Published in final edited form as:

J Psychopathol Behav Assess. 2017 March ; 39(1): 46–57.

Child Conduct Problems across Home and School Contexts: A Person-Centered Approach

Michael J. Sulik, Clancy Blair, and Mark Greenberg

Michael J. Sulik, Institute of Human Development and Social Change, New York University; Clancy Blair, Department of Applied Psychology, New York University; Mark Greenberg, Human Development and Family Studies, The Pennsylvania State University. The Family Life Project Phase II Key investigators include: Lynne Vernon-Feagans, The University of North Carolina at Chapel Hill; Mark T. Greenberg, The Pennsylvania State University; Clancy B. Blair, New York University; Margaret R. Burchinal, The University of North Carolina at Chapel Hill; Martha Cox, The University of North Carolina at Chapel Hill; Patricia T. Garrett-Peters, The University of North Carolina at Chapel Hill; Jennifer L. Frank, The Pennsylvania State University; W. Roger Mills-Koonce, University of North Carolina-Greensboro; and Michael T. Willoughby, RTI International

Abstract

To examine patterns of conduct problems across the home and school context, we used latent class analysis to analyze primary caregivers' and teachers' ratings on the conduct problems subscale of the Strengths and Difficulties Questionnaire (R. Goodman, 1997) in the Family Life Project ($N=1,292$), a prospective study of child development in rural and small town contexts. We found a similar four-class solution at five and seven years of age. In decreasing prevalence, the following classes were identified: (1) low symptoms reported by both informants (low cross-context); (2) high parent-reported symptoms, low teacher-reported symptoms (home context); (3) low parent-reported symptoms, moderate teacher-reported symptoms (school context); and (4) high symptoms reported by both informants (high cross-context). Classes exhibited stability from age five to age seven: children were more likely to remain in the same class than to transition to a different class, and longitudinal stability was especially high for children in the low cross-context class at age 5. A number of child and family characteristics measured in early childhood (executive function, verbal ability, poverty-related risk, sensitive parenting, and parental depressive symptoms) were associated with class membership at age five and age seven, but were generally not associated with longitudinal transitions between classes.

Correspondence concerning this article should be addressed to Michael J. Sulik, 196, Mercer St., Rm. 810, New York, NY 10012. michael.sulik@nyu.edu.

Contact Information for Other Authors. Clancy Blair, Address: 196 Mercer St., Rm. 811, New York, NY 10012, cbb5@nyu.edu; Mark Greenberg, Address: 306 BBH Building, The Pennsylvania State University, University Park, PA 16802, mxg47@psu.edu

Michael J. Sulik has no conflicts of interest to declare. Clancy Blair has no conflicts of interest to declare. Mark Greenberg has no conflicts of interest to declare.

Compliance with Ethical Standards: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Keywords

Conduct Problems; Externalizing Symptoms; Strengths and Difficulties Questionnaire; Longitudinal; Latent Class Analysis

Cross-Informant Agreement

Psychopathology is generally viewed as a stable individual characteristic. Nonetheless, correlations between parents' and teachers' ratings of children's externalizing symptomatology tend to be small-to-moderate in size, averaging approximately .28 (De Los Reyes et al., 2015; see also Achenbach, McConaughy, & Howell, 1987). This relatively low degree of cross-informant agreement means that many children will be rated as high in symptomatology by a parent and low by a teacher or low by a parent and high by a teacher. In what has been labeled *divergent operations*, discrepancies between parents' and teachers' ratings are believed to be influenced by meaningful differences in children's behavior across the home and school context (De Los Reyes, Thomas, Goodman, & Kunder, 2013). There are reasons to believe that children who are rated as high in symptoms by both parents and teachers could be especially at risk for negative outcomes (Campbell, 1994). These children exhibit difficulties that generalize across multiple contexts, which could indicate more severe or pervasive problems (Kraemer et al., 2003). Consistent with this interpretation, a multi-informant approach that incorporates parents' and teachers' ratings of child symptomatology is able to achieve better sensitivity and specificity in predicting clinical diagnosis than either informant alone (R. Goodman, Ford, Simmons, Gatward, & Meltzer, 2000; R. Goodman, Renfrew, & Mullick, 2000). In contrast, individuals who are rated as high in symptoms by a single informant (either by a teacher or a parent, but not both) could exhibit symptoms that are specific to a single context or that reflect bias related to child or informant characteristics (De Los Reyes & Kazdin, 2005). It is therefore of interest to investigate whether the risk factors for consistently high symptomatology across contexts (home and school)—as indicated by elevated ratings by both a parent and teacher—are different from the risk factors for symptomatology that is specific to a single context (either home or school, but not both).

Latent Class Analysis: A Person-Centered Approach

In this study, we use LCA (Collins & Lanza, 2010), which has been described as a person-centered method. In contrast to variable-centered methods (such as correlation and factor analysis), person-centered methods such as LCA are used to identify groups of individuals who are similar to each other, but distinct from individuals in other groups. Class membership is a latent variable and, similar to factor analysis, it is used to explain the relations among the set of measured variables. A particular strength of person-centered methods such as LCA is that they can address questions about how multiple traits co-occur within individuals. We capitalize on this strength to distinguish among different patterns of primary caregivers' and teachers' reports of child conduct problems.

We chose to use LCA because it has three important advantages relative to other methods for creating groups. One such method that has frequently been used in previous research is to

create high and low symptom groups for individuals who are respectively above or below a clinical-cut point (e.g., Lewinsohn, Seeley, Roberts, & Allen, 1997). Creating an artificial dichotomy from a continuous scale has been shown to *both* reduce power and generate spurious associations (MacCallum, Zhang, Preacher, & Rucker, 2002); LCA does not suffer from these limitations. In addition, LCA provides a number of ways of evaluating model adequacy that have been shown to perform well in simulation studies (Nylund, Asparouhov, & Muthén, 2007). Consequently, choosing the number of classes for the model can be empirically informed. Finally, because LCA is a latent-variable model, measures of association are disattenuated for measurement error, which provides more accurate parameter estimates (Bollen, 1989).

LCA has previously been used to investigate children's conduct problems in a study of the Dunedin, New Zealand, 1977 birth cohort (Fergusson, Boden, & Horwood, 2009). In that study, the authors dichotomized parents' and teachers' reports of child conduct problems (using a 90th percentile cut-point) when children were 7, 8, and 9 years old. They found that a four-class model fit better than models with fewer classes: the four classes identified were labeled as “no problems,” “mother-reported problems,” “teacher-reported problems” and “generalized problems.” Latent class analysis has also been used to investigate a mixed clinical and community sample of children's disruptive behavior during a structured laboratory observation with a parent and examiner (De Los Reyes, Henry, Tolan, & Wakschlag, 2009). In that study, the authors used a 75% cut-point to dichotomize the scales and found a similar four-class solution fit best. The four classes were labeled “not disruptive,” “disruptive with parent,” “disruptive with examiner,” and “pervasively disruptive.” These studies provide evidence that latent class analysis can be used to identify different patterns of conduct problems across informants and contexts, respectively.

Latent Transition Analysis

For longitudinal analyses, we use LTA, an extension of LCA to multiple assessments. In an LTA model, LCA models are estimated at each assessment and are connected using a latent Markov chain model (Collins & Wugalter, 1992). Like the LCA model, the LTA model parameters estimates include the conditional response probabilities (i.e., the probability of endorsing a response, given membership in a particular class) and latent class prevalences. These parameters can be freely estimated at each assessment or constrained to be equal across assessments (longitudinal measurement invariance). LTA models generate an additional set of parameters that are not estimated in an LCA model: latent transition probabilities (Collins & Lanza, 2010). The latent transition probabilities represent the probability of membership in a class at a later assessment, contingent on membership in a particular class at the previous assessment.

Risk for Early Conduct Problems

To examine the criterion validity of the latent class solution we will test whether class membership can be predicted from a number of child and family variables that were measured in early childhood in the Family Life Project. These variables were selected on the basis of their robust relations with children's conduct problems in previous research. In

testing these relations, our goal is to identify whether the predictors of high symptoms across the home and school contexts are similar to the predictors for high symptoms in either context alone.

We will test three aspects of the family environment as predictors of conduct problems: a cumulative index of poverty-related risk, sensitive parenting, and parental depression. Cumulative risk refers to the aggregation of a set of risk variables with the expectation that it is not any single risk factor that is important, but rather the simultaneous presence of multiple risk factors that is associated with negative outcomes (Sameroff, Seifer, Baldwin, & Baldwin, 1993). Although the conceptualization of cumulative risk differs from study to study, indices of poverty-related risk similar to the one used in this study have been positively associated with children's conduct problems in previous research (Deater-Deckard, Dodge, Bates, & Pettit, 1998; Trentacosta et al., 2008). We also consider sensitive parenting, which is negatively associated with conduct problems (Belsky, Fearon, & Bell, 2007; Campbell, 1995; Eisenberg et al., 2005). Finally, we examine parental depression, which has been identified as non-specific predictor of diverse aspects of child psychopathology and is positively associated with conduct problems (S. H. Goodman et al., 2011).

The child characteristics that we will test include executive function and verbal ability, which are associated with one another (Hughes, 1998) and with aspects of the childhood environment such as sensitive parenting (Hammond, Müller, Carpendale, Bibok, & Liebermann-Finestone, 2012; Landry, Miller-Loncar, Smith, & Swank, 2002). Previous studies comparing children with disruptive behavior disorders to comparison subjects have documented deficits in both executive function and verbal ability among clinical groups (Loney, Frick, Ellis, & McCoy, 1998; Moffitt, 1993; Speltz, DeKlyen, Calderon, Greenberg, & Fisher, 1999) and negative associations have also been reported in non-clinical samples (Ackerman, Brown, & Izard, 2003; Nigg, Quamma, Greenberg, & Kusche, 1999).

The Present Investigation

Using data from the Family Life Project, a prospective longitudinal study of children and families in rural and small town contexts in the United States, we investigate the *patterns* of child conduct problems across the home and school contexts. Children's symptoms were measured at approximately five and seven years of age. We first use LCA to analyze the cross-sectional data at each of these assessments and then estimate an LTA model to examine continuity in class membership across the two-year study period.

We hypothesized that a four-class LCA model would be able to accurately identify common and easily interpretable patterns of parent and teacher's ratings of child conduct problems that were also found in previous research using similar measurement strategies (De Los Reyes et al., 2009; Fergusson et al., 2009). Here we describe the four classes that we expected to find, in order of decreasing prevalence: (1) *low cross-context*: low symptom ratings for both primary caregivers and teachers; (2) *home context*: high parent-rated symptoms and low teacher-rated symptoms; (3) *school context*: high teacher-rated symptoms and low parent-rated symptoms; and (4) *high cross-context*: high symptoms in both contexts.

In community samples, the distribution of parent and teacher ratings of symptom scores tends to be positively skewed, with many children scoring at or close to the scale minimum. For this reason, the low symptom class was expected to have the largest prevalence, whereas the co-occurring symptom class was expected to have the smallest prevalence. Because parents have been found to endorse a greater number of symptoms relative to teachers (e.g., Miner & Clarke-Stewart, 2008), the home context class was expected to be more prevalent than the school context class. We also expected to observe substantial continuity in class membership over time (e.g., individuals in a class at the age-five assessment would be likely to remain in a similar class at the age-seven assessment).

Finally, we tested whether a broad set of child and family characteristics as predictors of class membership at both assessments. For analyses predicting the transition between classes from the first to the second assessment, we expected any effects to be relatively modest because controlling for prior symptoms greatly reduces the variance that can be explained by other predictors.

Method

Participants

The Family Life Project is a study of 1,292 children who were recruited at birth in 2003–2004 from six rural counties in Pennsylvania and North Carolina. Low-income families were oversampled in both states and African American families were oversampled in North Carolina. Extensive information about recruitment and sample characteristics has been published elsewhere (Vernon-Feagans, Cox, & The Family Life Project Key Investigators, 2013). Children and parents were assessed at home at approximately 2, 6, 15, 24, 36, 48, and 60, and 90 months of age, and teachers' completed questionnaires annually for four years starting when children were in preschool. Nearly all (99.6%) of the primary caregivers at the 2-month assessment were biological mothers, for whom the average age at the child's birth was 25.58 years ($SD = 5.78$ years). At the 2-month assessment, the average years of education completed by primary caregivers was 14.42 ($SD = 2.82$) and 65.7% of primary caregivers reported that their spouse or partner lived in the household. At the 6-month assessment, families reported on their household income and household size. Using this information, we computed the household income-to-needs ratio, which expresses household income relative to the federal poverty guidelines for a family of the same size. The average household income-to-needs ratio was 1.92 ($SD = 1.70$). One third of households reported a family income below the federal poverty line (\$18,850 for a family of four in 2004; U.S. Department of Health and Human Services, 2014)

Measures

Child demographic covariates—At the 2-month assessment, primary caregivers reported on the gender (0 = Male; 1 = Female) and race (0 = not African American; 1 = African American) of their child.

Child conduct problems—The age 4–10 version of the Strengths and Difficulties Questionnaire (SDQ) is a brief, widely-used screening questionnaire that includes a five-

item scale assessing children's conduct problems (R. Goodman, 1997). This scale is broadly representative of children's externalizing behavior problems. Each item (1: *often loses temper*; 2: *generally well behaved, usually does what adults request* (reversed); 3: *often fights with other children or bullies them*; 4: *often lies or cheats*; 5: *can be spiteful to others*) is rated on a three-point scale (0 = “Not true”; 1 = “Somewhat true”; 2 = “Certainly true”). Longitudinal measurement invariance for the SDQ conduct problems scale has previously been demonstrated in this sample for primary caregivers' and teachers' reports (results available upon request).

Primary caregivers completed the SDQ at the 60-, and 90-month home visits, whereas teachers completed the SDQ annually in pre-kindergarten, kindergarten, grade 1, and grade 2. Because the assessment schedule differed for parents and teachers, we matched parent data from the 60- and 90-month home assessments (henceforth, T1 and T2) to the closest teacher assessment that was within one year of the parent assessment (or, if parent data were missing, within one year of the average age at the parent assessment). Primary caregiver SDQ data were available for 1,086 children at T1 (84%) and for 1,088 children at T2 (84%), and teacher SDQ data were available for 960 children at T1 (74%) and 1,023 children at T2 (79%).

To evaluate our ability to match these visits on age, we computed the absolute value of the difference in age between the parent and teacher questionnaires. The average difference was 0.33 years ($SD = 0.21$) at T1 and 0.33 years ($SD = 0.19$) at T2. This indicates that we were able to closely match the primary caregiver and teacher assessments on age. Mean age for primary caregivers' reports at each assessment was 5.05 ($SD = 0.27$) and 7.27 ($SD = 0.35$) and the corresponding values for teachers' reports were 5.16 ($SD = 0.43$) and 7.13 ($SD = 0.50$).

We computed ordinal alpha reliability for the conduct problems scale because Cronbach's alpha underestimates reliability for ordinal scales, especially for scales with few response options (Zumbo, Gaderman, & Zeisser, 2007). For primary caregivers, T1 $\alpha = .67$ and T2 $\alpha = .73$; for teachers, T1 $\alpha = .79$ and T2 $\alpha = .80$.

Child executive function—Children's executive function was assessed using a battery of five tasks at the 36-month assessment and six tasks at the 48- and 60-month assessments. At the 48- and 60-month assessments, the executive function battery included three inhibitory control tasks, two working memory tasks, and one attention shifting task. The inhibitory control tasks included a Simon-like spatial conflict task, a Stroop-like silly sounds task, and a farm animal go no-go task. The working memory tasks included a span-like task and a self-ordered pointing task. The attention shifting task was an item selection task modeled on the Dimensional Change Card Sort task (Willoughby, Wirth, & Blair, 2011). At the 36-month assessment, a similar (but not identical) spatial conflict task was used and the self-ordered pointing task was not administered.

Each task was presented in an open spiral bound flipbook with pages that measured 8'' $\times$ 14''. For each task, research assistants established that the child knew colors and numbers and administered training trials and up to three practice trials if needed. If children failed to

demonstrate an understanding of the goals of the task during the practice trials, the examiner discontinued that task. Full details regarding the administration rules, psychometric properties, and scoring for each of these tasks have been presented elsewhere (Willoughby et al., 2011).

Item response theory was used to generate *expected a posteriori* scores for each task and these scores were equated over time using a calibration sample. For details about the generation of these scores and information about their longitudinal measurement invariance, see Willoughby, Wirth, Blair, & The Family Life Project Investigators (2012). Within each assessment, the scores were averaged to form a composite measure of executive function ability, α s = .37 at 36 months, .55 at 48 months, and .50 at 60 months. Although these reliability numbers are relatively low, the EF variables at each assessment have shown robust associations with other study variables (e.g., Sulik et al., 2015). We standardized the EF composites within each assessment ($M = 0$, $SD = 1$) and averaged these scores over time to create a single measure of executive function, standardized $\alpha = .68$.

Child verbal intelligence—At the preschool assessment, data collectors administered the Peabody Picture Vocabulary Test, 4th edition (Dunn & Dunn, 2007) as a measure of children's verbal intelligence. Scores are adjusted for age and are standardized relative to a population M of 100 and population SD of 15.

Primary caregiver depressive symptoms—Primary caregivers completed the 20-item Center for Epidemiological Studies Depression scale (CES-D; Radloff, 1977) at the 24-, 36-, and 48-month assessments, α s = .89, .90, and .90. Items for this scale (e.g., “I felt depressed”) are rated on a scale ranging from 0 = *Rarely or none of the time (less than 1 day)* to 3 = *Most or all of the time (5-7 days)*. Correlations across time ranged from .61 to .64. Consequently, we averaged across all three assessments to create a measure of early primary caregiver depressive symptoms, $\alpha = .83$. The proportion of primary caregivers reporting a CES-D score greater than or equal to the frequently-used clinical cut-off score of 16 (Lewinsohn et al., 1997) was .24 at 24 months, .29 at 36 months, and .25 at 48 months.

Parenting—Children and their primary caregiver completed a 10 minute semi-structured free-play task at the 6- and 15-month assessments and a 10 minute puzzle task at the 24- and 36-months assessments. The free-play interaction involved asking each caregiver to use a standardized set of toys and to play with the infant as he or she normally would. The puzzle task presented the child with a series of three increasingly difficult jig-saw puzzles to complete and asking the caregiver to assist the child in any way that he or she chose.

These parent-child interactions were video recorded, and the following seven aspects of parenting were later coded: sensitivity, detachment, intrusiveness, stimulation, positive regard, negative regard, and animation in interacting with the child (Cox & Crnic, 2002; NICHD Early Child Care Research Network, 1999). Ratings were made on a scale ranging from 1 (“not at all characteristic”) to 5 (“highly characteristic”). Each scale was coded by a team of 4–5 coders, which included 1–2 master coders. Each coder was trained to be reliable with the master coder(s) and each coder completed approximately 30% of the tapes with the master coder(s). For each of the parenting scales, the intraclass correlation exceeded .80 at

each assessment. Factor analysis of the parenting variables yielded two dimensions of parenting behavior (Mills-Koonce et al., 2011; Vernon-Feagans et al., 2013). Based on those results, we created a composite measure corresponding to the first factor (sensitive parenting), which was calculated as the average of the sensitivity, detachment (reversed), stimulation, positive regard, and animation codes. We averaged the parenting scores across these four assessments to create a composite measure of sensitive parenting, $\alpha = .86$.

Poverty-related risk—A cumulative risk index was used to assess poverty and associated risk factors at the 6-, 15-, 24-, and 36-month home visits. Following Vernon-Feagans and colleagues (2013), we included parent education, household income-to-needs ratio, work hours, job prestige, household density, and neighborhood safety as risk indicators. All variables were coded so that positive scores indicated more risk, and the variables were subsequently standardized and averaged within each assessment. Correlations among the risk scores across assessments ranged from .81 to .90. Consequently, we standardized and averaged these variables to create a composite measure of poverty-related risk in early childhood.

Analysis Plan

Mplus 7.3 (Muthén & Muthén, 2012) was used for all LCA and LTA analyses. After examining correlations and descriptive statistics, we estimated an unconditional, cross-sectional LCA model separately at each of the two assessments. To prevent information from being lost to aggregation, these models were based on the item-level data: there were five parent items and five teacher items at each assessment, for a total of ten observed indicators at each assessment. Next, we estimated an LTA model that combined the two LCA models and we evaluated measurement invariance of the class structure by constraining the conditional response probabilities for each item to be equal across assessments. Finally, we added a set of predictors to these models. For all predictive models, we used the three-step method to adjust for measurement error (Asparouhov & Muthén, 2014). For the cross-sectional LCA models, we predicted the probability of class membership; for the LTA model, we predicted class membership at T2 while controlling for the effect of the covariates on class membership at T1. To evaluate relative model fit, we considered the Bayesian Information Criterion (BIC), the Lo-Mendell-Rubin likelihood ratio test, and the bootstrap likelihood ratio test (Nylund et al., 2007), as well as the interpretability of the results.

Results

Correlations and descriptive statistics for the SDQ scales and all other study variables are presented in Table 1. Most African American respondents were from North Carolina, $r = .62$. There were modest negative correlations between being female and teachers' (but not primary caregivers') ratings of child conduct problems at T1 and T2. Being African American and, to a lesser extent, residing in North Carolina (relative to Pennsylvania), was related to higher poverty-related risk, less sensitive parenting, low child executive function, and low child verbal intelligence. All child and family risk factors were significantly intercorrelated in the expected directions. The strongest of these relations were between

poverty-related risk and sensitive parenting, $r = -.61$, and between child executive function and verbal intelligence, $r = .61$.

The correlation between the T1 and T2 measures of conduct problems was substantial for primary caregivers' reports, $r = .56$, and for teacher's reports, $r = .42$. Within each assessment, primary caregivers' and teachers' reports of conduct problems were more modestly related, $r = .24$ at T1 and $r = .33$ at T2. All child and family risk factors exhibited small to moderate associations in the expected direction with primary caregivers' and teachers' reports of conduct problems.

Latent Class Analysis

When LCA models were estimated separately using data at T1 and T2, the model fit statistics did not clearly favor a particular number of classes (see Table 2). The BIC was smallest (indicating the best fit) for the three-class model at T1 and the five-class model at T2, whereas the Lo-Mendell-Rubin likelihood ratio test indicated that the five-class model fit best at T1 and the four-class model fit best at T2, and the bootstrap likelihood ratio test indicated that the five-class model fit best at both assessments. Disagreement among these different fit statistics is common, and investigators have been advised to base model selection on the “theoretical appropriateness, interpretability, and parsimony” of the solution (Specht, Luhmann, & Geiser, 2014, p. 546; see also Marsh, Lüdtke, Trautwein, & Morin, 2009). Our four-class solution was similar across assessments and was similar to the four-class solution found in previous research (De Los Reyes et al., 2009; Fergusson et al., 2009). The four-class model also (as described below) exhibited good separation between classes, with average class assignment probabilities exceeding .80 in all classes, and all classes differed qualitatively from one another; neither of these was true for solutions with five or more classes. Finally, the four-class solution provided a clear distinction between the relative levels of primary caregivers' and teachers' ratings of conduct problems at both T1 and T2, which is the focus of our research question (see Figure 1). For these reasons, we chose the four-class solution for subsequent analyses. For comparison, the three-class model was composed of a low cross-context class, a high cross-context class, and a home context class, but the school context class was absent (see Figure S1). Adding a fifth class resulted in the addition of a class with moderate-to-high parent symptoms and moderate teacher symptoms that was similar to the high cross-context class, but lower in average severity: this class differed quantitatively, rather than qualitatively (see Figure S2).

Across models with varying numbers of classes, not all items discriminated equally between the classes. In particular, the fifth item, “can be spiteful to others,” was less strongly related to class membership. Supplemental confirmatory factor analyses have indicated that this item has a high threshold: it is indicative of severe conduct problems and is rarely endorsed among lower-scoring individuals. Even though this item was less strongly related to class membership, it nonetheless did provide unique information and was retained in our analyses.

In decreasing prevalence, we identified the following classes in the four-class model (see Figure 2): a class characterized by low primary caregiver and low teacher ratings (low cross-context); a class characterized by high primary caregiver ratings and low teacher ratings (home context), a class characterized by low primary caregiver ratings and moderate teacher

ratings (school context); and a class characterized by high primary caregiver ratings and high teacher ratings (high cross-context).

In LCA, average class assignment probabilities quantify classification uncertainty, and have a maximum value of 1.00, which indicates no uncertainty with respect to classification. In the four-class model, these values ranged from .82 to .91 at T1 and from .81 to .91 at T2. These numbers indicate that separation between classes was good.

Latent Transition Analysis

Based on the results of the four-class LCA models at T1 and T2, we estimated an LTA model with and without constraints for longitudinal measurement invariance. These constraints are desirable because they allow us to interpret class membership in the same way at T1 and T2. In the LTA model without the measurement invariance constraints, the conditional response probabilities were freely estimated; in the LTA model with these constraints, the conditional response probabilities were constrained to be equal at T1 and T2. A likelihood ratio test using the Satorra-Bentler (2001) adjustment indicated that the unconstrained model provided better fit to the data, $\chi^2(80) = 170.433, p < .001$. However, the statistical significance of nested model comparisons is sensitive to sample size, and trivial differences can produce significant test results in large samples (Chen, 2007). The BIC for the unconstrained model was 23,998.062 and the BIC for the constrained model was 23,614.949, with smaller values indicating better fit. Although the likelihood ratio test comparing these two models was significant, we interpret the smaller BIC for the constrained model as supporting the plausibility of measurement invariance and use the constrained model in subsequent analyses.

Class proportions—As expected, the low cross-context class was the largest, comprising 54% of the sample at T1 and 55% at T2. The home context class was the second largest class at both assessments, decreasing from 27% at T1 to 20% at T2. The next largest class was the school context class, which increased from 12% at T1 to 16% at T2. The high cross-context class was the smallest class, increasing from 6% at T1 to 9% at T2.

Latent transition probabilities—Children were more likely to remain in their T1 class at T2 than to transition to a different class at T2 (see Figure 3; for results from the unconstrained model, see Figure S3). Stability in class membership was particularly high for the low cross-context class: children in the low cross-context class at T1 had an 82% chance of remaining in that class at T2. The corresponding percentages were 53% for the home context class, 56% for the school context class, and 51% for the high cross-context class. For members of the low cross-context class at T1, onset of problems was relatively rare: 9% for the home context class at T2, 8% for the school context class at T2, and 1% for the high cross-context class at T2. In contrast, symptoms were more likely to remit in the high cross-context class: children in the high cross-context class at T1 exhibited a relatively large probability of moving into the home context class (28%) or into the school context class (17%) at T2, but were unlikely to transition into the low cross-context class (5%). Children in the home context class at T1 were most likely to transition into the low cross-context class (21%) and approximately equally likely to transition into the school context class (13%) and

the high cross-context class (12%). Children in the school context class at T1 were most likely to transition into the low cross-context class (28%), followed by the high cross-context class (12%) and the home context class (4%).

Prediction of Class Membership

In separate multinomial logistic regression models predicting membership in the high cross-context class relative to each of the other classes (low cross-context, home context, school context), class membership at T1 and at T2 was regressed on each of the child and family predictors while controlling for the demographic covariates (i.e., state, race, and gender). The odds ratios from these models with 95% confidence intervals are presented in Table 3.

The high cross-context class was clearly differentiated from the low cross-context class on all child and family predictors at both assessments, with high poverty-related risk, low sensitive parenting, high primary caregiver depressive symptoms, low child executive function, and low child verbal ability all associated with greater probability of membership in the high cross-context class. For the comparison between the home context class and the high cross-context class, there were few differences. Low child executive function was the only variable that was associated with membership in the high cross-context class at both assessments. In addition, low child verbal ability predicted membership in the high cross-context class at T2 only. The predictors of membership in the high cross-context class relative to the school context class were substantially different at T1 and T2. At T1, all five predictors distinguished these two classes; in fact, the school context class was similar to the low cross-context class at T1. At T2, however, none of these variables predicted membership in the high cross-context class relative to the school context class except for child verbal ability. This was the only predictor that differentiated between these two classes at both assessments.

We subsequently tested whether each predictor was related to the latent transition probabilities from T1 to T2.¹ These analyses allow us to investigate whether the predictors were associated with movement among classes from T1 to T2 while controlling for class membership at T1. Because there was substantial stability in class membership, we expected the prediction of the latent class transition probabilities to be limited. There was only one significant predictor of movement into or out of the high cross-context class: for children in the low cross-context class at T1, poverty-related risk was positively related to the probability of moving into the high cross-context class, OR = 1.51, 95% CI [1.07, 2.14], $p = .019$.

Discussion

In this paper, we used LCA, a person-centered analysis method, to examine patterns of children's conduct problems across the home and school contexts. We identified four relatively distinct patterns of symptoms: low cross-context, home context, school context, and high cross-context. These patterns showed substantial stability (> 50%) over a two-year

¹In the LTA model that included executive function as a predictor, there were problems with the standard errors for the executive function coefficients. Consequently, we do not report results for that model because they do not seem to be trustworthy.

period from age 5–7. Finally, we identified a number of child and family predictors that were associated with different patterns of symptomatology.

Descriptive Patterns of Conduct Problems

At each assessment, class separation was acceptable. For the four-class model, average classification probabilities exceeded .80 for each class. The classes were highly interpretable, and to a large degree conformed to our hypotheses: we identified a pattern characterized by low symptoms across contexts, a pattern characterized by high symptoms across contexts, a pattern characterized by greater conduct problems in the home context, and a class characterized by greater conduct problems in the school context. As expected, the low cross-context class was most common and the high cross-context class was the least common, which is consistent with the positively skewed distribution of parent and teacher ratings of child conduct problems. Furthermore, the class structure was similar at the five-year and seven-year assessments. The results suggest that children who are highest in symptomatology tend to exhibit symptoms in both the home and school context, whereas children who are more moderate in symptomatology tend to primarily exhibit symptoms in only one context. The classes did not perfectly conform to our predictions, however. Although the difference between parents' and teachers' reports of conduct problems was still meaningful for the school context class, this difference was relatively small for this class (especially at the age-five assessment; see Figure 2) in comparison to the difference between parents' and teachers' ratings of conduct problems in the home context class. A strength of our analytic approach is that these differences among classes are empirically informed, but a corresponding limitation is that interpretation of these classes can be complicated by these differences.

Stability in class membership from age 5 to age 7 was at least moderate for each of the classes. For all four classes, participants were more likely to remain in the same class across this two-year period than to transition to another class, which is especially remarkable because few children ($n = 7$) had the same teacher at both assessments. Some transition probabilities were more likely than others. In particular, we observed that children who do not exhibit problems in either context are unlikely to spontaneously develop symptoms—and when they did, it was almost always in only one context (home or school, but not both). In all three classes with elevated problems in at least one context, desistence was more likely that escalation, and children were relatively unlikely to switch from exhibiting conduct problems only at school to only exhibiting conduct problems at home, or vice versa. For children who exhibited high symptoms across both contexts, *partial* desistence was common, but *complete* desistence remained rare. This extends prior work on the stability of behavior problems based on a single informant (Pianta & Caldwell, 1990; Verhulst & Van Der Ende, 1992) by looking at the patterns of continuity and change in symptomatology across the home and school context simultaneously.

Prediction of High Conduct Problems across Contexts

By grouping individuals into different patterns of symptomatology, we were able to predict different patterns of conduct problems across the home and school contexts. All five child and family variables that we tested (i.e., poverty-related risk, sensitive parenting, parental

depression, child executive function, and child verbal ability) predicted membership in the high cross-context class relative to the low-cross context class. There were fewer differences between the classes that were characterized by conduct problems in a single context (i.e., the home context and the school context classes) and the high cross-context class, suggesting that risk factors for these different patterns were largely similar.

The exceptions to this include the child characteristics that we tested as predictors: executive function and verbal ability were protective against membership in the high cross-context class relative to the home context or school context class (or both) at the age 5 and the age 7 assessment. These results indicate that these child characteristics could play a particularly important role in protecting against the occurrence of severe and generalizable patterns of conduct problems that are more consistent with clinical diagnosis (R. Goodman, Renfrew, et al., 2000). There is evidence that executive functions such as inhibitory control are already associated with anger in early childhood (Gagne & Goldsmith, 2011). Executive functions such as strong attention regulation, inhibitory control, and planning could support effective emotion regulation strategies such as positive reappraisal, inhibition of emotional responses, and situation selection or modification (Gross, 1998). Furthermore, good verbal skills may facilitate communication or other alternatives to disruptive or aggressive behavior (Greenberg, Kusché, & Speltz, 1991). An important direction for future research will be to investigate the specific mechanisms linking neurocognitive deficits to conduct problems.

We also noted a shift in prediction for the high cross-context class relative to the school context class between the age 5 and age 7 assessments. At the age 5 assessment, the school context class was distinct from the high cross-context class on all three environmental predictors. This changed by the age 7 assessment, when the environmental predictors no longer distinguished between the school class and the high cross context class. At approximately age 5, many children are adjusting to entry into formal schooling. These results may suggest that—uniquely for children who do not already exhibit high conduct problems in the home context—teachers' ratings of conduct problems during this period of transition are less reliably related to environmental variables that have been found to confer risk for conduct problems (e.g., Campbell, 1995).

Limitations

The absence of a best-fitting model could be because the classes do not represent qualitatively distinct groups. Class separation in our models—although good—was not perfect. Some individuals were on the boundary between classes and consequently had a relatively high probability of membership in more than one class. This may have been exacerbated by the relatively low reliability of the Strengths and Difficulties Questionnaire Conduct Problems scale, which consists of only five items. To some extent, these issues were mitigated by the use of the three-step modeling procedure in the predictive analyses, which takes uncertainty in classification into account (Asparouhov & Muthén, 2014). Despite some uncertainty with respect to classification, we believe that the use of a person-centered approach such as LCA is useful because it allows us to address research questions about different patterns of symptoms across informants that are difficult to answer using variable-centered methods such as confirmatory factor analysis. As Nagin (2005) has

pointed out, mixture models (such as LCA) can be used to identify prototypical patterns that are broadly representative of a large number of individuals, even if these patterns are not qualitatively distinct in the population.

In addition, we were not able to distinguish between informant effects (parent vs. teacher) and context effects (home vs. school). The differences between parents' and teachers' social roles could contribute to differences among these informants that are independent of the context in which the behavior is examined. For example, teachers are often highly concerned with classroom management and could be more aware of children's disruptive behavior than are parents. Furthermore, parents and teachers can share information, and therefore their knowledge of child behavior problems may be influenced by the child's behavior in other contexts. For example, parents may become aware of a child's conduct problems at school when disciplinary actions are taken (e.g., suspension). Differentiating between informant and context effects is an important direction for future research (De Los Reyes et al., 2013; Kraemer et al., 2003).

Conclusions

Our results indicate that LCA can be a useful method for investigating patterns of child conduct problems across the home and school contexts to obtain a more holistic view of child symptomatology. We obtained a four-class model that was similar across a two-year period and for which the classes had a relatively straightforward interpretation, and were able to examine how the pattern of conduct problems across the home and school contexts changed from age 5 to age 7. By identifying a group of children who exhibited high symptoms across the home and school context, we were able to identify executive function and verbal ability as two predictors that are uniquely protective against this high-risk pattern.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the many families and research assistants who made this study possible. Support for this research was provided by the National Institute of Child Health and Human Development Grants R01 HD51502 and P01 HD39667 with co-funding from the National Institute on Drug Abuse.

Support for this research was provided by the National Institute of Child Health and Human Development Grants R01 HD51502 and P01 HD39667 with co-funding from the National Institute on Drug Abuse.

References

- 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. DOI: 10.1037/0033-2909.101.2.213 [PubMed: 3562706]
- Ackerman BP, Brown E, Izard CE. Continuity and change in levels of externalizing behavior in school of children from economically disadvantaged families. *Child Development*. 2003; 74:694–709. DOI: 10.1111/1467-8624.00563 [PubMed: 12795385]
- Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*. 2014; 21:329–341. DOI: 10.1080/10705511.2014.915181

- Belsky J, Fearon RM, Bell B. Parenting, attention and externalizing problems: Testing mediation longitudinally, repeatedly and reciprocally. *Journal of Child Psychology and Psychiatry*. 2007; 48:1233–1242. DOI: 10.1111/j.1469-7610.2007.01807.x [PubMed: 18093029]
- Bollen, KA. Structural equations with latent variables. New York, NY: Wiley; 1989.
- Campbell SB. Hard-to-manage preschool boys: Externalizing behavior, social competence, and family context at two-year followup. *Journal of Abnormal Child Psychology*. 1994; 22:147–166. DOI: 10.1007/BF02167897 [PubMed: 8064027]
- Campbell SB. Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry*. 1995; 36:113–149. DOI: 10.1111/j.1469-7610.1995.tb01657.x [PubMed: 7714027]
- Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*. 2007; 14:464–504. DOI: 10.1080/10705510701301834
- Collins, LM., Lanza, ST. Latent class and latent transition analysis for the social, behavioral, and health sciences. New York, NY: Wiley; 2010.
- Collins LM, Wugalter SE. Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research*. 1992; 27:131–157. DOI: 10.1207/s15327906mbr2701_8
- Cox, M., Crnic, K. Qualitative ratings for parent-child interaction at 3–12 months of age. Unpublished Manuscript, Department of Psychology, University of North Carolina at Chapel Hill; 2002.
- De Los Reyes A, Augenstein TM, Wang M, Thomas SA, Drabick DA, Burgers DE, Rabinowitz J. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*. 2015; 141:858–900. DOI: 10.1037/a0038498 [PubMed: 25915035]
- De Los Reyes A, Henry DB, Tolan PH, Wakschlag LS. Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*. 2009; 37:637–652. DOI: 10.1007/s10802-009-9307-3 [PubMed: 19247829]
- De Los Reyes A, Kazdin AE. Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*. 2005; 131:483–509. DOI: 10.1037/0033-2909.131.4.483 [PubMed: 16060799]
- De Los Reyes A, Thomas SA, Goodman KL, Kundey SM. Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*. 2013; 9:123–149. DOI: 10.1146/annurev-clinpsy-050212-185617
- Deater-Deckard K, Dodge KA, Bates JE, Pettit GS. Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology*. 1998; 10:469–493. DOI: 10.1017/S0954579498001709 [PubMed: 9741678]
- Dunn, LM., Dunn, DM. Peabody picture vocabulary test. 4th. Minneapolis, MN: Pearson; 2007.
- Eisenberg N, Zhou Q, Spinrad TL, Valiente C, Fabes RA, Liew J. Relations among positive parenting, children's effortful control, and externalizing problems: A three-wave longitudinal study. *Child Development*. 2005; 76:1055–1071. DOI: 10.1111/j.1467-8624.2005.00897.x [PubMed: 16150002]
- Fergusson DM, Boden JM, Horwood LJ. Situational and generalised conduct problems and later life outcomes: evidence from a New Zealand birth cohort. *Journal of Child Psychology and Psychiatry*. 2009; 50:1084–1092. DOI: 10.1111/j.1469-7610.2009.02070.x [PubMed: 19298467]
- Gagne JR, Goldsmith HH. A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. *Developmental Science*. 2011; 14:112–124. DOI: 10.1111/j.1467-7687.2010.00969.x [PubMed: 21159093]
- Goodman R. The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*. 1997; 38:581–586. DOI: 10.1111/j.1469-7610.1997.tb01545.x [PubMed: 9255702]
- Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry*. 2000; 177:534–539. DOI: 10.1192/bjp.177.6.534 [PubMed: 11102329]
- Goodman R, Renfrew D, Mullick M. Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka.

- European Child & Adolescent Psychiatry. 2000; 9:129–134. DOI: 10.1007/s007870050008 [PubMed: 10926063]
- Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal depression and child psychopathology: A meta-analytic review. *Clinical Child and Family Psychology Review*. 2011; 14:1–27. DOI: 10.1007/s10567-010-0080-1 [PubMed: 21052833]
- Greenberg, MT., Kusché, CA., Speltz, M. Emotional regulation, self-control and psychopathology: The role of relationships in early childhood. In: Cicchetti, D., Toth, S., editors. *Rochester Symposium on Developmental Psychopathology*. Vol. 2. Hillsdale, NJ: Lawrence Erlbaum Associates; 1991.
- Gross JJ. The emerging field of emotion regulation: An integrative review. *Review of General Psychology*. 1998; 2:271–299. DOI: 10.1037/1089-2680.2.3.271
- Hammond SI, Müller U, Carpendale JI, Bibok MB, Liebermann-Finestone DP. The effects of parental scaffolding on preschoolers' executive function. *Developmental Psychology*. 2012; 48:271–281. DOI: 10.1037/a0025519 [PubMed: 21928877]
- Hughes C. Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*. 1998; 16:233–253. DOI: 10.1111/j.2044-835X.1998.tb00921.x
- Kraemer HC, Measelle JR, Ablow JC, Essex MJ, Boyce WT, Kupfer DJ. A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry*. 2003; 160:1566–1577. DOI: 10.1176/appi.ajp.160.9.1566 [PubMed: 12944328]
- Landry SH, Miller-Loncar CL, Smith KE, Swank PR. The role of early parenting in children's development of executive processes. *Developmental Neuropsychology*. 2002; 21:15–41. DOI: 10.1207/S15326942DN2101_2 [PubMed: 12058834]
- Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging*. 1997; 12:277–287. DOI: 10.1037/0882-7974.12.2.277 [PubMed: 9189988]
- Loney BR, Frick PJ, Ellis M, McCoy MG. Intelligence, callous-unemotional traits, and antisocial behavior. *Journal of Psychopathology and Behavioral Assessment*. 1998; 20:231–247. DOI: 10.1023/A:1023015318156
- MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychological Methods*. 2002; 7:19–40. DOI: 10.1037//1082-989X.7.1.19 [PubMed: 11928888]
- Marsh HW, Lüdtke O, Trautwein U, Morin AJ. Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling*. 2009; 16:191–225. DOI: 10.1080/10705510902751010
- Mills-Koonce WR, Garrett-Peters P, Barnett M, Granger DA, Blair C, Cox MJ. The Family Life Project Key Investigators. Father contributions to cortisol responses in infancy and toddlerhood. *Developmental Psychology*. 2011; 47:388–395. DOI: 10.1037/a0021066 [PubMed: 21142362]
- Miner JL, Clarke-Stewart KA. Trajectories of externalizing behavior from age 2 to age 9: Relations with gender, temperament, ethnicity, parenting, and rater. *Developmental Psychology*. 2008; 44:771–786. DOI: 10.1037/0012-1649.44.3.771 [PubMed: 18473643]
- Moffitt TE. The neuropsychology of conduct disorder. *Development and Psychopathology*. 1993; 5:135–151. DOI: 10.1017/S0954579400004302
- Muthén, LK., Muthén, BO. 7th. Los Angeles, CA: Muthén & Muthén; 2012. Mplus User's Guide. Retrieved from <http://www.statmodel.com>
- Nagin, DS. Group-based modeling of development. Cambridge, MA: Harvard University Press; 2005.
- NICHD Early Child Care Research Network. Chronicity of maternal depressive symptoms, maternal sensitivity, and child functioning at 36 months. *Developmental Psychology*. 1999; 35:1297–1310. DOI: 10.1037/0012-1649.35.5.1297 [PubMed: 10493655]
- Nigg JT, Quamma JP, Greenberg MT, Kusche CA. A two-year longitudinal study of neuropsychological and cognitive performance in relation to behavioral problems and competencies in elementary school children. *Journal of Abnormal Child Psychology*. 1999; 27:51–63. DOI: 10.1023/A:1022614407893 [PubMed: 10197406]

- Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*. 2007; 14:535–569. DOI: 10.1080/10705510701575396
- Pianta RC, Caldwell C. Stability of externalizing symptoms from kindergarten to first grade and factors related to instability. *Development and Psychopathology*. 1990; 2:247. doi: 10.1017/S0954579400000754
- Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977; 1:385–401. DOI: 10.1177/014662167700100306
- Sameroff AJ, Seifer R, Baldwin A, Baldwin C. Stability of intelligence from preschool to adolescence: The influence of social and family risk factors. *Child Development*. 1993; 64:80–97. DOI: 10.2307/1131438 [PubMed: 8436039]
- Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*. 2001; 66:507–514. DOI: 10.1007/BF02296192
- Specht J, Luhmann M, Geiser C. On the consistency of personality types across adulthood: Latent profile analyses in two large-scale panel studies. *Journal of Personality and Social Psychology*. 2014; 107:540–556. DOI: 10.1037/a0036863 [PubMed: 25133730]
- Speltz ML, DeKlyen M, Calderon R, Greenberg MT, Fisher PA. Neuropsychological characteristics and test behaviors of boys with early onset conduct problems. *Journal of Abnormal Psychology*. 1999; 108:315–325. DOI: 10.1037/0021-843X.108.2.315 [PubMed: 10369042]
- Sulik MJ, Blair C, Mills-Koonce R, Berry D, Greenberg M. The Family Life Project Investigators. Early parenting and the development of externalizing behavior problems: Longitudinal mediation through children's executive function. *Child Development*. 2015; 86:1588–1603. DOI: 10.1111/cdev.12386 [PubMed: 26082032]
- Trentacosta CJ, Hyde LW, Shaw DS, Dishion TJ, Gardner F, Wilson M. The relations among cumulative risk, parenting, and behavior problems during early childhood. *Journal of Child Psychology and Psychiatry*. 2008; 49:1211–1219. DOI: 10.1111/j.1469-7610.2008.01941.x [PubMed: 18665880]
- U.S. Department of Health and Human Services. 2014. Retrieved from <http://aspe.hhs.gov/poverty/figures-fed-reg.cfm>
- Verhulst FC, Van Der Ende J. Six-year developmental course of internalizing and externalizing problem behaviors. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1992; 31:924–931. DOI: 10.1097/00004583-199209000-00022 [PubMed: 1400127]
- Vernon-Feagans L, Cox M. The Family Life Project Key Investigators. The Family Life Project: An epidemiological and developmental study of young children living in poor rural communities. *Monographs of the Society for Research in Child Development*. 2013; 78(5):vii–150.
- Willoughby MT, Wirth RJ, Blair CB. Contributions of modern measurement theory to measuring executive function in early childhood: An empirical demonstration. *Journal of Experimental Child Psychology*. 2011; 108:414–435. DOI: 10.1016/j.jecp.2010.04.007 [PubMed: 20553690]
- Willoughby MT, Wirth RJ, Blair CB. The Family Life Project Investigators. Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*. 2012; 24:418–431. DOI: 10.1037/a0025779 [PubMed: 22023561]
- Youthinmind Ltd. Scoring the Strengths & Difficulties Questionnaire for age 4–17. 2015. Retrieved from <http://www.youthinmind.com/>
- Zumbo BD, Gaderman AM, Zeisser C. Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*. 2007; 6:21–29.

Figure 1.

Conditional response probabilities of responding to the second and third response categories (combined) for the four-class LCA model. The conditional response probabilities indicate the probability that an individual in a given class will endorse an item. P = Primary Caregiver. T = Teacher. Classes are numbered in decreasing prevalence.

Figure 2.

Means and standard deviations for primary caregivers' and teachers' ratings of conduct problems in each class at Time 1 and Time 2. The range of this scale is 0–10, with a score of 3 indicating “possible” risk for conduct disorder and a score of 4 or more indicating “probable” risk for conduct disorder (R. Goodman, Renfrew, et al., 2000; Youthmind Ltd., 2015).

Figure 3. Latent Transition Probabilities from Time 1 to Time 2 for the Latent Transition Analysis Model with Measurement Invariance Constraints

Table 1

Correlations and Descriptive Statistics

	1	2	3	4	5	6	7	8	9	10	11	12
1 Female		0.00	0.07	0.02	0.02	0.04	0.15	0.04	-0.05	-0.14	-0.09	-0.13
2 African American			0.62	0.43	-0.41	0.11	-0.35	-0.44	0.12	0.17	0.05	0.20
3 North Carolina				0.27	-0.28	0.01	-0.31	-0.36	0.02	0.09	-0.03	0.08
4 Poverty-Related Risk					-0.61	0.30	-0.41	-0.49	0.27	0.16	0.22	0.20
5 Sensitive Parenting						-0.26	0.43	0.48	-0.26	-0.17	-0.22	-0.20
6 Primary Caregiver Depression							-0.17	-0.23	0.35	0.05	0.29	0.12
7 Child Executive Function								0.61	-0.22	-0.26	-0.19	-0.25
8 Child Verbal Intelligence									-0.26	-0.20	-0.16	-0.22
9 Child Conduct Problems, (P, T1)										0.24	0.56	0.25
10 Child Conduct Problems, (T, T1)											0.26	0.42
11 Child Conduct Problems, (P, T2)												0.33
12 Child Conduct Problems, (T, T2)												
<i>M</i>	0.50	0.43	0.60	0.00	2.85	0.57	0.00	94.10	0.39	0.23	0.38	0.28
<i>SD</i>	–	–	–	1.00	0.66	0.43	1.00	15.88	0.36	0.37	0.39	0.41

P = Primary Caregiver; T = Teacher. T1 = Time 1; T2 = Time 2. Bold = $p < .001$

Table 2
Fit Statistics for Latent Class Analysis Models at Time 1 and Time 2

Time 1				
Number of Classes	BIC	Entropy	Lo-Mendell-Rubin LRT	Bootstrap LRT
2	12065.412	.769		
3	11763.934	.760	< .001	< .001
4	11776.457	.736	.007	< .001
5	11818.862	.714	.003	< .001
Time 2				
Number of Classes	BIC	Entropy	Lo-Mendell-Rubin LRT	Bootstrap LRT
2	12850.007	.827		
3	12619.702	.780	.002	< .001
4	12561.512	.760	.002	< .001
5	12546.276	.752	.075	< .001

BIC = Bayesian Information Criterion. LRT = Likelihood ratio test. Bold indicates the model that was selected for further analysis.

Table 3
Prediction of Membership in the High Cross-Context Class at Time 1 and Time 2: Odds Ratios with 95% Confidence Intervals

Reference Group						
	Low Cross Context		Home Context		School Context	
	Time 1					
	OR	95% CI	OR	95% CI	OR	95% CI
Poverty-Related Risk	1.97	[1.29, 3.02]	0.94	[0.61, 1.47]	1.80	[1.06, 3.08]
Sensitive Parenting	0.35	[0.22, 0.57]	0.82	[0.49, 1.37]	0.35	[0.20, 0.62]
Primary Caregiver Depression	6.56	[3.08, 13.94]	0.62	[0.33, 1.16]	5.18	[1.81, 14.84]
Child Executive Function	0.37	[0.26, 0.52]	0.62	[0.44, 0.87]	0.46	[0.30, 0.72]
Child Verbal Intelligence	0.94	[0.92, 0.97]	0.99	[0.96, 1.01]	0.95	[0.92, 0.98]
	Time 2					
	OR	95% CI	OR	95% CI	OR	95% CI
Poverty-Related Risk	1.89	[1.44, 2.48]	1.03	[0.76, 1.38]	1.32	[0.96, 1.80]
Sensitive Parenting	0.38	[0.25, 0.59]	0.84	[0.52, 1.34]	0.63	[0.38, 1.04]
Primary Caregiver Depression	3.71	[2.06, 6.69]	0.85	[0.47, 1.52]	1.81	[0.90, 3.61]
Child Executive Function	0.46	[0.36, 0.58]	0.72	[0.55, 0.95]	0.77	[0.58, 1.02]
Child Verbal Intelligence	0.95	[0.93, 0.97]	0.97	[0.95, 0.99]	0.97	[0.95, 1.00]

Odds ratios greater than 1 indicate that increases in a variable are associated with a greater likelihood of being in the high cross-context class relative to the reference group. Bold = $p < .05$.