Equivalence of Family Functioning and Externalizing in Adolescent Substance Users of Different Race/Ethnicity

Daniel J. Feaster[1], **Michael S. Robbins**[2], **Craig Henderson**[3], **Viviana Horigian**[1], **Marc J. Puccinelli**[1], **A. Kathy Burlew**[4], and **José Szapocznik**[1]

[1]Department of Epidemiology and Public Health, University of Miami Miller School of Medicine

[2]Oregon Research Institute

[3]Sam Houston State University

[4]University of Cincinnati

Abstract

Brief Strategic Family Therapy (BSFT) for Adolescent Drug Abuse clinical trial of 480 adolescents boys and girls age 12 to 17 and their parents was designed to maximize the chance that a sufficient number of Hispanic and Black adolescents would be included to allow valid subgroup comparisons. Examination of measurement invariance is an important step to ensure valid analysis. Two construct areas important to the analysis of trial results, adolescent problem behaviors and family functioning showed a high degree of measurement invariance, which allowed valid comparisons of mean baseline differences across groups. Results showed that Black families had significantly higher initial levels of family functioning and lower levels of adolescent externalizing than either Hispanic or White non-Hispanic families. This pattern is consistent with an increased likelihood of referral of Black adolescents with more severe problems to restricted setting rather than to outpatient drug abuse treatment. This possibility highlights the importance of considering differing baseline characteristics of subgroups prior to assessing differential treatment effectiveness to prevent confounding.

Keywords

Measurement Invariance; Ethnic Minorities; Racial/Ethnic Variability; Adolescent problem behaviors

---

The most recent data from the US Bureau of the Census (2000, 2004) indicates that racial/ethnic minorities are projected to account for one third of the US population as soon as 2010. The adolescent population is even more diverse, with 36% of 10-19 years-old coming from racial/ethnic minority backgrounds, compared to 28% of the general population (US Census Bureau, 2000).

Understanding racial/ethnic variation in important adolescent outcomes such as substance use has historically been overlooked. For example, Hall and Maramba (2001) report that

---

"Publications involving cross-cultural issues represented 1% and racial/ethnic minority issues represented 3% of all the publications in the English language in the PyscINFO database" (p. 15), although content analyses of journals have shown an increasing trend in authors reporting at least their samples' racial/ethnic characteristics (Imade & Schiavo, 2005).

As socio-behavioral aspects of race/ethnicity continue to emerge as an area of substantive inquiry, clinical researchers will increasingly need to address them in their research. Indeed, the National Institutes of Health (NIH) have developed guidelines for the inclusion of women and minorities in clinical research. These recommendations focus on two broad areas: inclusion and valid analysis (NIH Guidelines, 2001). While the citations above indicate that researchers have begun to engage more ethnic minorities in research studies (including developing structured protocols for engaging and retaining minorities in treatment; Robbins, Horigian, Feaster, Szapocznik, Burlew, et al., 2009, Santisteban et al., 1996; Szapocznik et al., 1988), Burlew, Feaster, Brecht, and Hubbard (2009) make clear that research that takes race and ethnicity seriously must also focus on valid analysis (NIH guidelines 2001).

Burlew et al. (2009) focus on two primary issues in valid analysis, measurement and statistical analysis, and provide researchers with practical guidance and specific recommendations intended to assist researchers in translating the NIH mandates into research practice, which we implement in the current study. At a molar level, Burlew et al. address the cultural appropriateness of assessment instruments, focusing more specifically on the issue of whether items on assessment instruments mean the same thing--measurement equivalence--across racial/ethnic groups. Though what is of interest is establishing the equivalence of measurement instruments across groups, statistically this involves tests of whether various aspects of the measures are invariant across groups. Fundamentally, lack of "equivalence on such measures may lead to misleading conclusions about the etiology of substance abuse, its effects, or appropriate treatments" (p. 5). Any analysis of comparative effectiveness of treatments across groups must first ensure that the measures used to assess treatment effectiveness have the same meaning across groups.

The research on treatment outcomes for ethnic minority adolescents remains limited. Studies that evaluate the efficacy of treatment modalities specifically for ethnic minorities are rare (Strada, Donohue & Lefforge, 2006). In fact, a review of the literature by Szapocznik, Prado, Burlew, Williams, & Santisteban (2007) reported that the only treatments with demonstrated efficacy for Hispanic or Black adolescents are Brief Strategic Family Therapy and Multisystemic Therapy, respectively. Moreover, the available evidence suggests that the treatment outcomes for ethnic minority youth are less favorable than the treatment outcomes for their White non-Hispanic counterparts. Specifically, ethnic minority youth are less likely to become engaged in treatment (Jackson-Gilfort, Liddle, Tejeda & Dakof, 2001, Perrino, Coatsworth, Briones, Pantin, & Szapocnik, 2001), less likely to complete treatment (Campbell, Weisner, & Sterling, 2006), and more likely to have a negative termination from treatment (Shillington & Clapp, 2003).

Differential treatment outcomes may reflect several important differences between ethnic minority and White non-Hispanic adolescents that may potentially be related to the fit between many existing treatment modalities and ethnic minority client characteristics. First, ethnic minorities appear to differ from White adolescents in their drug histories and their route to treatment. For example, Black adolescents are more likely than White non-Hispanic adolescents to report marijuana as their primary drug and to be mandated to treatment (Shillington & Clapp, 2003). In addition, the available research suggests that family factors and involvement in externalizing behaviors warrant further consideration in the design of effective treatments. Although the unique aspects of family that are related to substance use may vary across ethnic groups, family factors have been demonstrated to be related to substance

use among Black (Strada & Donohue, 2006), Hispanic (Gil, Wagner, & Vega, 2000), American Indian/Alaska Native (Hawkins, Cummins, & Marlatt, 2004), and Asian (Kim, Zane & Hong (2002) adolescents. Comparing these predictors of treatment outcome will help to prevent the confounding of initial differences with differential treatment effectiveness. The establishment of measurement invariance is a necessary first step to examining these pre-existing difference across groups.

There have been investigations of the equivalence of particular externalizing and family measures. Examination of the Child-Behavior Checklist have shown that the externalizing scale is largely equivalent across racial-ethnic groups (Achenbach et al., 2008, Gross et al., 2006). There is also some evidence of reliability of delinquency measures across race/ethnicity (Huizenga & Elliot, 1986, Meyers et al., 2006, Smith & Krohn, 1995), though these have not utilized confirmatory factor analysis. Finally, there has also been some examination of cross-cultural reliability and validity of family scales (Tolan et al., 1997, Munet-Vilaro, 1990), though these again have not utilized confirmatory factor analysis. Whereas there is evidence that some particular scales have been equivalent in particular studies, the equivalence of measures may be sample-specific. In addition, the emphasis of most investigations has been on particular scales, rather than on higher-order constructs.

This study reports the results of a measurement invariance analysis of adolescent externalizing behaviors and family functioning from NIDA's Clinical Trials Network Protocol 14 (CTN 0014), Brief Strategic Family Therapy (BSFT™) for Adolescent Drug Abusers. This effectiveness protocol is one of the largest adolescent psychotherapy trials completed to date including data from 480 adolescents and their parent/guardians. The protocol was implemented by community-based treatment providers and put no restrictions, limits or quotas on the race/ethnicity of enrollees. However, every effort was made to maximize racial and ethnic representation by inclusion of treatment providers that serve Hispanics and Blacks.

The BSFT protocol (Feaster, Robbins, Horigian, & Szapocznik, 2004, Robbins, Szapocznik, Horigian, Feaster, Puccinelli, et al., 2009) tests the effectiveness of an empirically validated family therapy intervention (BSFT) in reducing adolescent problem behaviors and drug use. BSFT is a strategic and structural family therapy that targets family interactions; the hypothesized mechanism of action is improvement in family functioning. Both family functioning and adolescent problem behaviors are assessed by multiple established measures reported by both the adolescent and his/her parent guardian. It is important to understand whether the multiple indicators of both family functioning and adolescent problem behaviors are similarly inter-related across racial/ethnic groups for valid comparisons across these groups on the higher-order constructs of family functioning and externalizing.

The purpose of this manuscript is three-fold. The first aim is to give a practical example of how to examine measurement equivalence of assessment instruments across subgroups to ensure appropriate analysis. The second aim is to test whether specific measures (i.e., adolescent externalizing behavior and family functioning) are invariant across racial/ethnic groups at the initial assessment of the BSFT study. The final aim is to examine differences in adolescent externalizing and family functioning across racial/ethnic groups. This final aim will be addressed qualitatively, if the measurement model is found to differ across groups and quantitatively if there are sufficient levels of invariance across groups in the measurement model. Given prior research findings, we anticipated finding sufficient evidence of equivalence to allow at least some quantitative comparisons.

Methods

Study Procedures

Drug using adolescents and their parents or guardian were enrolled into a randomized controlled trial comparing BSFT to treatment as usual (TAU) at each of 8 outpatient community treatment programs across the country (Robbins et al., 2009). Sites were volunteers from within the National Institute of Drug Abuse' National Drug Abuse Treatment Clinical Trials Network. This research was approved by the Institutional Review Boards at all participating sites.

To enroll in the study, adolescent participants had either to self-report illicit drugs (other than alcohol and tobacco) in the 30-day period that preceded the baseline assessment *or* be referred from an institution (e.g., detention, residential treatment, court etc.) for the treatment of substance abuse. Also, the adolescent had to assent and a parent or legal guardian had to provide informed consent to participate in therapy. Further eligibility criteria included living with a family (defined to include any parental/adult guardian, except foster) and in the same geographical area as the agency (because home-based services were expected). Adolescents with current or pending severe criminal offences that would likely result in short-term or long-term incarceration were excluded from participation. Adolescents in non-restricted settings were excluded if they were already receiving drug treatment.

Study Sample

Four hundred and eighty adolescents and their family members participated in this study. Family members consist of biological relations as well as other individuals that resided in the adolescent's home and that were involved in the adolescent's life on a daily basis. Adolescents were predominately male ($n = 377$ versus 103 female), with a mean age of 16.02 ($SD = 1.32$). Based on adolescent self reports, the sample included 213 Hispanic/Latino, 148 White non-Hispanic, and 110 Black/African Americans and 9 other race/ethnicity. Note that the nine individuals reporting other race/ethnicity are excluded from the analyses with a resulting sample size of 471. Most families were biological one-parent ($n = 224$), two-parent ($n = 120$) households, or extended/blended families ($n = 117$) with the remaining families describing themselves as other. and Approximately half of the families reported a household income of less than $25,000.

Measures

The BSFT trial included multiple measures from multiple reporters of both family functioning and adolescent problem behaviors. These scales are utilized as indicators of a measurement model for the externalizing behaviors and family functioning constructs (explained below). The scales scores are utilized rather than the individual items to minimize the number of indicators within the measurement models; it would not be possible to test invariance of the measurement model described below utilizing items with the current sample size. An additional advantage of the use of scales rather than items is that scales composed of sums of items are more likely to have an approximately normal distribution. The utilization of scale scores may mask some sources of invariance (Chen, Sousa & West, 2005), however, as noted in the introduction, there has been research showing that many of the scales utilized in this investigation are themselves invariant across racial/ethnic groups. The reliability for these indicator scales is generally adequate or good, though a few are in the marginal range. This further highlights the need to utilize measurement modeling both to improve reliability and ensure that the underlying constructs work equally well for all groups.

Externalizing Behavior Measures—The *Diagnostic Interview Schedule for Children (DISC) Predictive Scales* are a series of diagnosis-specific short-forms of the DISC IV composed of yes-no questions, that were administered to both youth and parents to assess the

adolescent's psychological functioning. Both parent and adolescent reports of oppositional defiant disorder (10 items parent, 7 adolescent and conduct disorder (15 items parent, 8 adolescent) are used as indicators of externalizing behaviors of the adolescent. Separated by race/ethnicity, reliability scores (Cronbach's Alpha) for adolescent report of conduct disorder were: Hispanic α=.69, Black α=.55, and White non-Hispanic α=.60. For the parent report of conduct disorder, alpha's were: Hispanic α=.73, Black α=.81, and White non-Hispanic α=.79. The oppositional defiant disorder reliability scores were: adolescent report - Hispanic α=.65, Black α=.65, White non-Hispanic α=.66, and parent report - Hispanic α=.81, Black α=.86, and White non-Hispanic α=.80.

The *Youth Self-Report* (YSR)(Achenbach & Rescorla, 2001) is a two part measure used to asses a child's view of both his/her own social competence as well as problem behaviors. Problem behaviors can be scored along the dimensions "internalizing" and "externalizing" behaviors. The externalizing scale composed of 32 items on a 3 point scale (0-not true, 1 somewhat true, 2-very or often true) is used as an indicator of the externalizing factor, herein. Cronbach's alphas for externalizing domain of the YSR were: Hispanic α=.88, Black α=.86, and White non-Hispanic α=.89.

The *Self-Report Delinquency Scale* (Elliot et al., 1983) consists of 23 items from the National Youth Survey (Huizinga & Elliot, 1983). The primary purpose of this measure is to provide self-reports of adolescent delinquent behaviors. The *Total Delinquency variety score* (Huizinga & Elliott, 1986), which is a sum of the items with any endorsement is used as an indicator of adolescent externalizing behaviors. Cronbach's alphas for the Total Delinquency variety subscale were: Hispanic α=.84, Black α=.75, and White non-Hispanic α=.83.

Family Functioning—The *Family Environmental Scale* (FES; Moos & Moos, 1986) is a widely used measure of yes-no items developed to determine social and environmental characteristics of families. In this study, we used only the cohesion and conflict subscales (9 items each) of the FES, administered to both parents and adolescents. Cronbach's alphas for the FES conflict subscale (adolescent report) were: Hispanic α=.66, Black α=.66, and White non-Hispanic α=.74, and for the parent report: Hispanic α=.70, Black α=.72, and White non-Hispanic α=.75. Alphas for the Coherence subscale were: (adolescent report) Hispanic α=.72, Black α=.60, White non-Hispanic α=.75, and (parent report) Hispanic α=.74, Black α=.79, and White non-Hispanic α=.75.

The Parenting Practices Questionnaire was derived from the parental supervision and discipline interview used in the Oregon Youth Study and Pittsburgh Youth Study (Thornberry, Huizinga, & Loeber, 1995). Items are a mixture of 3 and 5 point ordered likert-type response categories. Both parent and adolescent reports were used in this study. The indicators of family functioning used herein are four scales for parenting practices: (a) Positive Parenting (6 items), (b) Discipline Effectiveness (5 items), (c) Avoidance of Discipline (7 items), and (d) Monitoring (13 items parent, 12 items adolescent) (Gorman-Smith et al., 1996). Reports of discipline effectiveness and avoidance of discipline were gathered from parents only. Estimates of positive parenting and extent of monitoring were gathered from both parent and child. Internal consistency reliabilities (Cronbach's α) of each of the subscales ranged from .66 to .89. Divided by ethnic group, alphas for the *positive parenting* subscale of the adolescent report were: Hispanic α=.84, Black α=.87, White non-Hispanic α=.78, and for the parent report: Hispanic α=.74, Black α=.73, and White non-Hispanic α=.66. For the *monitoring by parent* subscale, alphas were: (adolescent report) Hispanic α=.85, Black α=.84, White non-Hispanic α=.89, and (parent report) Hispanic α=.78, Black α=.80, and White non-Hispanic α=.72. For the two subscales administered to parents only, alphas were as follows: (*avoidance of discipline*) Hispanic α=.75, Black α=.84, White non-Hispanic α=.80, and (*discipline effectiveness*) Hispanic α=.72, Black α=.81, and White non-Hispanic α=.82.

Hypothesized Measurement Model: The hypothesized measurement model is depicted in Figure 1. There are four hypothesized latent constructs: Family Functioning and Externalizing from both the parent and adolescent perspectives. From prior experience of the authors as well as others (Achenbach, McConaughy & Howell, 1987,De Los Reyes & Kazdin, 2004,Youngstrom, Loeber, & Strouthamer-Loeber, 2000), it was decided to fit separate latent constructs for parent and adolescent reports. The latent factor for parent report of externalizing included two indicators both from the parent-reported version of DISC Predictive Scales-- Conduct Disorder and Oppositional Defiant Disorder. The latent factor for adolescent report of externalizing is measured by four indicators: DISC Predictive Scales Conduct Disorder and Oppositional Defiant Disorder, the YSR Externalizing scale and the NYS Delinquency Variety score. The latent factor for parent report of family functioning was composed of two subscales from the Family Environment Scale, Cohesion and Conflict and 4 indicators from the Parenting Practices Scale: Positive Parenting, Parental Monitoring, Effective Discipline and Discipline Avoidance. The final latent factor for adolescent reported family functioning consisted of two subscales from the Family Environment Scale, Cohesion and Conflict, as well as two subscales from the Parenting Practices Scale, Parental Monitoring and Positive Parenting. All four latent factors were freely inter-correlated as depicted be the connecting lines with two-headed arrows.

Analytic Methods—All analyses were conducted using Mplus 5.21 (Muthén & Muthen, 1998-2009). In initial analyses scales utilized within the measurement model were examined for normality. The steps for the assessment of the equivalence of measures by race/ethnicity were as follows. First, the unstructured variance-covariance matrix of the indicators was tested for differences across race/ethnicity. This is an overall test of whether the items in the measurement model are invariant across racial/ethnic groups. After comparing unstructured covariance matrices across the three /racial/ethnic categories, the hypothesized measurement model was fit as a multi-group model with parameters free to vary across groups. This model is used to establish configural invariance, which means that the same items measure the same latent constructs across the three groups. The model fit was examined using the Confirmatory Fit Index (*CFI*) and the root mean square error of approximation (RMSEA; Kline, 2002). A structural equation model with excellent fit to the data will have *CFI* ≥.95 and *RMSEA* ≤.06; whereas, an adequate fit has *CFI* ≥.90 and *RMSEA* ≤.08 (McDonald & Ho, 2002). For this initial unconstrained multi-group measurement model any necessary minor modifications to improve fit were assessed. Preference was given to modifications that involved correlating residuals of scales from the same assessment instrument that were within the same latent construct which is an approach for controlling for a methods effect. This modified model was then subjected to a series of increasingly stringent restrictions on the parameters across groups to identify the source and qualities of any invariance between the groups (Cheong & Rensvold, 2002, Vandenberg & Lance, 2000). This series of models was compared to ascertain to what extent the measurement of the proposed latent factors was invariant across the three major racial/ ethnic categories represented in the sample: Hispanic, Black, and White non-Hispanic. Qualitative differences in the final measurement model were enumerated and implied composite reliability was calculated (Reuterberg & Gustafsson, 1992). Finally, comparisons of variances, covariances and means of latent factors across the three ethnic/racial groups were performed where possible given the results of the measurement invariance analysis.

Steps to determine level of invariance: There are numerous approaches to ascertaining measurement invariance (Burlew, Feaster, Brecht & Hubbard, 2008, Vandenberg & Lance, 2000). Some (Vandenberg & Lance) recommend an initial omnibus test of equivalent variance-covariance matrices across groups. According to the logic of performing this test, if there is no statistical evidence that these raw covariance matrices of indicators is different across groups there is no reason to go further in examining measurement invariance. Others (Burlew et al. & Woehr, Arciniega & Fowler, 2003) believe that this test is overly restrictive because there may

be important differences in smaller pieces of the measurement model that would be masked in this overall test. We include the test of equality of covariance matrices for completeness but planned a series of measurement models with even more stringent constraints on parameters across groups regardless of the results of the test. When faced with a rejection of the null hypothesis associated with one of these steps, modification indices were examined to find the most likely indicators contributing to the rejection. These particular parameters were then allowed to vary across the appropriate groups and the model reassessed. If this model with a few of the most recently imposed restrictions relaxed was not rejected relative to the prior even less restrictive model, then the procedure continued and a new set of restrictions were tested.

The first multiple group model to be estimated with the hypothesized measurement structure allowed all parameters across the three racial/ethnic groups to differ. An acceptable level of model fit for this model is indicative of configural invariance. Configural invariance implies that the same set of items measure the same underlying factors across the groups. The first set of restrictions to be tested, equal item loadings across groups, is referred to as metric invariance (Vandenberg & Lance, 2000) or weak invariance (Meredith, 1993). It implies that a unit increase on a particular item implies the same proportionate increase in the level of the latent construct for all groups (where the estimated loading is the coefficient of proportionality). Metric invariance implies that the scale of the latent factors are the same across the groups being compared; a unit increase in one group implies the same change in the underlying factor as it does for the other groups. This alone is not sufficient to make comparisons of levels of the factor across groups because it does not imply that the zero-point on the scales are the same across groups (Chen, Sousa & West, 2005). The next set of restrictions tested, equal item intercepts, is referred to as scalar invariance (Vandenberg & Lance) or strong invariance (Meredith). The item intercept is the mean level of endorsement of the item when the latent factor is equal to zero. Therefore, equal item intercepts imply that the zero-point on the underlying factor is the same across groups. If the item intercepts differ across groups, it implies that there are different normative levels of the item across groups. That is, different groups require different levels of the item to have the same underlying level of the latent factor. Note that at least one of the indicators within each construct must have the same item loading and the same item intercept across all groups to make comparisons of factor means across the groups (Byrne, Muthén & Shavelson, 1989, Vandenberg & Lance, 2000).

The establishment of what Meredith calls strict equivalence requires both of the prior two levels of equivalence (weak and strong) and equal item variances. The final two sets of restrictions to be tested were equal item variance which were then followed by testing of equal factor variances. Note that failure of equal item variances and/or factor variances across groups will cause standardized loadings and factor reliability to differ between groups even when metric invariance or equal (unstandardized) loadings holds. The final tests completed include tests of equal factor covariances and then equal factor means. Note that the test of means depends on having at least one invariant item loading and item intercept across groups for each of the latent factors, as described above.

Results

The results in the manuscript focus on the steps of invariance as described above. Additional detail concerning the measurement model including the item (scale) means and raw variance covariance estimates by racial ethnic groups as well as Mplus input and output files are available from the corresponding author. Initial examination of the scale scores showed some deviation from normality and therefore the Mplus MLR estimator was used which yields standard errors and scaled chi-square difference tests that are robust to non-normality (Satorra, 2000; Muthén & Muthen, 1998-2007)

Table 1 shows the indicator scales that differed by race/ethnicity. The majority of indicators in each latent construct other than adolescent report of family functioning showed an overall difference amongst the three racial/ethnic groups. In addition, the unstructured covariance matrices across the three racial/ethnic groups were significantly different ($\chi^2(272, n = 471) = 328.2, p < .02$).

Configural Invariance

The initial multi-group fit of the model constrained one factor loading per latent construct to be equal 1. For Parent Report of Externalizing, the loadings of both of the indicators were constrained to one, to identify the latent factor. In the initial model factor means for all subgroups were set to zero and item intercepts were allowed to vary. This initial multi-group model had a *CFI* of .86 and *RMSEA* of .090. All loadings were significant in all ethnic/racial groups, indicating configural invariance. It was decided to examine a similar model which correlated the 5 pairs of items that were subscales from the same assessment instrument to control for methods effects and to improve the fit of the model. These pairs were: Family cohesion and family conflict (for both adolescent and parent report), parental monitoring and positive parenting (for both adolescent and parent report) and effective discipline and discipline avoidance (parent report only). The model including these five pairs of residual correlation had a *CFI* of .91 and an *RMSEA* of .072.

Metric Invariance

The next step was to constrain all the factor loadings to be equal across groups to test metric invariance. This model had a *CFI* of .91 and an *RMSEA* of .070. In addition, this model was not significantly different from the model in which loadings were unconstrained ($\chi^2(22, n = 471) = 14.7, p < .88$). This means that we fail to reject that the item loadings are significantly different across ethnic/racial groups. Hence, there is metric invariance. This metric invariance implies that a unit increase in any one indicator implies an equal difference in the latent construct across all the racial/ethnic groups.

Scalar Invariance

The following step was to constrain the remaining item intercepts to be equal across groups to test scalar invariance. This model had a *CFI* of .90 and *RMSEA* of .074, however, this model did show a significantly poorer fit to the data than the model with equal loadings and free intercepts ($\chi^2(24, n = 471) = 74.7, p < .0001$). Examination of the modification indices revealed that intercepts for 5 indicators needed to be freed to improve fit. Intercepts for adolescent report of conduct disorder, parent reports of both parental monitoring and family cohesion needed to be free to vary amongst all three groups. For adolescent report of conduct disorder Hispanic families had the largest intercept (1.72), followed by Black families (1.64) and White non-Hispanic families with the smallest intercept (1.29). For the parent report of family cohesion the intercepts were largest for Hispanic families (6.19), with Black families in the middle (5.84) and the intercept for White non-Hispanic families (5.57) the smallest. In addition the intercepts for adolescent report of oppositional defiant disorder (3.28) and parental monitoring (35.01) were higher for White non-Hispanic families than for Hispanic and Black families intercepts which remained equated (2.96 and 33.08 respectively). The model freeing just these intercepts had a *CFI* of .91 and *RMSEA* of .069 and was not significantly different from the final model in *Metric Invariance* section, above ($\chi^2(16, n = 471) = 21.2, p < .17$). This indicates that we have partial scalar invariance--for most indicators the same level of the indicator scales were associated with the same value on the latent variable across racial/ethnic groups. This also implies that is possible to make valid factor mean comparisons across racial ethnic groups.

Invariance of Residual Variances

The next step was to test equal residual variances of the latent construct indicator variables across racial/ethnic groups. If the final measurement model shows the factor variances to be equal then this test also is the test of equal reliability of factors across race/ethnicity (Vandenburg and Lance, 2000). The model imposing equal residual variances had comparable model fit indices to the final model in the scalar invariance section, *CFI* of .91, *RMSEA* of .069, however the equal residual variance restrictions as a group were rejected by the likelihood ratio test (χ^2(32, $n = 471$) = 66.8, $p < .0001$). This implies that reliabilities of the latent factors do vary by race/ethnicity. Modification indices indicated five indicators had unequal residual variances. The residual variance of parental report of monitoring was different across all ethnic/racial groups. Hispanic families had the largest residual variability (18.2), followed by Black families (14.6) with White non-Hispanic families showing the smallest residual variance (9.5). In addition the residual variances for positive parenting (5.7) and avoidance of discipline (5.0) were higher for Hispanic families than for either White non-Hispanic and Black families whose residual variances were held equal to each other (4.0 and 3.0, respectively). Finally, the residual variances of the adolescents report of the NYS delinquency variety score (2.7) and the DISC predictive scale conduct disorder (.8) were lower for Black families than for White non-Hispanic and Hispanic families whose residual variances were held equal to each other (6.2 and 1.3, respectively). The resulting model, *CFI* of .91, *RMSEA* of .067, and was not significantly different from the final model in the *Scalar Invariance* section (χ^2(26, $n = 471$) = 32.8, $p < .17$).

Equality of Factor Variances and Covariances

The initial test for equal factor variances was rejected (χ^2(8, $n = 471$) = 18.38, $p < .02$). Examination of the model showed that for Black families, the factor variances for adolescent's report of externalizing (6.80) and family functioning (2.16) were lower than those for Hispanic and White non-Hispanics (7.43 and 2.54, respectively). In addition, again for Black families, the factor variance for parent report of externalizing had a greater variance (6.40) than the factor variance for Hispanic and White non-Hispanic families (2.54). Family functioning by parent report had the same factor variance (.60) across all racial/ethnic groups. The model which imposed equality of factor variances except for those explicitly described above was not significantly different from the final model in the *Equal Residual Variances* section (χ^2(5, $n = 471$) = 3.84, $p < .58$). The model which imposed equal factor covariances across racial/ethnic groups did not differ significantly from the model in the prior step with partially invariant factor variances (χ^2(12, $n = 471$) = 18.5, $p < .11$).

Final Measurement Model

The standardized loadings and implied reliability of the latent constructs for this final model are presented in Table 2. Recall that the final model showed metric invariance and all unstandardized loadings to be equal across racial/ethnic groups. However, there were several item (residual) and factor variances that did differ across racial/ethnic groups. This caused the standardized loadings of the items involved as well as the reliability of the factors involved to vary slightly by race/ethnicity because these variances are involved in the standardization. Those that differed because of this reason are shown in italics in Table 2. The loadings are mostly quite good as are the implied reliabilities of the factors. The indicators with marginal loadings in the family functioning latent constructs were retained to ensure that the hypothesized sub-domains of family functioning were represented. The composite reliabilities based on the loadings of the measurement model for externalizing varied from .71 to .72 across race/ethnicity for adolescent report and from .75 to .81 for parent report. The composite reliabilities for family functioning varied from .70 to .73 across race/ethnicity for adolescent report and from .74 to .78 for parent report.

Factor Mean Differences

The factor means implied by the final measurement model did differ significantly by race ethnicity ($\chi^2(8, n = 471) = 26.4, p < .001$). Table 3 shows the standardized mean differences on each of the latent factors by race/ethnic group. Adolescent report of externalizing was significantly lower in Black families than Hispanic families (*Standardized Difference* = -.42, $z = 3.03, p < .002$). Parent report of family functioning was significantly higher in Black families than in Hispanic families (*Standardized. Difference* = .56, $z = 4.18, p < .0001$). The only significant difference between White non-Hispanic and Hispanic families was adolescent report of family functioning which was significantly lower in White non-Hispanic families (*Standardized. Difference* = -.29, $z = 2.16, p < .04$). All four factor means differed between Black and White non-Hispanic families. Externalizing was significantly lower in Black families than White non-Hispanic families as measured by both adolescent report (*Standardized. Difference* = -.60, $z = 3.70, p < .0001$) and parent report (*Standardized. Difference* = -.35, $z = 2.57, p < .01$). Family functioning was significantly higher in Black families than in White non-Hispanic families as measured by both adolescent report (*Standardized. Difference* = .44, $z = 2.57, p < .01$) and by parent report (*Standardized. Difference* = .56, $z = 3.80, p < .0001$).

Discussion

The current study, building on the recommendations and guidance of Burlew et al. (2009), demonstrates procedures for how to systematically address measurement invariance in a multi-racial/ethnic sample of adolescent drug abusers and their families, It is the case in the BSFT trial data that there is sufficient equivalence of measures across racial/ethnic groups to make valid quantitative comparisons across these groups. The pattern of mean difference results observed in this study illustrate that ethnically diverse adolescents with drug use problems may enter treatment with very different levels of family functioning and problem behaviors.

One of the most interesting aspects of the current study was that, even though the variance-covariance matrices of our indicators were significantly different across racial/ethnic groups, the step-by-step decomposition of these group differences indicate that there is a surprising degree of invariance in the measurement of family functioning and adolescent externalizing across racial/ethnic adolescents and parents. Both configurally (i.e. the same indicators are important) and metrically (i.e. equal factor loadings), there is complete invariance across Black, Hispanic, and White non-Hispanic families. There was a statistically significant difference between the model that imposed metric invariance and the model that additionally imposed full scalar invariance, though the change to model fit parameters was quite modest. Together, these results imply that comparisons of factor means are possible and meaningful across the racial/ethnic groups with this measurement model, as long as the measurement model includes differential estimates of the items that were found not to be equivalent. In fact, it would be possible to analyze these baseline variables using a composite variables based on this model without great concern for the minor measurement differences across racial/ethnic groups. For example, composites based on equally weighted sums of standardized scales of each of the indicators within these constructs showed Cronbach's alpha in the range .89-.93 across race/ ethnicity.

Factor Mean Differences

Examination of factor means shows differences amongst the racial ethnic groups on the underlying family functioning and externalizing constructs. The comparison of factor means across groups showed that Black families reported significantly lower externalizing and significantly higher family functioning than did White non-Hispanic families regardless of reporter. Black adolescents also reported lower externalizing than Hispanic and White non-

Hispanic adolescents, and Black parents reported higher family functioning than Hispanic and White non-Hispanic parents. These differences were not small and in most cases are around what most call a moderate standardized difference of .50. This pattern of findings is consistent with prior research with drug using adolescents (Robbins et al., 2002), in which it was observed that Hispanic adolescents and their parents reported significantly higher rates of externalizing disorders than Black adolescents and their parents. One hypothesis generated (but not tested) by this pattern of differences is that there is a fundamental difference between racial/ethnic groups in which kids get referred for outpatient treatment. For example, in both the current study and the Robbins et al. (2002) study, the samples consisted of adolescents referred for outpatient treatment of substance use problems. It is possible that a group of Black substance users that have higher rates of problem behaviors are not referred to outpatient treatment, but rather are sent to a more intensive (residential) treatment setting or to detention/jail and this possibility should be examined in further research. Support for this argument about this subtle but pervasive type of institutionalized racism is based on the fact Blacks are disproportionately over-represented in jails and prisons (Radosh, 2008) and by examination of criminal justice decision-making (Leiber, Johnson, Fox & Lacks, 2007, Leiber & Johnson, 2008). This clearly has implications for any differences in other outcomes (such as drug abuse) across ethnicity in this sample and is information that will inform clinical comparisons of case response in process analyses. Generalizations to even the population of families with adolescents experiencing drug and problem behaviors, however, must be tempered by the fact that this is a clinical, non-random sample. Nevertheless, it is interesting that in this sample of adolescents and their families seeking drug treatment in community treatment settings that the Black adolescents appear to have fewer problem behaviors and the families have higher family functioning than do Hispanic and White non-Hispanic families.

Regardless of the reasons that may underlie between group differences, the differences in factor means should have no biasing effects for the analysis of trial outcomes in a study in which assignment to treatment condition is balanced across the racial/ethnic groups. If there were imbalance across conditions, then controlling for these mean differences across racial/ethnic groups might be very important. Given these differences, future clinical studies with multi-racial/ethnic drug using adolescents that include similar measures of family functioning and externalizing problems should include stratification procedures to ensure that treatment condition are balanced by race/ethnicity.

Factor Variance Differences

On three of the four factors, all but parent report of family functioning, the factor variances were lower for Black families than they were for Hispanic or White non-Hispanics, for whom the variances were found not to differ. This implies that not only are the Black families higher functioning with less externalizing, they are more homogeneous than are other families. This difference, as well, may be a result of differential probabilities of receiving drug treatment (given a specific level of problem behaviors). If Black adolescents with higher problem behaviors are more likely to be in more restrictive environments, then this would reduce the variability seen in this group.

Differences on Intercepts

Despite the remarkable measurement consistencies that were observed across groups in this sample, there were three indicators for which the intercepts appeared to vary across racial/ethnic groups and two additional indicators that were different in White non-Hispanic families than in Black and Hispanic families. These differences are potentially interesting culturally and clinically because they suggest different ways in which racial/ethnic groups may respond to the measures, which has important implications for (a) understanding the meaning of the

items and scales that are measured within each group, and (b) for examining differences between groups.

One example of these intercept differences was observed in parent reports of family cohesion. For example, in this sample, the parent reported cohesion level that corresponds to zero on mean family functioning is highest in Hispanic, followed by Black, and then White non-Hispanic. Thus, in this clinical sample, a Hispanic parent would need higher levels of cohesion to have an equivalent level of family functioning as a White non-Hispanic parent. These results suggest that Hispanic parents have a tendency to report higher levels of cohesion than can be explained by (higher) levels of the latent variable of family functioning (relative to the other racial/ethnic groups). The implication of this type of invariance is that there are different levels of the indicator scale associated with the zero level of the factor. Thus, any between group comparisons of the latent factor means should be made while controlling for the effects of these differences in intercepts. This intercept differences imply that there are different normative levels for this indicator scale across racial/ethnic groups in the current sample. Future research is needed to confirm this in other samples.

Differences in Uniqueness (item residual variance)

The residual variances of three of the indicators of parent report of family functioning. Positive Parenting, Parental Monitoring and Avoidance of Discipline were largest in Hispanic families. Of these, only the residual variance of parental monitoring was significantly different across all three racial/ethnic groups with White non-Hispanic families having the lowest variability. In addition, the residual variances of Delinquency and Conduct Disorder were significantly smaller in Black families than either Hispanic or White non-Hispanic families. One reason why these findings are relevant is that the larger the variance, the smaller the standardized loading and resulting reliability of the factor. Thus, this difference points out that the parent report of the three parenting related indicators are least reliable in Hispanic families, whereas the two adolescent reports of problem behaviors are most reliable in Black adolescents.

Limitations

This manuscript has provided an example of how larger drug abuse treatment trials might approach making racial/ethnic comparisons of their sample composition with measurement integrity. As with most clinic research samples, a limitation of this research is that there is no guarantee the sample is representative of the population of families in each of these ethnic groups. In fact, we have noted both in this sample and others in our program of research, that Black youth with more severe symptoms may be systematically less likely to be given the opportunity to participate in research trials. Second, unfortunately because of the relatively small size with respect to the requirements of measurement modeling, it was not possible to stratify the sample into an exploratory sample with a confirmatory target sample. This more extensive approach would allow stronger statements about the statistical generalizability of the specific types of invariance observed. Second, this study took the evidence of adequate Cronbach's α on the indicator scales by racial/ethnic groups as a starting point. The very approach to invariance taken herein could be taken with each individual scale, if sample size would have allowed it. Thus, we cannot make statements about potential differential item functioning within each of the scales used as factor indicators. Despite these limitations, this is a very large sample of treatment seeking adolescents and their families and there is no reason to believe that the sample herein is unusual with respect to the population of drug treatment seeking adolescents and their families.

Conclusion

This very careful analysis has shown racial/ethnic differences in the presenting levels of problem behaviors and family functioning in a large cohort of community-based adolescent

drug treatment clients. This difference is worthy of future study and verification. Similarly, there were minor differences in the measurement structure across ethnic/racial groups that might lead to better understanding of cultural differences if confirmed in other cohorts and by other researchers. Because of the relatively small samples of most drug treatment trials, it is unlikely that definitive answers to these cultural measurement differences can be provided by a single trial. However, if the field is conscientious about planning and reporting these studies, evidence can be cumulated over time concerning the generalizability of these differences.

Acknowledgments

References

Achenbach TM, Becker A, Dopfner M, Heiervang E, Roessner V, Steinhausen H-C, Rthenberger A. Multicultural assessment of child and adolescent psychopathology with ASEBA and SDC instruments: research findings, applications, and future directions. The journal of Child Psychology and Psychiatry 49:251–275.

Achenbach TM, McConaughy SH, Howell CT. Child adolescent behavioral and emotional problems: Implications of cross informant correlations for situational specificity. Psychological Bulletin 1987;101:213–232. [PubMed: 3562706]

Achenbach, TM.; Rescorla, LA. Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families; 2001.

Burlew AK, Feaster D, Brecht M, Hubbard R. Measurement and data analysis in research addressing health disparities in substance abuse. Journal of Substance Abuse Treatment 2009;36:25–43. [PubMed: 18550320]

Campbell CI, Weisner C, Sterling S. Adolescents entering chemical dependency treatment in private managed care: ethnic differences in treatment initiation and retention. Journal of Adolescent Health 2006;38:343–350. [PubMed: 16549294]

Chen FF, Sousa KH, West SG. Testing measurement invariance of second-order factor models. Structural Equations Modeling 2005;12:471–492.

De Los Reyes A, Kazdin AE. Measuring informant discrepancies in clinical child research. Psychological Assessment 2004;16:330–334. [PubMed: 15456389]

Elliot, DS.; Ageton, SS.; Huizinga, D.; Knowles, BA.; Cantor, RJ. The National Youth Survey Rep. No. 26. Boulder, CO: Behavioral Research Institute; 1983. The prevalence and incidence of delinquent behavior: 1976-1980.

Feaster DJ, Robbins MS, Horigian V, Szapocznik J. Statistical issues in multisite effectiveness trials: the case of brief strategic family therapy for adolescent drug abuse treatment. Clinical Trials 2004;1(5): 428–439. [PubMed: 16279281]

Gil AG, Wagner EF, Vega WA. Acculturation, familism, and alcohol use among Latino adolescent males: Longitudinal relations. Journal of Community Psychology 2000;28:443–458.

Gorman-Smith D, Tolan PH, Zelli A, Huesmann LR. The relation of family functioning to violence among inner-city minority youths. Journal of Family Psychology 1996;10:115–129.

Gross D, Fogg L, Young M, Ridge A, Cowell JM, Richardson R, Sivan A. The equivalence of the child behavior checklist/1 1/2-5 across parent race/ethnicity, income level, and language. Psychology of Addictive Behaviors 2006;18:313–323.

Hall GC, Maramba GG. In search of cultural diversity: recent literature in cross-cultural and ethnic minority psychology. Cultural Diversity & Ethnic Minority Psychology 2001;7(1):12–26. [PubMed: 11244901]

Hawkins E, Cummins L, Marlatt G. Preventing Substance Abuse in American Indian and Alaska Native Youth: Promising Strategies for Healthier Communities. Psychological Bulletin 2004;130:304–323. [PubMed: 14979774]

Huizinga, D.; Elliott, DS. National Youth Survey Project Report 27. Behavioral Research Institute; Boulder, Colo: 1983. A preliminary examination of the reliability and validity of the national youth survey self-reported delinquency indices.

Huizinga D, Elliott DS. Reassessing the reliability and validity of self-report delinquency measures. Journal of Quantitative Criminology 1986;2:293–327.

Imada T, Schiavo RS. The use of ethnic minority populations in published psychological research, 1990-1999. Journal of Psychology: Interdisciplinary and Applied 2005;139:389–400. [PubMed: 16285210]

Jackson-Gilfort A, Liddle HA, Tejeda MJ, Dakor GA. Facilitating engagement of African American male adolescents in family therapy: A cultural theme process study. Journal of Black Psychology 2001;27:321–340.

Kim B, Zane, Hong S. Protective factors against substance use among Asian American youth: a test of peer cluster theory. Journal of Community Psychology 2002;30:565–584.

Kline, RB. Principles and Practice of Structural Equation Modeling. 2. New York: Guilford Press; 2002.

Leung PW, Lucas CP, Hung SF, Kwong SL, Tang CP, Lee CC, et al. The test-retest reliability and screening efficiency of DISC predictive scales-version 4.32 with Chinese children/youths. European Child & Adolescent Psychiatry 2005;14(8):461–465. [PubMed: 16341503]

Lieber MJ, Johnson JD. Being Young and Black: What are their effects on juvenile justice decision making? Crime & Delinquency 2008;54:560.

Lieber MJ, Johnson J, Fox K, Lacks R. Differentiating among racial/ethnic groups and its implications for understanding juvenile justice decision making. Journal of Criminal Justice 2007;35:471–484.

McDonald RP, Ho M. Principles and practice in reporting structural equation analyses. Psychological Methods 2002;7:64–82. [PubMed: 11928891]

Meredith W. Measurement invariance, factor analysis, analysis and factorial invariance. Psichometrika 1993;58(4):525–543.

Meyer K, Hagan TA, McDermott P, Webb A, Randall M, Frantz J. Factor structure of the Comprehensive Adolescent Severity Inventory (CASI): Results of reliability, validity, and generalizability analysis. The American Journal of Drug and Alcohol Abuse 2006;32:287–310. [PubMed: 16864465]

Miranda J. Introduction to the Special Section of Recruiting and Retaining Low Income Latinos in Psychotherapy Research. Journal of Consulting and Clinical Psychology 1996;64:848–850. [PubMed: 8916609]

Moos, R.; Moos, B. Family Environment Scale Manual. 2. Palo Alto, Calif: Consulting Psychologists Press; 1986.

Munot-Vilaro F, Egan M. Reliability Issues of the Family Environment Scale for Cross-Cultural Research. Nursing Research 1990;39:244–247. [PubMed: 2367207]

National Institutes of Health. NIH guidelines for the inclusion of women and minorities as subjects in clinical research. 2001. http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm

Perrino T, Coatsworth J, Briones E, Pantin H, Szapocznik J. Initial engagement in parent-centered interventions: A family systems perspective. Journal of Primary Prevention 2001;22:21–44.

Radosh PF. War on drugs: gender and race inequities in crime control strategies. Criminal Justice Studies: A Critical Journal of Crime. Law and Society 2008;21:167–178.

Reuterberg S, Gustafsson J. Confirmatory Factor Analysis and Reliability: Testing Measurement Model Assumptions. Educational & Psychological Measurement 1992;52:795–810.

Robbins, MS.; Horigian, V.; Feaster, DJ.; Szapocznik, J.; Burlew, AK.; Hodgkins, C.; Schindler, E.; Rohrbaugh, M.; Shoham, V. Brief Strategic Family Therapy versus Treatment as Usual: Engagement and Retention in a Community-Based Sample of Drug Using Adolescents. University of Miami; 2009. Unpublished Manuscript

Robbins MS, Kumar S, Walker-Barnes C, Feaster D, Briones E, Szapocznik J. Ethnic differences in psychiatric comorbidity among substance-abusing adolescents referred to outpatient therapy. Journal of the American Academy of Child & Adolescent Psychiatry 2002;41(4):394–401. [PubMed: 11931595]

Robbins MS, Szapocznik J, Horigian VE, Feaster DJ, Puccinelli M, Jacobs P, Burlew K, Werstlein R, Bachrach K, Brigham G. Brief strategic family therapy™ for adolescent drug abusers: A multi-site effectiveness study. Contemporary Clinical Trials 2009;30(3):269–278. [PubMed: 19470315]

Santisteban DA, Szapocznik J, Perez-Vidal A, Kurtines WM, Murray EJ, Laperriere A. Efficacy of intervention for engaging youth and families into treatment and some variables that may contribute to differential effectiveness. Journal of Family Psychology 1996;10(1):35–44.

Satorra, A. Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In: Heijmans, RDH.; Pollock, DSG.; Satorra, A., editors. Innovations in multivariate statistical analysis. A Festschrift for Heinz Neudecker. London: Kluwer Academic Publishers; 2000. p. 233-247.

Shillington A, Clapp J. Adolescents in Public Substance Treatment Programs: The Impact of Sex and Race on Referrals and Outcomes. Journal of Child and Adolescent Substance Abuse 2003;12:69–91.

Smith C, Krohn MD. Delinquency and family life among male adolescents: The role of ethnicity. Journal of Youth and Adolescence 1995;24:69–93.

Strada MJ, Donohue B. Substance Use in Ethnic Minority Youth. Journal of Ethnicity in Substance Abuse 2006;5:67–89. [PubMed: 16537338]

Strada M, Donohue B, Lefforge N. Examination of Ethnicity in Controlled Treatment Outcome Studies involving Adolescent Substance Abusers: A Comprehensive Literature Review. Psychology of Addictive Behaviors 2006;20:11–27. [PubMed: 16536661]

Szapocznik J, Prado G, Burlew AK, Williams R, Santisteban D. Drug Abuse in African American and Hispanic Adolescents: Culture, Development, and Behavior. Annual Review of Clinical Psychology 2007;3:77–105.

Szapocznik J, Perez-Vidal A, Brickman A, Foote FH, Santisteban DA, Hervis O, Kurtines WM. Engaging adolescent drug abusers and their families into treatment: A strategic structural systems approach. Journal of Consulting and Clinical Psychology 1988;56(4):552–557. [PubMed: 3198813]

Tolan PH, Gorman-Smith D, Huesmann LR, Zelli A. Assessment of family relationship characteristics: A measure to explain risk for antisocial behavior and depression among urban youth. Psychological Assessment 1997;3:212–223.

Thornberry, T.; Huizinga, D.; Loeber, R. The prevention of serious delinquency and violence: Implications from the Program of Research on the Causes and Correlates of Delinquency. In: Howell, J.; Krisberg, B.; Hawkins, J.; Wilson, J., editors. Sourcebook on Serious, Violent, and Chronic Juvenile Offenders. Thousand Oaks, CA: Sage; 1995. p. 213-237.

U.S. Bureau of the Census. Projections of the resident population by race, Hispanic origin, and nativity: Middle series, 1999–2100. 2000. (online). www.census.gov/population/www/projections/popproj.html (October 2001)

U.S. Census Bureau. U.S Interim Projections by Age, Sex, Race and Hispanic Origin. 2004. http://www.census.gov/ipc/www/usinterimprojinternet release date, March 18, 2004

Vandenberg R, Lance C. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods 2000;3:4–70.

Woehr, DJ.; Arciniega, LM.; Fowler, O. Measuring Work Ethic: An Examination of the Measurement Equivalences and Spanish Versions of the Multidimensional Work Ethic Profile. Proceedings of the 2003 European Applied Business Research Conference; Venice, Italy. June 2003; 2003.

Youngstrom E, Loeber R, Stouthamer-Loeber M. Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. Journal of Consulting and Clinical Psychology 2000;68:1038–105. [PubMed: 11142538]
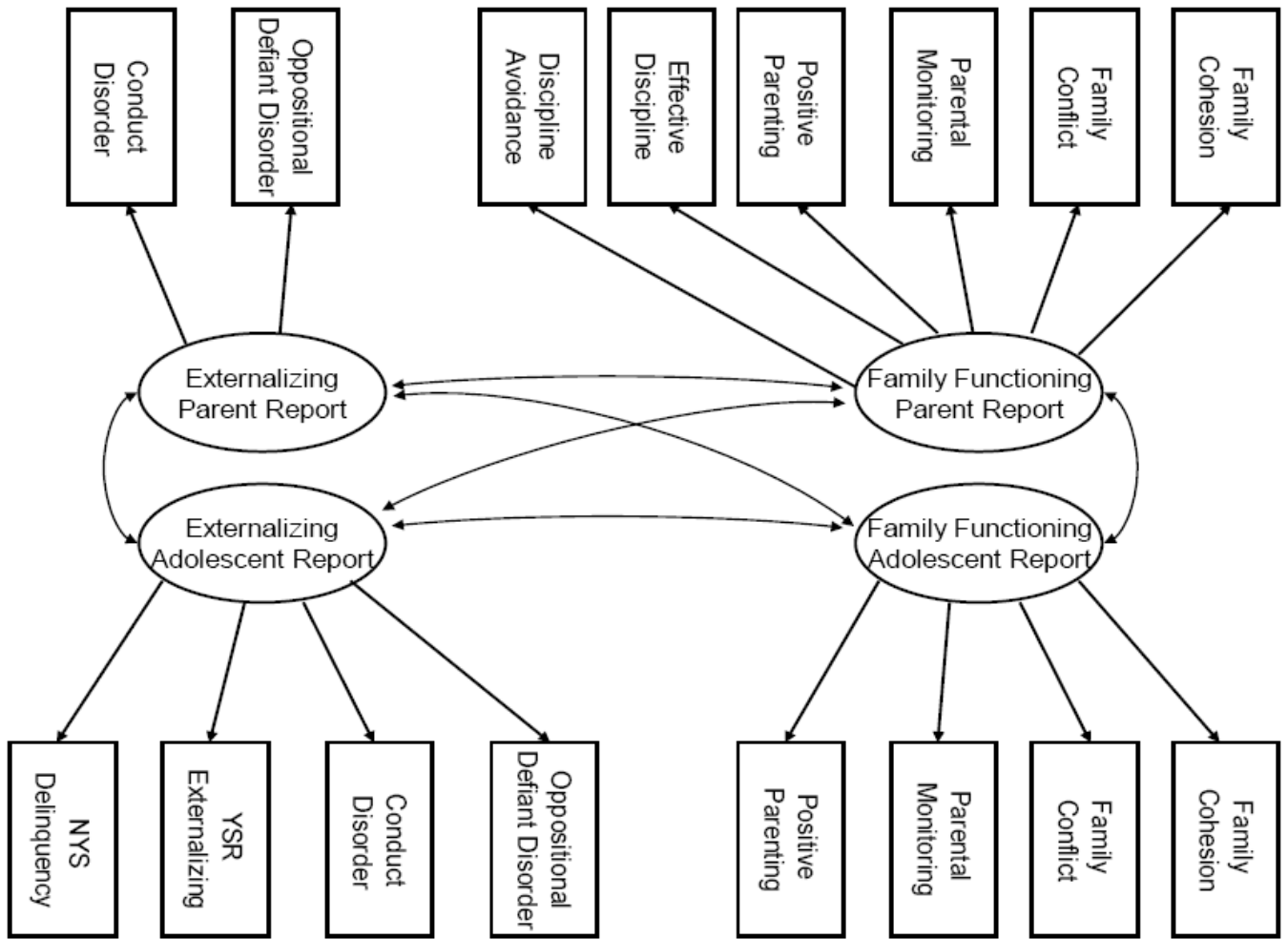
Figure 1.
Hypothesized Measurement Model

Table 1

ANOVA Results on Individual Indicator Scales (Means)

	Hispanic	Black	White[1]	Pooled Variance	p-value
Externalizing					
Adolescent Report					
NYS Delinquency	3.95	2.65	3.98	12.88	0.004[b,c]
YSR Externalizing	18.11	15.15	20.06	86.71	0[b,c]
DISC-Conduct Disorder	1.73	1.18	1.49	2.6	0.015[c]
DISC Oppositional Defiant Disorder	2.93	2.47	3.51	3.41	0[a,b,c]
Parent Report					
DISC-Conduct Disorder	3.08	3.07	3.81	8.38	0.04[a,b]
DISC Oppositional Defiant Disorder	5.48	5.12	5.97	7.6	0.023[b]
Family Functioning					
Adolescent Report					
Positive Parenting	12.12	12.84	12.54	10.34	0.615
Parental Monitoring	33.09	33.44	34.21	44.87	0.308
Family Conflict	3.64	3.7	4.18	5.24	0.076
Family Cohesion	6.38	6.83	5.8	4.61	0.001[a,b]
Parent Report					
Positive Parenting	14.4	14.72	14.18	5.3	0.18
Parental Monitoring	38.82	40.63	39.88	20.66	0.002[a,c]
Commitment to Discipline	17.4	18.69	17.56	8.43	0.001[b,c]
Discipline Avoidance	10.92	11.97	10.93	6.95	0.001[b,c]
Family Conflict	3.46	3.34	3.98	5.27	0.049[a,b]
Family Cohesion	6.29	6.55	5.51	5.39	0.001[a,b]
Sample Size	213	110	148		

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

[1]Non-Hispanic,

[a]Hispanics and White non-Hispanics are significantly different from each other,

[b]Blacks and White non-Hispanics are significantly different from each other,

[c]Blacks and Hispanics are significantly different from each other.

Table 2

Standardized Factor Loadings by Race/Ethnicity

Externalizing			
Adolescent Report	**Hispanic**	**Black[1]**	**White [1]**
NYS Delinquency	0.74	*0.85*	0.74
YSR Externalizing	0.79	*0.77*	0.79
DISC-Conduct Disorder	0.72	*0.78*	0.72
DISC Oppositional Defiant Disorder	0.72	*0.71*	0.72
Implied Reliability	.79	*.80*	.79
Parent Report			
DISC-Conduct Disorder	0.73	*0.79*	0.73
DISC Oppositional Defiant Disorder	0.81	*0.85*	0.81
Implied Reliability	.75	*.81*	.75
Family Functioning			
Adolescent Report			
Positive Parenting	0.49	*0.46*	0.49
Parental Monitoring	0.56	*0.53*	0.56
Family Conflict	-0.71	*-0.68*	-0.71
Family Cohesion	0.76	*0.73*	0.76
Implied Reliability	.73	*.70*	.73
Parent Report			
Positive Parenting	*0.31*	0.36	0.36
Parental Monitoring	*0.50*	*0.54*	*0.63*
Commitment to Discipline	*0.69*	0.78	0.78
Discipline Avoidance	-0.82	-0.82	-0.82
Family Conflict	-0.44	-0.44	-0.44
Family Cohesion	0.58	0.58	0.58
Implied Reliability	*.74*	*.77*	*.78*
Sample Size	213	110	148

Note. The final model incorporated full metric invariance--all unstandardized loadings are equal across racial/ethnic groups. There are a few residual variances as well as three factor variances for Black Families that were not invariant and this causes the standardized loading to differ slightly. Loadings in italics are the loadings that vary slightly due to lack of invariance in the item and factor variances.

[1] Non-Hispanic.

Table 3

Standardized Mean Differences in Latent Factors by Race/Ethnicity

	Black[1]-Hispanic	White[1]-Hispanic	Black-White[1]
Externalizing, Adolescent Report	**-0.42** ***	0.17	**-0.60** ****
Externalizing, Parent Report	-0.15	*0.24[a]*	**-0.35** **
Family Functioning, Adolescent Report	0.12	**-0.29** *	**0.44** **
Family Functioning, Parent Report	**0.56** ***	-0.004	**0.56** ***

[1]Non-Hispanic,

[a].05 < p ≤ .10,