order, Independent Cluster Model (ICM-CFA; Marsh et al., 2014; Howard et al., 2016) with two correlated factors examined (but not proposed) by Elgar et al., (2007). This model had the original PP factor and a second factor with all the non-positive-parenting items (2, 4, 9, 3, 5, 8); (C) the first order ICM-CFA model with three correlated factors proposed by Elgar et al. (2007). Regarding the model fit, the hypothesis of maximum parsimony was rejected (MODEL A). The two-factor ICM-CFA model also performed poorly (MODEL B). The 3-factor model (MODEL C) had adequate fit, with all fit statistics and factor loadings within acceptable limits. The fit statistics and the standardized loadings of all models are presented in Table 3 and the path of this optimal model in Figure 3. A second-order 3-factor Bifactor model (Harman, 1976; Holzinger & Swineford, 1937) was also tested but it failed to converge. This model had PP, ID and PS items in three specific factors tapping simultaneously in a general factor.

3.4. Measurement Invariance

The configural, weak, strong and strict full measurement invariance were evaluated across the gender of the child, the 621 parents had completed the APQ-9 for. The nested models were compared using the cutoffs of ΔCFI $\leq$ 0.01 (Cheung & Rensvold, 2002; Chen, 2007) and ΔRMSEA $\leq$ 0.015 (Chen, 2007). The 3-factor optimal solution was tested separately for each child-gender (Table 4). These models showed an adequate fit both for girls ($N = 337$) and for boys ($N = 284$). Nested invariance models (1 - 4) also fit the data well (Table 5). The weak to configural model comparison and the strong to weak model comparison yielded ΔCFIs and ΔRMSEAs below the cutoffs of non-invariance. However, in the strict to strong model comparison, only the ΔRMSEA cutoff supported invariance.

3.5. Internal Consistency Reliability, Model-Based Reliability, and Validity

Cronbach's alpha $\geq$ 0.70 is generally acceptable (Hair et al., 2010). Omega values

Table 3. Goodness of fit measures, factor loadings and Inter-correlations for the APQ models specified in the CFA.

$N = 434$ Model	χ^2	df	χ^2/df	CFI	TLI	RMSEA	Lower	Higher	SRMR	BIC	AIC	Factor Loadings Range	Factor Inter-correlations
MODEL A Single factor	480.10	27	17.78	0.381	0.174	0.197	0.180	0.214	0.138	10,760.8	10,687.5	0.027 - 0.753	-
MODEL B 2-factor	159.25	26	6.13	0.672	0.546	0.064	0.109	0.095	0.123	10,596.3	10,518.9	F1 = 0.432 - 00.948 F2 = 0.215 - 00.702	−0.050
MODEL C 3-factor	**44.103**	**24**	**1.84**	**0.951**	**0.926**	**0.044**	**0.024**	**0.063**	**0.043**	**10,453.2**	**10,367.7**	**F1 = 0.457 - 0.896 F2 = 0.598 - 0.681 F3 = 0.379 - 0.716**	**F1➔F2 = 0.022 F1➔F3 = −0.353 F2➔F3 = 0.224**

Note. Estimator = MLR; Bold typeface indicates optimal fit. df = Degrees of freedom; CFI = Comparative fit index; TLI = Tucker-Lewis index; RMSEA = Root mean square error of approximation; CI = Confidence interval; SRMR= Standardized root mean square residual. FI = Factor 1 (items 1, 6, 7), F2 = Factor 2 (items 2, 4, 9), F3= Factor 3 (items 3, 5, 8).

Table 4. Goodness-of-fit measures for the baseline model for testing measurement invariance across child gender for the 3-factor APQ-9 model.

N = 621 (337 girls & 284 boys) Models	Chi-Square Value	Chi-Square df	Chi-square /df	CFI	TLI	RMSEA	RMSEA Lower CI	RMSEA Higher CI	SRMR
MODEL 1 3-factors correlated (GIRLS)	47.72	24	1.98	0.949	0.924	0.054	0.032	0.076	0.048
MODEL 2 3-factors correlated (BOYS)	27.70	24	1.15	0.984	0.976	0.023	0.000	0.061	0.039

Note. Estimator = MLR.

Table 5. Goodness-of-Fit measures for the nested APQ-9 models to validate full measurement invariance across the child gender of the parents.

N = 621 Models	Chi-Square	Df	CFI	RMSEA	Model comparison	ΔCFI	ΔRMSEA
1. Full Configural Invariance	74.19	48	0.962	0.042	-	-	-
2. Full Weak Invariance	78.54	54	0.965	0.038	Model 2 vs 1	0.003	−0.004
3. Full Strong Invariance	85.00	60	0.964	0.037	Model 3 vs 2	−0.001	−0.001
4. Full Strict Invariance	104.32	69	0.949	0.041	Model 4 vs 3	-0.015	0.004

Note. Estimator = MLR.
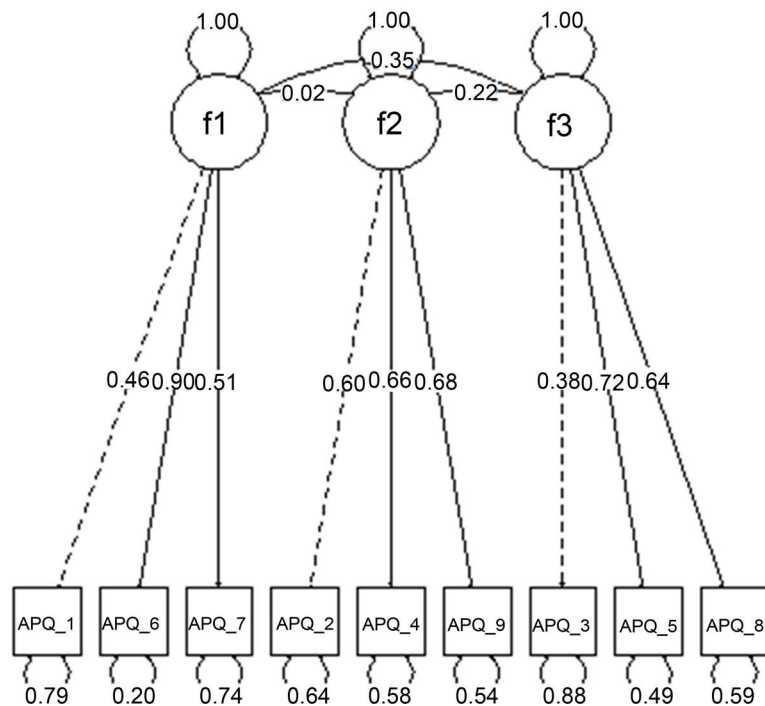
Figure 3. Path diagram of the optimal CFA solution for the APQ-9. Conventionally, cycles are latent factors, rectangles represent manifest variables.

≥ 0.70 are also acceptable (Hair et al., 2010). Average Variance Extracted (AVE; Fornell & Larcker, 1981) ≥ 0.50 are satisfactory (Fornell & Larcker, 1981).

The internal consistency reliability of the APQ-9 PP, ID and PS scales was es-

timated in the total sample. Cronbach's α coefficients ranged from 0.61 - 0.68 (Table 6). On average ω coefficients ranged from 0.64 - 0.65, 0.68 and 0.62 for the PP, ID and PS scales respectively. AVE ranged from 0.35 - 0.41 (Table 6).

3.6. Convergent and Discriminant Validity with CFA Multitrait-Multimethod Model (CFA MTMM)

The hypothesized Correlated Traits/Correlated Methods model (Model 1-CTCM, Table 7) was compared to three alternatives commonly used MTMM Models (Byrne, 2012): No Traits/Correlated Methods (Model 2-NTCM), Perfectly Correlated Traits/Freely Correlated Methods (Model 3-PCTCM) and Freely Correlated Traits/Uncorrelated Methods (Model 4-CTUM). The CFA MTMM model was parameterized with 3 Traits and 3 Methods. The 3 Traits were composed by 1) Positive Parenting containing the APQ Positive Parenting factor (items 1, 6, 7) and PBDQ Emotional Warmth factor (items 1, 2, 3, 4, 5, 6); 2) Inconsistent Discipline that contained the APQ Inconsistent Discipline factor (items 4, 9) and PBI Hostile/Coercive Parenting factor (items 1, 3, 5, 7, 9, 13, 15, 17, 19, 20); 3) Poor Supervision that contained the APQ Poor Supervision factor (items 3, 5, 8) and PBDQ Punitive Discipline factor (items 7, 8, 9, 10, 11). The 3 methods comprised: 1) Alabama Parenting Questionnaire (items 1, 3, 4, 5, 6, 7, 8, 9); 2)

Table 6. Internal consistency reliability and model-based reliability and validity for the three APQ-9 scales in the optimal CFA model.

N = 621	APQ-9 Factors		
	Positive Parenting (PP)	Inconsistent Discipline (ID)	Poor Supervision (PS)
Cronbach's Alpha (α)	0.63	0.68	0.61
Bollen's Omega (ω)	0.64	0.68	0.62
Bentler's Omega (ω)	0.64	0.68	0.62
McDonald's Omega (ω_t)	0.65	0.68	0.62
Average Variance Extracted (AVE)	0.38	0.41	0.35

Note. PP = items 1, 6, 7, ID = items 2, 4, 9, PS = items 3, 5, 8.

Table 7. Goodness-of-Fit measures of CFA MTMM models specified for the APQ-9.

N = 621 CFA MTMM Models	Chi-Square	Df	CFI	RMSEA	SRMR
Model 1 (CTCM) Correlated traits and methods	776.64	342	0.909	0.045	0.049
Model 2 (NTCM) No traits, correlated methods	2258.48	374	0.607	0.090	0.126
Model 3 (PCTCM) Perfectly correlated traits, correlated methods	976.61	345	0.868	0.054	0.048
Model 4 (CTUM) Correlated traits, uncorrelated methods	894.25	345	0.886	0.051	0.075

Note. Estimator = MLR.

Parent Behavior Inventory (items 1, 3, 5, 7, 9, 13, 15, 17, 19, 20) and 3) Parenting Behaviours & Dimensions Questionnaire (items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11). The $\Delta\chi^2$ test (based on MLR) and the ΔCFI criteria were used to compare the fit difference of the nested models (Cheung & Rensvold, 2002; Byrne, 2010, 2012).

The fit of the baseline model (MODEL 1, CTCM) to the data was good. The fit of the rest MTMM models is presented in Table 7. Regarding model comparison, the $\Delta\chi^2$ was highly significant ($p < 0.0001$), and the fit difference (ΔCFI) supported the traits convergent and discriminant validity. The findings of methods discriminant validity were conflicting. While ΔCFI (0.023) was within acceptable limits (Byrne, 2010: p. 291), $\Delta\chi^2$ was statistically significant. All the results of model comparisons are summarized in Table 8. The factor loadings of the CTCM Model are presented in Table 9. The path diagrams of the CFA MTMM model are presented in Figure 4.

3.7. Convergent and Discriminant Validity with Correlation Analysis

The validation measures were arranged in two groups: Positive and Non-Positive Parenting Practices (Table 10). PP was positively correlated (at $p < 0.01$ level) with the scales in the Positive Parenting Practices Group at a magnitude ranging from $r_S(619) = 0.17$, $p < 0.01$ (Kansas Parental Satisfaction Scale) to $r_S(619) = 0.29$, $p < 0.01$ (PBDQ Autonomy Support and PBDQ Democratic Discipline). PP was low to moderately correlated with the scales in the Non-positive Parenting Practices Group from $r_S(619) = 0.11$, $p < 0.01$ (PBDQ Anxious Intrusiveness) to $r_S(619) = -0.12$, $p < 0.01$ (PBDQ Punitive Discipline). ID showed statistically significant, negative correlations with the Positive Parenting Practices Group ranging from $r_S(619) = -0.12$, $p < 0.01$ (PSS Positive Parenting Themes) to $r_S(619) = -0.18$, $p < 0.01$ (PBDQ Autonomy Support) and positive correlations with the scales of Non-Positive Parenting Practices Group varying from $r_S(619) = 0.03$, *ns* (*PCQ* Family/Environmental problems) to $r_S(619) = 0.43$, $p < 0.01$ (PBDQ Punitive Discipline). Similarly, PS showed low to moderate negative correlations with all scales contained in the Positive Parenting Practices Group from $r_S(619) = -0.20$, $p < 0.01$ (KPSS) to $r_S(619) = -0.29$, $p < 0.01$ (PBDQ Emotional Warmth). PS showed positive (with one exception), low to moderate correlations

Table 8. Differential goodness-of-fit statistics for CFA MTMM nested model comparisons.

$N = 621$ CFA MTMM Model	ΔChi-Square	Δdf	Exact p Value	ΔCFI	Model Comparison
Convergent Validity Traits	2896.527	32	**0.0000**	0.302	Model 1 vs 2
Discriminant Validity Traits	5713.429	3	**0.0000**	0.041	Model 1 vs 3
Discriminant Validity Methods	309.516	3	**0.0000**	0.023	Model 1 vs 4

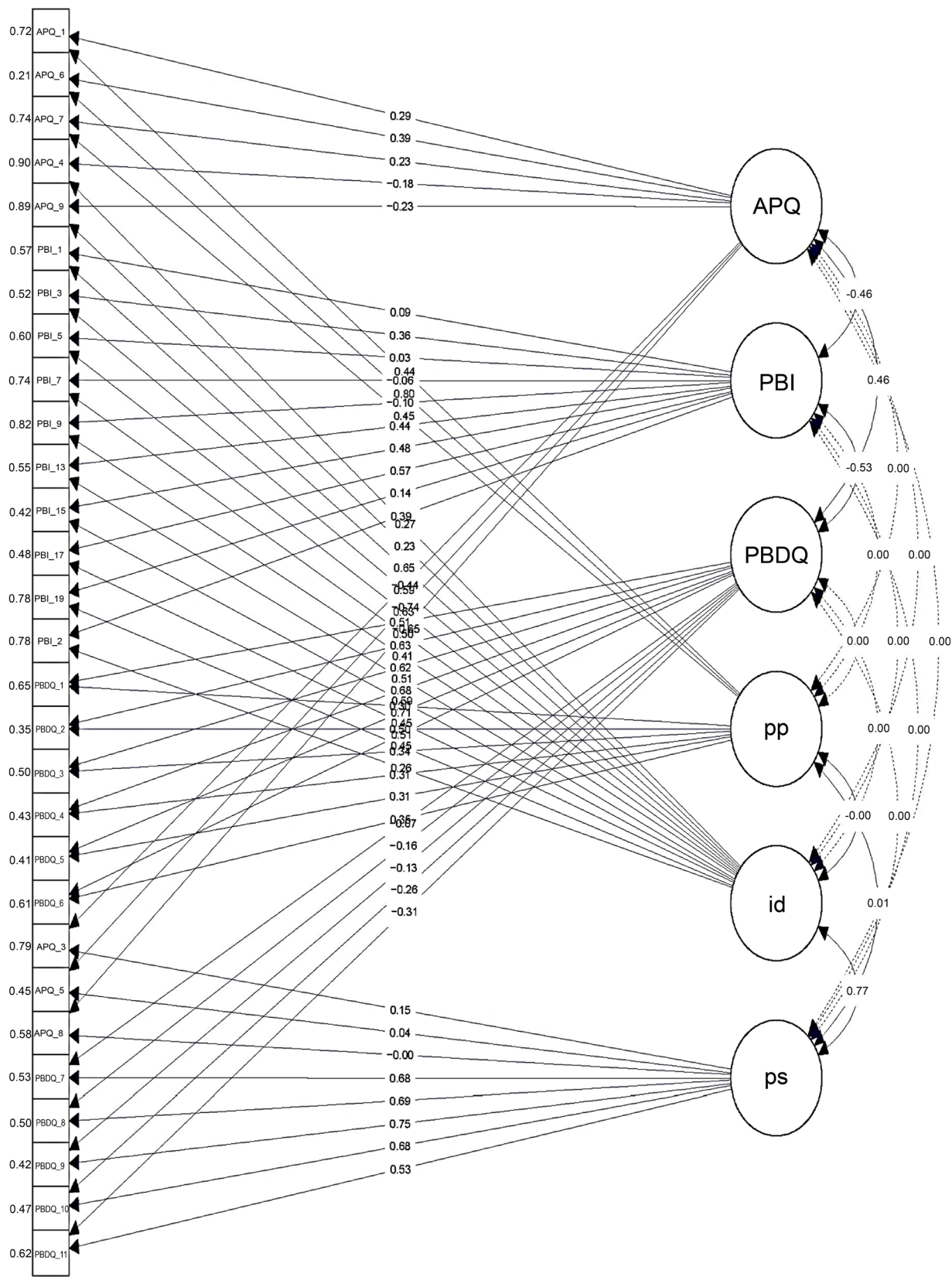*Note.* ΔChi-Square was based on MLR estimator.

Figure 4. CFA MTMM model (Model 1-CTCM): Path diagram of the correlated traits (latent variables in lowercase)/correlated methods (latent variables in uppercase).

Table 9. Factor loadings of the CFA MTMM.

Item	PP	ID	PS	APQ	PBI	PBDQ
APQ 1	0.441	-	-	0.287	-	-
APQ 6	0.798	-	-	0.389	-	-
APQ 7	0.452	-	-	0.229	-	-
PBDQ 1	0.297	-	-	-	-	0.508
PBDQ 2	0.497	-	-	-	-	0.508
PBDQ 3	0.341	-	-	-	-	0.619
PBDQ 4	0.315	-	-	-	-	0.684
PBDQ 5	0.305	-	-	-	-	0.706
PBDQ 6	0.355	-	-	-	-	0.515
APQ 4	-	0.266	-	−0.182	-	-
APQ 9	-	0.231	-	−0.74	-	-
PBI 1	-	0.652	-	-	0.095	-
PBI 3	-	0.589	-	-	0.359	-
PBI 5	-	0.632	-	-	0.029	-
PBI 7	-	0.505	-	-	−0.062	-
PBI 9	-	0.412	-	-	−0.098	-
PBI 13	-	0.509	-	-	0.438	-
PBI 15	-	0.594	-	-	0.48	-
PBI 17	-	0.446	-	-	0.568	-
PBI 19	-	0.447	-	-	0.137	-
PBI 20	-	0.264	-	-	0.389	-
APQ 3	-	-	0.152	−0.436	-	-
APQ 5	-	-	0.041	−0.74	-	-
APQ 8	-	-	−0.004	−0.646	-	-
PBDQ 7	-	-	0.68	-	-	−0.068
PBDQ 8	-	-	0.692	-	-	−0.157
PBDQ 9	-	-	0.75	-	-	−0.128
PBDQ 10	-	-	0.676	-	-	−0.265
PBDQ 11	-	-	0.535	-	-	−0.31

Table 10. Bivariate correlations of APQ-9 with validation scales.

	Spearman rho		
Validation Scales	**APQ-9 Positive Parenting (PP)**	**APQ-9 Inconsistent Discipline (ID)**	**APQ-9 Poor Supervision (PS)**
Positive Parenting Practices Group			
Kansas Parental Satisfaction Scale	0.17**	−0.15**	−0.20**
PBDQ Emotional Warmth	0.38**	−0.15**	−0.29**
PBDQ Autonomy Support	0.29**	−0.18**	−0.15**

Continued

PBDQ Democratic Discipline	0.29**	−0.14**	−0.17**
PBI Supportive/Engaged Parenting	0.26**	−0.17**	−0.26**
PSS Positive Parenting Themes	0.21**	−0.12**	−0.23**
Non-Positive Parenting Practices Group			
PBDQ Anxious Intrusiveness	0.11**	0.13**	−0.08
PBDQ Punitive Discipline	−0.12**	0.43**	0.23**
PBDQ Permissive Discipline	−0.03	0.34**	0.09*
PCQ Child Development Problems	−0.01	0.07	0.16**
PCQ Parenting Capacity Problems	−0.09*	0.12**	0.17**
PCQ Family/Environmental Problems	−0.03	0.03	0.17**
PBI Hostile/Coercive Parenting	−0.10*	0.26**	0.20**
PSS Stressful Parenting Themes	−0.10*	0.15**	0.16**

Note. **Significant at $p < 0.01$ level. *Significant at $p < 0.5$ level.

with the scales of Non-Positive Parenting Practices Group, from $r_S(619) = −0.08$, *ns* (PBDQ Anxious Intrusiveness) to $r_S(619) = 0.23$, $p < 0.01$ (PBDQ Punitive Discipline). All correlations are presented in Table 10.

3.8. Descriptive Statistics and Normative Data

APQ-9 factor scores for PP, ID and PS factors were $M = 4.48$ ($SD = 0.71$), $M = 2.73$ ($SD = 0.81$), and $M = 1.54$ ($SD = 0.76$) respectively. The 10th, 25th, 50th, 75th and 90th percentile of the factor scores were calculated ($N = 621$). For PP, ID, and PS, 50% of the respondents had $M \leq 4.67$, ≤ 2.67 and ≤ 1.33 respectively. For each APQ-9 measured variable the highest means were observed on item 6 ($M = 4.66$, $SD = 0.75$) and 1 ($M = 4.44$, $SD = 1.01$), equivalent to *often—always* Likert points. The lowest mean was found on item 3 ($M = 1.77$, $SD = 1.18$ (or *never—almost never*). All percentile means are presented in Table 11 and the measured variables means were presented in Table 1.

Regarding the correlations of the APQ-9 factors, the correlation of PP with ID was $r_S(619) = 0.01$, *ns*. The correlation of PP with PS was $r_S(619) = −0.23$, $p < 0.01$. Finally, the correlation of ID with PS was $r_S(619) = −0.20$, $p < 0.01$.

4. Discussion

The purpose of this study was to evaluate the factor structure of APQ-9 in a Greek sample of the general population with EFA and CFA. The aim of the study was also: 1) to examine measurement invariance; 2) to evaluate convergent and discriminant validity of APQ-9 based on CFA Multitrait Multimethod Matrix (CFA MTMM); 3) to examine convergent and discriminant validity further with correlation analysis; 4) to estimate internal consistency (with coefficient alpha Cronbach, 1951), model-based reliability (with coefficient omega, McDonald, 1999, 1970), and model-based convergent validity (using Average Variance

Table 11. Percentiles of the APQ-9 factor means.

Total Sample ($N = 621$)	Mean (SD)	Range	Percentile				
			10	25	50	75	90
Positive Parenting	4.48 (0.71)	1 - 5	3.67	4	**4.67**	5.00	5.00
Inconsistent Discipline	2.73 (0.81)	1 - 5	1.67	2.33	**2.67**	3.33	3.67
Poor Supervision	1.54 (0.76)	1 - 5	1.00	1.00	**1.33**	2.00	2.67

Extracted/AVE, Fornell & Larcker, 1981), finally 5) to calculate normative data for the mean factor scores.

The sample was recruited using a variation of the network sampling method (APA, 2014), with the difference that those who recruited volunteers did not participate in the sample themselves. The sample was randomly divided into two subsamples. EFA was carried out in the first subsample and CFA followed in the second one. Sample-splitting (Guadagnoli & Velicer, 1988; MacCallum, Browne, & Sugawara, 1996) is considered a construct validity cross-validation method (Byrne, 2012; Brown, 2015; see also Kyriazos, 2018a, 2018b). Sample to measured variables ratios was higher than the proposed minimums for both the EFA (Costello & Osborne, 2005) and the CFA subsample (Bentler & Chou, 1987; Bollen, 1989). The CFA sample to estimated parameters ratio was also higher than the proposed minimums of adequacy (Kline, 2016). A post hoc estimation of CFA sample power (Wang, Watts, Anderson, & Little, 2013) suggested that sample size was larger than the proposed CFA sample at 80% probability level for rejecting a false null hypothesis (Cohen, 1988, 1992).

Moving to research findings, EFA factorability of the correlation matrix was evaluated with multiple methods and they suggested satisfactory factorability. The three factors were extracted with Principal Axis Factoring method and an oblique rotation because of the APQ-9 factor correlations. The number of factors to retain was three. The fit of this 3-factor model was good using multiple fit indicators (Brown, 2015). Communalities suggested that the shared common variance of the items was adequate. All the factor loadings were good forming three robust factors (Positive Parenting, Inconsistent Discipline, and Poor Supervision) with no cross-loadings. This EFA solution verified the structure originally proposed both by Elgar et al. (2007) subsequently by Gross et al. (2015) in a longitudinal study.

CFA followed in the second subsample with the evaluation of three alternative models. The fit was evaluated adopting the multiple assessment approaches (Bentler & Bonett, 1980), for more conservative results (Brown, 2015). Apart from the commonly accepted goodness of fit statistics, the chi-square/df ratio was calculated, although it received criticism (e.g. Kline, 2016) because its inclusion is a common practice. All chi-square-based criteria used were interpreted in tandem with the rest fit indicators as a result of chi-square over-sensitivity to samples $n > 200$ (Little, 2013; see Kyriazos, 2018b). A CFA Bifactor model (Harman, 1976; Holzinger & Swineford, 1937) was also specified. Generally, testing a Bi-

factor structure is considered good practice (Hammer & Toland, 2016). Unfortunately, the Bifactor model failed to converge and it lacked a theoretical background to attempt troubleshooting the convergence problem with recommended solutions (Byrne, 2012; Heck & Thomas, 2015). We could not test a higher-order model either, because of the inherent under-identification problems for $m \leq 3$ (e.g. Wang & Wang, 2012). After examining the combined evidence of model fit, factor loadings and factor inter-correlations, the 3-factor model with correlated factors was the optimal solution. This finding confirmed both the preceding EFA model and the structures proposed in the literature (Elgar et al., 2007; Gross et al., 2015). The factor loadings and inter-correlations of this optimal 3-factor solution were satisfactory and comparable to those of the APQ-9 model propose by Elgar et al. (2007). Additionally, three factors are consistent for APQ-42 validation studies (Hinshaw et al., 2000; Randolph & Radey, 2011; Zlomke et al., 2014; Molinuevo et al., 2011), except for Robert (2009) and Święcicka et al. (2019) who extracted five factors and Zlomke et al. (2014) who found four factors (see Maguin et al., 2016). However, interpreting these results is complicated by the variation of the allocation of the measured variables to factors (Maguin et al., 2016; Esposito et al., 2016).

APQ-9 measurement invariance across child gender was evaluated in the total sample using the three-factor model as a baseline model. Full invariance was examined to the strict level, i.e. the strictest possible measurement invariance level (Wang & Wang, 2012). The comparison of the nested models showed that configural, Weak and Strong invariance were fully supported and Strict invariance was partially supported. Actually, this level is often hard to establish in practice (Timmons, 2010). Thus, factor structure factor loadings and indicator means can be safely compared between parents that either care for a girl or a boy. However, indicator residuals comparisons between parents of girls and parents of boys must be made cautiously. Generally, the heterogeneity of the existing studies, along with the lack of reported results details blur the assessments of invariance across samples (Maguin et al., 2016) and family types (Adams, 2015).

Convergent and discriminant validity of APQ-9 parenting practices were evaluated with the CFA Multitrait-Multimethod method (Widaman, 1985), using three traits and 3 methods. Findings suggested strong tenability for the traits convergent and discriminant validity, and less strong for methods discriminant validity, as expected based on methods used. Convergent and discriminant validity were also examined with correlations of APQ-9 with five validity measures having 13 dimensions were examined. The validity measures were arranged in two broad categories: 1) Positive parenting practices and 2) Negative parenting practices. A fairly consistent pattern or relationships emerged for all three APQ-9 factors, in agreement with the existing literature (Elgar et al., 2007; Gross et al., 2015 and Dadds et al., 2003 for the original APQ). As expected, APQ-9 Positive Parenting Scale consistently showed almost the opposite pattern of relationships,

in comparison to the pattern of relationships of Inconsistent Discipline and Poor Supervision Scales. Almost all relationships were statistically significant with low to moderate magnitude, abiding by the criteria specified by Cohen (1988, 1992). The strength of associations is discussed in parenting literature (e.g. Seabridge, 2012; Hershkowitz et al., 2017; Burlaka et al., 2017).

Internal consistency reliability and factor-based reliability (Mair, 2018) were measured with Cronbach's alpha (1951) and three omega methods (Bollen, 1980; see also Raykov, 2001; Bentler, 1972, 2009; McDonald, 1999, 1970; Werts, Lim, & Joreskog, 1974). Multiple methods were calculated because Cronbach's alpha may generate inaccurate estimates in multidimensional constructs, although in unidimensional ones it produces similar results to factor-based reliability measures (Sha & Ackerman, 2018). In this study, internal consistency reliability and the factor-based reliability estimates were comparable, corroborating each other. However, AVE stayed below the levels of acceptability, maybe due to inherent dichotomy of the APQ dimensions (positive and non-positive). Their results were also generally comparable to the original results of APQ-9 and APQ-42 (>0.60). Genarally, the parenting measures are notorious for internal consistency in the 0.60 range due to the complexity and broadness of parenting construct (or lower; see Maguin et al., 2016) for the APQ-42 (Shelton et al., 1996; Frick et al., 1999), APQ-15 (Badahdah & Le, 2015) and the APQ-9 (Elgar et al., 2007). For the broad constructs, these findings are not uncommon (Kline, 1999; Boyle, 1991), especially taking into account the sensitivity of alpha to the number of items (Green, Lissitz, & Mulaik, 1977; Nunnally & Bernstein, 1994). Finally, average internal consistency reliability for the APQ-42 scales is $\alpha = 0.68$ (Dadds et al., 2003). The Spearman-Brown formula predicts 3-item subscales with the internal consistency of $\alpha = 0.44.247$ (Smith, McCarthy, & Anderson, 2000; Elgar et al., 2007).

Lastly, given the violation of the normality assumption, percentiles, factor means, and item means were also calculated. The findings were also comparable to the values of the original APQ-9 (Elgar et al., 2007). Future research directions could include the comparison of different models for mothers and fathers, measurement invariance in other demographics like parent age, or gender. Longitudinal measurement invariance could be also tested to replicate Gross et al., (2015) findings. The present solution could be examined in children older than 13 years. Additionally, multi-cultural studies are necessary to assess measurement invariance further. Likewise, assessments of invariance under demographic variation are also needed (Maguin et al., 2016).

Finally, the sample size didn't allow the full implementation of the 3-faced construct validation method (Kyriazos, 2018a; Kyriazos, Stalikas, Prassa, & Yotsidi, 2018). Anyhow, the findings of this study—in line with literature demands for shorter assessment (Scott, Briskman, & Dadds, 2011; Gross et al., 2015)—make the use of APQ-9 more reliable for use in future parenting interventions in Greece and provide normative data for professionals.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Adams, L. M. (2015). *Utilization of the Alabama Parenting Questionnaire across Family Structures: Do the Same Constructs Apply?* LSU Doctoral Dissertations. 152. https://digitalcommons.lsu.edu/gradschool_dissertations/152

Akaike, H. (1987). Factor Analysis and AIC. *Psychometrika, 52,* 317-332. https://doi.org/10.1007/BF02294359

APA (2014). *APA Dictionary of Statistics and Research Methods.* Washington DC: American Psychological Association.

Badahdah, A., & Le, K. T. (2015). Parenting Young Arab Children: Psychometric Properties of an Adapted Arabic Brief Version of the Alabama Parenting Questionnaire. *Child Psychiatry & Human Development, 47,* 486-493. https://doi.org/10.1007/s10578-015-0581-8

Bartlett, M. S. (1954). A Note on the Multiplying Factors for Various χ^2 Approximations. *Journal of the Royal Statistical Society (Series B), 16,* 296-298. https://doi.org/10.1111/j.2517-6161.1954.tb00174.x

Baumrind, D. (1967). Child Care Practices Anteceding Three Patterns of Preschool Behavior. *Genetic Psychology Monographs, 75,* 43-88.

Belsky, J. (2015). Social-Contextual Determinants of Parenting. In M. Boivin, R. De V. Peters, & R. E. Tremblay (Eds.), *Encyclopedia on Early Childhood Development* (pp. 60-64). Montreal: Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development.

Bentler, P. M. (1972). A Lower-Bound Method for the Dimension-Free Measurement of Internal Consistency. *Social Science Research, 1,* 343-357. https://doi.org/10.1016/0049-089X(72)90082-8

Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin, 107,* 238-246. https://doi.org/10.1037//0033-2909.107.2.238

Bentler, P. M. (2009). Alpha, Dimension-Free, and Model-Based Internal Consistency Reliability. *Psychometrika, 74,* 137-143. https://doi.org/10.1007/s11336-008-9100-1

Bentler, P. M., & Bonett, D. G. (1980). Significance Tests and Goodness-of-Fit in the Analysis of Covariance Structures. *Psychological Bulletin, 88,* 588-600. https://doi.org/10.1037/0033-2909.88.3.588

Bentler, P. M., & Chou, C. P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research, 16,* 78-117. https://doi.org/10.1177/0049124187016001004

Berry, J. D., & Jones, W, H. (1995) The Parental Stress Scale: Initial Psychometric Evidence. *Journal of Social and Personal Relationships, 12,* 463-472. https://doi.org/10.1177/0265407595123009

Bollen, K. A. (1980). Issues in the Comparative Measurement of Political Democracy. *American Sociological Review, 45,* 370-390. https://doi.org/10.2307/2095172 https://www.jstor.org/stable/2095172

Bollen, K. A. (1989). *Structural Equations with Latent Variables.* New York: Jon Wiley & Sons. https://doi.org/10.1002/9781118619179

Boyle, G. J. (1991). Does Item Homogeneity Indicate Internal Consistency or Item Re-

dundancy in Psychometric Scales? *Personality and Individual Differences, 12,* 291-294. https://doi.org/10.1016/0191-8869(91)90115-R

Brislin, R. W. (1970). Back-Translation for Cross-Cultural Research. *Journal of Cross-Cultural Psychology, 1,* 185-216. https://doi.org/10.1177/135910457000100301

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd Ed.). New York: Guilford Publications.

Browne, M. W., & Cudeck, R. (1993). Alternative Ways of Assessing Model Fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Newbury Park, CA: Sage.

Burlaka, V., Graham-Bermannb, S. A., & Delvac, J. (2017). Family Factors and Parenting in Ukraine. *Child Abuse & Neglect, 72,* 154-162. https://doi.org/10.1016/j.chiabu.2017.08.007

Byrne, B. M. (2010). *Structural Equation Modeling with Amos* (2nd ed.). New York: Routledge.

Byrne, B. M. (2012). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming.* London: Routledge. https://doi.org/10.4324/9780203807644

Campbell, D. T., & Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin, 56,* 81-105. https://doi.org/10.1037/h0046016

Cattell, R. B. (1966). The Scree Test for the Number of Factors. *Multivariate Behavioral Research, 1,* 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling, 14,* 464-504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling, 9,* 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Clerkin, S. M., Marks, D. J., Policaro, K. L., & Halperin, J. M. (2007). Psychometric Properties of the Alabama Parenting Questionnaire-Preschool Revision. *Journal of Clinical Child and Adolescent Psychology, 36,* 19-28. https://doi.org/10.1207/s15374424jccp3601_3

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Mahwah, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112,* 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998). The Revised Conners' Parent Rating Scale (CPRS-R): Factor Structure, Reliability and Criterion Validity. *Journal of Abnormal Child Psychology, 26,* 257-268. https://doi.org/10.1023/A:1022602400621

Costello, A. B., & Osborne, J. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. *Practical Assessment Research & Evaluation, 10,* 1-9.

Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika, 16,* 297-334. https://doi.org/10.1007/BF02310555

Dadds, M. R., Maujean, A., & Fraser, J. A. (2003). Parenting and Conduct Problems in Children: Australian Data and the Psychometric Properties of the Alabama Parenting Questionnaire. *Australian Psychologist, 38,* 238-241. https://doi.org/10.1080/00050060310001707267

Darling, N., & Steinberg, L. (1993). Parenting Style as Context: An Integrative Model. *Psychological Bulletin, 113,* 487-496. https://doi.org/10.1037/0033-2909.113.3.487

de la Osa, N., Granero, R., Penelo, E., Domènech, J. M., & Ezpeleta, L. (2014). Psychometric Properties of the Alabama Parenting Questionnaire-Preschool Revision (APQ-Pr) in 3 Year-Old Spanish Preschoolers. *Journal of Child and Family Studies, 23,* 776-784. https://doi.org/10.1007/s10826-013-9730-5

Doornik, J. A., & Hansen, H. (2008). An Omnibus Test for Univariate and Multivariate Normality. *Oxford Bulletin of Economics and Statistics, 70,* 927-939. https://doi.org/10.1111/j.1468-0084.2008.00537.x

Elgar, F. J., Waschbusch, D. A., Dadds, M. R., & Sigvaldason, N. (2007). Development and Validation of a Short Form of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies, 16,* 243-259. https://doi.org/10.1007/s10826-006-9082-5

Epskamp, S. (2019). *R Package semPlot v1.1.1.*

Esposito, A., Servera, M., Garcia-Banda, G., & Del Giudice, E. (2016). Factor Analysis of the Italian Version of the Alabama Parenting Questionnaire in a Community Sample. *Journal of Child and Family Studies, 25,* 1208-1217. https://doi.org/10.1007/s10826-015-0291-7

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Psychometric Properties of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies, 15,* 595-614. https://doi.org/10.1007/s10826-006-9036-y

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research, 18,* 382-388. https://doi.org/10.1177/002224378101800313

Frick, P. J. (1991). *The Alabama Parenting Questionnaire.* Unpublished Rating Scale, Tuscaloosa, AL: University of Alabama. https://doi.org/10.1037/t58031-000

Frick, P. J., Christian, R. E., & Wootton, J. M. (1999). Age Trends in the Association between Parenting Practices and Conduct Problems. *Behavior Modification, 23,* 106-128. https://doi.org/10.1177/0145445599231005

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of Coefficient Alpha as an Index of Test Unidimensionality. *Educational and Psychological Measurement, 37,* 827-838. https://doi.org/10.1177/001316447703700403

Gross, T. J., Fleming, C. B., Mason, W. A., & Haggerty, K. P. (2015) Alabama Parenting Questionnaire-9: Longitudinal Measurement Invariance Across Parents and Youth During the Transition to High School. *Assessment, 24,* 646-659. https://doi.org/10.1177/1073191115620839

Guadagnoli, E., & Velicer, W. F. (1988). Relation to Sample Size to the Stability of Component Patterns. *Psychological Bulletin, 103,* 265-275. https://doi.org/10.1037/0033-2909.103.2.265

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Hammer, J. H., & Toland, M. D. (2016). *Bifactor Analysis in Mplus.* Lexington, KY: University of Kentucky.

Harman, H. H. (1962). *Modern Factor Analysis* (2nd ed.). Chicago, IL: University of Chicago Press.

Harman, H. H. (1976). *Modern Factor Analysis* (3rd ed.). Chicago, IL: University of Chicago Press.

Hawes, D.J., & Dadds, M. R. (2006). Assessing Parenting Practices through Parent-Report and Direct Observation during Parent-Training. *Journal of Child and Family Studies,*

15, 555-567. https://doi.org/10.1007/s10826-006-9029-x

Heck, R. H., & Thomas, S. L. (2015). *An Introduction to Multilevel Modeling Techniques: MLM and SEM Approaches Using Mplus* (3rd ed.). New York: Routledge. https://doi.org/10.4324/9781315746494

Henze, N., & Zirkler, B. (1990). A Class of Invariant Consistent Tests for Multivariate Normality. *Communications in Statistics: Theory and Methods, 19,* 3595-3617. https://doi.org/10.1080/03610929008830400

Hershkowitz, M., Dekel, R., Fridkin, S., & Freedman, S. (2017). Posttraumatic Stress Disorder, Parenting, and Marital Adjustment among a Civilian Population. *Frontiers in Psychology, 8,* 1655. https://doi.org/10.3389/fpsyg.2017.01655

Hinshaw, S. P., Owens, E. B., Wells, K. C., Kraemer, H. C., Abikoff, H. B., Arnold, L. E., Wigal, T. et al. (2000). Family Processes and Treatment Outcome in the MTA: Negative/Ineffective Parenting Practices in Relation to Multimodal Treatment. *Journal of Abnormal Child Psychology, 28,* 555-568. https://doi.org/10.1023/A:1005183115230

Holden, G. W., & Edwards, L. A. (1989). Parental Attitudes toward Childrearing: Instruments, Issues, and Implications. *Psychological Bulletin, 106,* 29-58. https://doi.org/10.1037/0033-2909.106.1.29

Holzinger, K. J., & Swineford, F. (1937). The Bifactor Method. *Psychometrika, 2,* 41-54. https://doi.org/10.1007/BF02287965

Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika, 30,* 179-185. https://doi.org/10.1007/BF02289447

Howard, J., Gagné, M., Morin, A. J. S., Wang, Z. N., & Forest, J. (2016). Using Bifactor Exploratory Structural Equation Modeling to Test for a Continuum Structure of Motivation. *Journal of Management, 44,* 2638-2664.

Hu, L.T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling, 6,* 1-55. https://doi.org/10.1080/10705519909540118

James, D., Schumm, W., Kennedy, C., Grigsby, C., Shectman, K., & Nichols, C. (1985). Characteristics of the Kansas Parental Satisfaction Scale among Two Samples of Married Parents. *Psychological Reports, 57,* 163-169. https://doi.org/10.2466/pr0.1985.57.1.163

Jorgensen, T. (2019). *R Package SemTools v0.5-1.*

Kaiser, H. (1970). A Second Generation Little Jiffy. *Psychometrika, 35,* 401-415. https://doi.org/10.1007/BF02291817

Kaiser, H. (1974). An Index of Factorial Simplicity. *Psychometrika, 39,* 31-36. https://doi.org/10.1007/BF02291575

Kelley, T. L. (1935). *Essential Traits of Mental Life, Harvard Studies in Education* (Vol. 26). Cambridge, MA: Harvard University Press.

Kline, P. (1999). *Handbook of Psychological Testing.* London: Routledge.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York: The Guilford Press.

Korkmaz, S. (2019). *MVN 5.7: An R Package for Assessing Multivariate Normality (R Package).* Edirne, Turkey: Trakya University.

Kuhn, M. (2019). *R Package Caret V.6.0-84.*

Kyriazos, T. A. (2018a). Applied Psychometrics: The 3-Faced Construct Validation Method, a Routine for Evaluating a Factor Structure. *Psychology, 9,* 2044-2072. https://doi.org/10.4236/psych.2018.98117

Kyriazos, T. A. (2018b). Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology, 9,* 2207-2230. https://doi.org/10.4236/psych.2018.98126

Kyriazos, T. A., & Stalikas, A. (2019a). *Validation of the Greek Version of the Parenting Behaviours and Dimensions Questionnaire (PBDQ).* Under Preparation.

Kyriazos, T. A., & Stalikas, A. (2019b). *Validation of the Greek Version of the Parent Behavior Inventory (PBI).* Under Preparation.

Kyriazos, T. A., & Stalikas, A. (2019c). *Validation of the Greek Version of the Parent Concerns Questionnaire (PCQ).* Under Preparation.

Kyriazos, T. A., & Stalikas, A. (2019d). *Validation of the Greek Version of the Parental Stress Scale (PSS).* Under Preparation.

Kyriazos, T. A., & Stalikas, A. (2019e). *Validation of the Kansas Parental Satisfaction Scale (KPSS) in a Greek sample.* Under Preparation.

Kyriazos, T. A., Stalikas, A., Prassa, K., & Yotsidi, V. (2018). Can the Depression Anxiety Stress Scales Short Be Shorter? Factor Structure and Measurement Invariance of DASS-21 and DASS-9 in a Greek, Non-Clinical Sample. *Psychology, 9,* 195-1127.

Little, T. D. (2013). *Longitudinal Structural Equation Modeling.* New York: Guilford Press.

Locke, L. M., & Prinz, R. J. (2002). Measurement of Parental Discipline and Nurturance. *Clinical Psychology Review, 22,* 895-929. https://doi.org/10.1016/S0272-7358(02)00133-2

Lorezo-Seva, U., & Ferrando, J. P. (2013). *Factor v.9.20 Computer Software.* Tarragona, Spain.

Lovejoy, M. C., Weis, R., O'Hare, E., & Rubin, E. C. (1999). Development and Initial Validation of the Parent Behaviour Inventory. *Psychological Assessment, 11,* 534-545. https://doi.org/10.1037/1040-3590.11.4.534

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods, 1,* 130-149. https://doi.org/10.1037/1082-989X.1.2.130

Maguin, E., Nochajski, T. H., De Wit, D. J., & Safyer, A. (2016). Examining the Validity of the Adapted Alabama Parenting Questionnaire—Parent Global Report Version. *Psychological Assessment, 28,* 613-625. https://doi.org/10.1037/pas0000214

Mahalanobis, P. C. (1936). On the Generalized Distance in Statistics. *Proceedings of the National Institute of Science, India, 12, 49-55.*

Mair, P. (2018). *Modern Psychometrics with R.* Cham, Switzerland: Springer International. https://doi.org/10.1007/978-3-319-93177-7

Mardia, K. V. (1970). Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika, 57,* 519-530. https://doi.org/10.1093/biomet/57.3.519

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. *Annual Review of Clinical Psychology, 10,* 85-110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

McDonald, R. P. (1970). The Theoretical Foundations of Principal Factor Analysis, Canonical Factor Analysis, and Alpha Factor Analysis. British *Journal of Mathematical and Statistical Psychology, 23,* 1-21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x

McDonald, R. P. (1999). *Test Theory: A Unified Treatment.* Mahwah, NJ: Erlbaum.

Molinuevo, B., Pardo, Y., & Torrubio, R. (2011). Psychometric Analysis of the Catalan

Version of the Alabama Parenting Questionnaire (APQ) in a Community Sample. *The Spanish Journal of Psychology, 14,* 944-955. https://doi.org/10.5209/rev_SJOP.2011.v14.n2.40

Nitsch, E., Hannon, G., Rickard, E., Houghton, S., & Sharry, J. (2015). Positive Parenting: A Randomized Controlled Trial Evaluation of the Parents plus Adolescent Programme in Schools. *Child and Adolescent Psychiatry and Mental Health, 9,* 43. https://doi.org/10.1186/s13034-015-0077-0

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.

Pickering, J. A., & Sanders, M. R. (2016). Reducing Child Maltreatment by Making Parenting Programs Available to All Parents: A Case Example Using the Triple P-Positive Parenting Program. *Trauma, Violence, & Abuse, 17,* 398-407. https://doi.org/10.1177/1524838016658876

R Development Core Team (2019). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Randolph, K. A., & Radey, M. (2011). Measuring Parenting Practices among Parents of Elementary School-Age Youth. *Research on Social Work Practice, 21,* 88-97. https://doi.org/10.1177/1049731509353048

Rasmussen, K. N. (2009). Effective Parenting. In S. Lopez (Ed.), *The Encyclopedia of Positive Psychology* (pp. 291-296). Chichester: Blackwell Publishing Ltd.

Raykov, T. (2001). Estimation of Congeneric Scale Reliability Using Covariance Structure Analysis with Nonlinear Constraints. *British Journal of Mathematical and Statistical Psychology, 54,* 315-323. https://doi.org/10.1348/000711001159582

Reid, C. A. Y., Roberts, L. D., Roberts, C. M., & Piek, J. P. (2015). Towards a Model of Contemporary Parenting: The Parenting Behaviours and Dimensions Questionnaire. *PLoS ONE, 10,* e0114179. https://doi.org/10.1371/journal.pone.0114179

Revelle, W. (2019). *Package Psych V1.8.12: Procedures for Psychological, Psychometric, and Personality Research.* Evanston, IL: Northwestern University.

Revelle, W., & Rocklin, T. (1979) Very Simple Structure—Alternative Procedure for Estimating the Optimal Number of Interpretable Factors. *Multivariate Behavioral Research, 14,* 403-414. https://doi.org/10.1207/s15327906mbr1404_2

Robert, C. J. (2009). *Parenting Practices and Child Behavior in Mexico: A Validation Study of the Alabama Parenting Questionnaire.* ProQuest Dissertations & Theses Full Text.

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48,* 1-36. https://doi.org/10.18637/jss.v048.i02 http://www.jstatsoft.org/v48/i02/

Schumacker, R. E., & Lomax, R. G. (2016). *A Beginner's Guide to Structural Equation Modeling* (4th ed.). New York: Routledge.

Schwartz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics, 6,* 461-464. https://doi.org/10.1214/aos/1176344136

Scott, S., Briskman, J., & Dadds, M. R. (2011). Measuring Parenting in Community and Public Health Research Using Brief Child and Parent Reports. *Journal of Child and Family Studies, 20,* 343-352. https://doi.org/10.1007/s10826-010-9398-z

Seabridge, S. D. (2012). *Examining the Link between Parenting and Child Problem Behaviors in American Indian Families.* Stillwater, OK: Oklahoma State University.

Sha, S., & Ackerman, T. (2018). The Performance of Five Reliability Estimates in Multi-

dimensional Test Situations. In L.A. van der Ark et al. (Eds.), *Quantitative Psychology* (Vol. 196, pp. 173-181). Cham: Springer. https://doi.org/10.1007/978-3-319-56294-0_16

Shelton, K. K., Frick, P. J., & Wootton, J. (1996). Assessment of Parenting Practices in Families of Elementary School-Age Children. *Journal of Clinical Child Psychology, 25,* 317-329. https://doi.org/10.1207/s15374424jccp2503_8

Sheppard, M. (2010). The Parent Concerns Questionnaire: A Reliable and Valid Common Assessment Framework for Child and Family Social Care. *The British Journal of Social Work, 40,* 371-390. https://doi.org/10.1093/bjsw/bcn163

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the Sins of Short-Form Development. *Psychological Assessment, 12,* 102-111. https://doi.org/10.1037/1040-3590.12.1.102

Steiger, J. H., & Lind, J. C. (1980). Statistically Based Tests for the Number of Common Factors. Paper Presented at the Psychometric Society Annual Meeting, Iowa City, IA.

Święcicka, M., Woźniak-Prus, M., Gambin, M., & Stolarski, M. (2019). *Confirmation of the Five-Factor Structure of the Parent Global Report Version of the Alabama Parenting Questionnaire in a Polish Community Sample.* Current Psychology, 1-13. https://doi.org/10.1007/s12144-019-00340-8

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.) Boston, MA: Allyn & Bacon/Pearson Education.

Timmons, A. C. (2010). *Establishing Factorial Invariance for Multiple-Group Confirmatory Factor Analysis.* KU Guide No. 22.1.

Tucker, L. R., & Lewis, C. (1973). A Reliability Coefficient for Maximum Likelihood Factor Analysis. *Psychometrika, 38,* 1-10. https://doi.org/10.1007/BF02291170

Ullman, J. B. (2013). Structural Equation Modeling (Chapter 14). In B. Tabachnick, & L. Fidell (Eds.), *Using Multivariate Statistics* (pp. 681-785). Boston, MA: Pearson Education Inc.

Velicer, W. F. (1976). Determining the Number of Components from the Matrix of Partial Correlations. *Psychometrika, 41,* 321-327. https://doi.org/10.1007/BF02293557

Wang, J., & Wang, X. (2012). *Structural Equation Modeling.* Beijing: Higher Education Press. https://doi.org/10.1002/9781118356258

Wang, L. L., Watts, A. S., Anderson, R. A., & Little, T. D. (2013). Common Fallacies in Quantitative Research Methodology. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods* (pp. 718-758). New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199934898.013.0031

Werts, C. E., Linn, R. N., & Joreskog, K. G. (1974). Interclass Reliability Estimates: Testing Structural Assumptions. *Educational & Psychological Measurement, 34,* 25-33. https://doi.org/10.1177/001316447403400104

Wickham, H. (2019a). *Package Haven V 2.1.1.*

Wickham, H. (2019b). *R Package Dplyr v0.7.8.*

Widaman, K. F. (1985). Hierarchically Nested Covariance Structure Models for Multi-trait-Multimethod Data. *Applied Psychological Measurement, 9,* 1-26. https://doi.org/10.1177/014662168500900101

Xie, Y. (2019). *R Package knitr V. 1.2: Dynamic Report Generation.*

Yuan, K. H., & Bentler, P. M. (2000). Three Likelihood-Based Methods for Mean and Covariance Structure Analysis with Nonnormal Missing Data. *Sociological Methodology, 30,* 165-200. https://doi.org/10.1111/0081-1750.00078

Zlomke, K. R., Lamport, D., Bauman, S., Garland, B., & Talbot, B. (2014). Parenting Adolescents: Examining the Factor Structure of the Alabama Parenting Questionnaire for Adolescents. *Journal of Child and Family Studies, 23,* 1484-1490. https://doi.org/10.1007/s10826-013-9803-5