HHS Public Access

Author manuscript

Dev Psychopathol. Author manuscript; available in PMC 2015 August 01.

Published in final edited form as:

Dev Psychopathol. 2015 August ; 27(3): 791–818. doi:10.1017/S0954579414000789.

Describing and Predicting Developmental Profiles of Externalizing Problems from Childhood to Adulthood

Isaac T. Petersen,
Indiana University

John E. Bates,
Indiana University

Kenneth A. Dodge,
Duke University

Jennifer E. Lansford, and
Duke University

Gregory S. Pettit
Auburn University

Abstract

This longitudinal study considers externalizing behavior problems from ages 5 to 27 (N = 585). Externalizing problem ratings by mothers, fathers, teachers, peers, and self-report were modeled with growth curves. Risk and protective factors across many different domains and time frames were included as predictors of the trajectories. A major contribution of the study is in demonstrating how heterotypic continuity and changing measures can be handled in modeling changes in externalizing behavior over long developmental periods. On average, externalizing problems decreased from early childhood to preadolescence, increased during adolescence, and decreased from late adolescence to adulthood. There was strong nonlinear continuity in externalizing problems over time. Family process, peer process, stress, and individual characteristics predicted externalizing problems beyond the strong continuity of externalizing problems. The model accounted for 70% of the variability in the development of externalizing problems. The model's predicted values showed moderate sensitivity and specificity in prediction of arrests, illegal drug use, and drunk driving. Overall, the study showed that by using changing, developmentally-relevant measures and simultaneously taking into account numerous characteristics of children and their living situations, research can model lengthy spans of development and improve predictions of the development of later, severe externalizing problems.

Keywords

Externalizing/antisocial conduct behavior problems; heterotypic continuity; developmental trajectories; changing measures; developmental actuarial prediction model

Correspondence concerning this article should be addressed to: Isaac T. Petersen, Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47405. itpeters@indiana.edu.

The ultimate goal of developmental psychopathology is to understand the whole trajectory of an individual's life, and not just transitory outcomes at a particular point in life. Thus, research should strive to build a bridge that spans from childhood to adulthood (Rutter & Sroufe, 2000). Very few studies have measured change in externalizing problems across long spans of development in one piece, from childhood to adulthood. There are several reasons for the difficulties in measuring change across lengthy developmental spans: 1) it is costly, 2) it takes a long time, and most importantly, 3) there are difficult conceptual and statistical issues that need to be addressed with respect to measurement. The difficulty is in comparing measurements in childhood with measurements in adulthood in a way that allows one to infer that differences in scores on a measure across time reflect true change rather than differences in the meaning of the measure. The present study considers externalizing behavior problems across childhood to adulthood. We describe forms of growth and we use early risk factors to predict who will be at greatest risk for problem trajectories. We also employ risk and protective factors from successive developmental periods to shed light on the mechanisms of externalizing behavior across development.

Predicting Externalizing Problems

Traditionally, research has been quite poor in predicting behavior (e.g., Sutton, 1998; Underwood, 1979). The issue of prediction of later behavior problems is crucial from a policy perspective because early identification of at-risk individuals may be crucial to prevention. Accurate prediction provides important tools for many aspects of public policy, including prison parole, personnel hiring, and security agency clearance. Prediction is generally made by judgments (by experts or laypeople) or by a formula that weighs various characteristics of the individual (Underwood, 1979). One type of formula employs an actuarial model that takes into account many different characteristics of the individual, including risk and protective factors. Because of the actuarial model's ability to weigh many risk and protective factors simultaneously, the actuarial approach has consistently been shown to be more accurate than judgment-based predictions (Dawes, Faust, & Meehl, 1989). The present study uses an actuarial approach to predict the development of later externalizing problems.

Development and Heterotypic Continuity of Externalizing Problems

Externalizing behavior problems do not emerge suddenly. A high degree of age-to-age stability in externalizing problems appears to be the rule (Fergusson, 1998; Olweus, 1979). Therefore, developmental psychopathology research often describes trajectories of externalizing problems across development. However, to do so across major developmental eras, one must deal with changes in how externalizing behavior is manifested from childhood to adolescence (Olson et al., 2013). For example, young children exhibit more physical aggression (e.g., biting, kicking) and adolescents engage in different forms of externalizing (e.g., drug use, delinquency, indirect aggression; Achenbach, Howell, McConaughy, & Stanger, 1995; Miller, Vaillancourt, & Boyle, 2009). A particular scale may not actually measure the same construct at different ages (Widaman, Ferrer, & Conger, 2010). Thus, changes in measurement may need to accompany changes in externalizing behavior in order for the measures to remain developmentally relevant. The different

externalizing behavior manifestations over time do show heterotypic continuity or coherence (for excellent discussions of coherence, see Caspi, 1998; Miller et al., 2009).

McArdle and colleagues (McArdle, Grimm, Hamagami, Bowles, & Meredith, 2009) have argued that it is not theoretically desirable or necessary for developmental studies to require the same measures over time. Many researchers have supported the use of changing, developmentally-appropriate measures over time (e.g., Eddy, Dishion, & Stoolmiller, 1998; Knight & Zerr, 2010; McArdle et al., 2009; Owens & Shaw, 2003), and others have suggested that the measurement of heterotypic continuity should receive more attention (Schulenberg & Maslowsky, 2009). Several previous studies have implemented growth curves with changing measures. For example, similar to the approach of the present study, Owens and Shaw (2003) predicted growth curves of externalizing problems using different Achenbach scales across different time frames to maintain developmental relevance of the measures. Other studies have examined trajectories of externalizing problems with changing measures (Brame, Nagin, & Tremblay, 2001; Patterson, 1993) or trajectories of other phenotypes (e.g., Pettit, Keiley, Laird, Bates, & Dodge, 2007). However, to infer heterotypic continuity and changes in behavior problems over time with different measures, the measures must be comparable conceptually and empirically. Otherwise, apparent changes in behavior could be due to changes in the functioning of measures.

Ensuring statistical equivalence for comparing scores on different measures

Several statistical approaches have been used to increase comparability of measures. One approach is to standardize or age-norm measures across time (e.g., a *T*- or *z*-score), which places variables on a standard normal metric. Researchers are often cautioned against standardizing variables in longitudinal designs, however (Stoolmiller, 1995; Willett, Singer, & Martin, 1998). This practice would actually prevent observing changes in means or variances across time because standardization holds them constant.

A more promising approach recommended by Little (2013) for longitudinal designs may be a proportional scoring metric, such as proportion-of-maximum (POM) scoring. POM scoring divides each individual's score on a measure by the total possible score, rendering the individual's score a proportion of the maximum possible, with the assumption that similar proportions correspond to similar trait levels. Because all proportions have the same possible range (0–1), they have greater comparability than the raw metric, and unlike standardization, still allow growth over time. Another advantage of POM scoring over standardization in growth curve models is that it does not distort any of the fundamental statistics of the variable to provide a "reasonably comparable scale" (McArdle, Hamagami, Meredith, & Bradway, 2000, p. 60).

Ensuring conceptual equivalence for comparing scores on different measures

POM rescaling approaches do not ensure, however, that variables at different ages are on the same conceptual metric. In order to ensure this, construct validity invariance is also necessary (Knight & Zerr, 2010). In other words, although identical measures over time are unnecessary, the measures should have identical meaning across the time frame of the study (Owens & Shaw, 2003). There are many ways to develop construct validity of a set of

measures for a given construct. First, the items selected for the measures should be based on theory—they should be judged to reflect the same construct and the items should adequately sample the different facets of the construct (content or face validity). Second, despite heterotypic continuity in the long-term in the case of externalizing behavior, there should be short-term test-retest reliability of the measures across time. Third, the measures should show convergent validity with each other and discriminant validity with measures of distinct constructs. Fourth, the measures should demonstrate a similar factor structure across time, yet might not be expected to have an invariant structure because of qualitative changes in the factor structure with age. Fifth, the measures should have high internal consistency. In sum, in order to model externalizing trajectories, it is important for measures to have theoretical relevance to the construct at each age examined and to be on a comparable metric for measurement equivalence (as opposed to measurement invariance, which is unnecessary in cases of heterotypic continuity; Knight & Zerr, 2010).

Studies of Trajectories of Externalizing Problems

In the present study, we apply the preceding considerations to the study of trajectories of externalizing behavior problems across many years of development. Several studies have examined trajectories (defined here as 3 or more measurement occasions) of externalizing problems, including Odgers et al. (2008), which examined trajectories of antisocial conduct problems in individuals from New Zealand from ages 7–26 with 8 measurement occasions. A study using six longitudinal data sets examined trajectories of disruptive behaviors and delinquency, with the longest trajectory spanning ages 7–15 with 7 measurement occasions (Broidy et al., 2003). Another notable study examined trajectories of various externalizing problems in children from the Netherlands from ages 4–18 with 5 measurement occasions (Bongers, Koot, Ende, & Verhulst, 2004). Other studies have examined trajectories of other externalizing phenotypes, including aggression from ages 8–30 with 3 measurement occasions (Huesmann, Eron, Lefkowitz, & Walder, 1984) and from ages 8–42 with 4 measurement occasions (Kokko, Pulkkinen, Huesmann, Dubow, & Boxer, 2009). Studies have also examined trajectories of delinquency from ages 7–19 with 13 measurement occasions (Keijsers, Loeber, Branje, & Meeus, 2012) and from ages 8–46 with 9 measurement occasions (Farrington, 2003). In general, previous studies have shown that, on average, externalizing problems decrease from early to middle childhood (Keiley, Bates, Dodge, & Pettit, 2000; Leve, Kim, & Pears, 2005), increase during adolescence (Sampson & Laub, 2003), and decrease from adolescence to adulthood (Sampson & Laub, 2003). We seek to extend prior studies by examining externalizing problems of children from the United States annually from ages 5–27 (except ages 18, 25, and 26) with 20 measurement occasions, and to evaluate multiple risk factors as predictors of these trajectories. Similar to Odgers et al. (2008), we obtained measures from different, although developmentally-appropriate, sources at different ages.

Choosing the Developmental Model of Growth

We modeled the trajectories with growth curve models. A limitation of growth curve models is that they assume that all individuals can be described by the same parameters of change (e.g., everyone shows a quadratic trajectory; Connell & Frye, 2006). Growth curve models

do not assume, however, that individuals' change is homogeneous. In the present study, the quadratic form was allowed to vary across all individuals, allowing each individual to have a different trajectory (with different intercepts, slopes, and curvatures). This form of describing trajectories is in contrast to previous studies that have examined trajectories of subgroups of people. For example, similar to the developmental taxonomy proposed by Moffitt (1993), Odgers et al. (2008) identified four different subgroups that followed different trajectories: life-course persistent, adolescent-onset, childhood-limited, and low. At the very least, subgroup modeling can be useful as a heuristic of some general patterns of externalizing trajectories and simplified characterizations of continuous trajectories. An assumption in modeling subgroups' trajectories, however, is that all individuals within a subgroup follow a similar (though not necessarily the same) trajectory, and yet qualitatively different from the trajectories of other subgroups (Bauer & Reyes, 2010). Subgroup modeling has been known to identify illusory subgroups (Bauer & Curran, 2003).

We opted to model growth curves by treating individual differences in externalizing behavior trajectories as dimensional rather than categorical because we find the dimensional models richer. Moreover, there is evidence that externalizing problems are dimensional, not categorical (Coghill & Sonuga-Barke, 2012; Krueger, Markon, Patrick, & Iacono, 2005; Markon & Krueger, 2005; Walton, Ormel, & Krueger, 2011). Researchers have argued that theory provides stronger support for modeling the heterogeneity of developmental trajectories by allowing trajectories to differ continuously rather than categorically in order to describe individual trajectories more parsimoniously and accurately, both conceptually and empirically (Bauer, 2007; Little, Card, Preacher, & McConnell, 2009). Three recent studies found that externalizing problem trajectories are more accurately modeled dimensionally than with Moffitt's (1993) subgroups (Walters, 2011, 2012; Walters & Ruscio, 2013). Moreover, Burt, Donnellan, Iacono, and McGue (2011) found that subdimensions of externalizing problems were more strongly predictive of later antisocial behavior than was the age of antisocial behavior onset, which has often been used in determining subtypes. These findings suggest that the identified subtypes differ quantitatively in degree of severity rather than qualitatively in kind.

Risk Factors in the Development of Externalizing Problems

In addition to modeling trajectories of externalizing problems, we also examined many risk factors as predictors of individuals' trajectories. Risk factors tend to co-occur, so considering them together may provide a more accurate estimate of the unique contributions of individual risk factors. It is also important to consider the timing of risk factors to clarify the developmental process. Some risk factors such as child maltreatment appear to have stronger effects on externalizing problems in later than early childhood (Keiley, Howe, Dodge, Bates, & Pettit, 2001), whereas many other risk factors appear to have stronger effects earlier in childhood (Appleyard, Egeland, Van Dulmen, & Sroufe, 2005). In their review, Dodge, Coie, and Lynam (2006) highlighted the importance of genetics, child temperament, language ability, pregnancy complications, poverty, non-family child care, family processes, and peer processes in the etiology of externalizing behavior problems. They also observed, however, that these risk factors may be markers of other causal processes or may mediate the effects of each other (e.g., poverty may influence externalizing

behavior in part via its effect on parenting). To clarify the independent roles of these risk factors in the development of externalizing problems, we evaluated many of these domains of processes identified by Dodge and colleagues. Risk domains evaluated in the present study include demographic characteristics, aspects of parenting, parental adjustment, peer influences, child characteristics, stress, pregnancy, family background, and child activities. We selected risk and protective factors on the basis of previous studies' findings. Even if their mechanisms in the development of externalizing problems are not fully explained, successful prediction promotes prevention, targeted intervention, better choice of intervention and also aids, ultimately, in the development of process models (Sutton, 1998). Below we list risk factors selected for the present study, and the research supporting their selection:

Demographic characteristics selected as predictors included sex (Deater-Deckard, Dodge, Bates, & Pettit, 1998), ethnicity (Deater-Deckard et al., 1998), family socioeconomic status (SES; Dodge, Pettit, & Bates, 1994; Odgers et al., 2008), and other indicators of SES during adulthood including educational attainment and length of unemployment.

Aspects of parenting received in childhood included parental values toward aggression (Deater-Deckard et al., 1998), positive parenting (Deater-Deckard et al., 1998), parental involvement (Beyers, Bates, Pettit, & Dodge, 2003), parental monitoring (Beyers et al., 2003), interparental conflict (Buehler et al., 1997; Odgers et al., 2008), and exposure to violence (Dodge et al., 1994), harsh discipline (Dodge et al., 1994), spanking (Gershoff, Lansford, Sexton, Davis-Kean, & Sameroff, 2012), and physical harm (Deater-Deckard et al., 1998; Odgers et al., 2008).

Parental adjustment included mothers' and fathers' alcohol/drug use (Connell & Goodman, 2002) and arrests (Dallaire & Wilson, 2010; Odgers et al., 2008). Peer influences included social preference with peers (Keiley, Lofthouse, Bates, Dodge, & Pettit, 2003) and peer deviance (Goodnight, Bates, Newman, Dodge, & Pettit, 2006; Odgers et al., 2008).

Child characteristics included intelligence (Nigg & Huang-Pollock, 2003), social information processing (Dodge, Pettit, Bates, & Valente, 1995), reward sensitivity (Goodnight et al., 2006), internalizing problems (Keiley et al., 2000), language ability (Petersen et al., 2013; Petersen, Bates, & Staples, in press), and aspects of temperament (Keiley, Bates, Dodge, & Pettit, 2001). Stress included individual stress (Kim, Conger, Elder, & Lorenz, 2003), family stress (Deater-Deckard et al., 1998), and sleep problems (Goodnight, Bates, Staples, Pettit, & Dodge, 2007). Pregnancy risks for the target child included medical complications (Deater-Deckard et al., 1998), and having been born to a teenage mother (Wakschlag et al., 2000) or from an unplanned pregnancy (Hayatbakhsh et al., 2011). Family background characteristics included the ratio of children to adults in the home (Deater-Deckard et al., 1998), and whether or not the mother was a single mother (Ackerman, D'Eramo, Umylny, Schultz, & Izard, 2001), the father was low in caregiving (Mott, Kowaleski-Jones, & Menaghan, 1997), the mother was cohabiting with a non-marital partner (Ackerman et al., 2001), the parents divorced (Lansford et al., 2006), or the individual himself or herself divorced his or her spouse. Other experiences of interest

included the amount of television watched (Manganello & Taylor, 2009) and amount of non-maternal childcare (Bates et al., 1994; Deater-Deckard et al., 1998).

The Present Study

We sought to describe developmental profiles of externalizing problems and to predict them using an actuarial model of risk and protective factors. In other words, given inputs for each child (various child and environmental characteristics), we predicted his or her most likely output (i.e., trajectory of externalizing problems from childhood to adulthood). To do this, we combined different sources of ratings (mother-, teacher-, father-, peer- and self-reports) and different scales of externalizing problems, capitalizing on all of the available information to create a more robust externalizing profile. To render the ratings from different sources and scales conceptually and empirically comparable while still retaining mean-level change to observe meaningful change over time, we used the POM proportional scoring metric along with additional theoretical and empirical considerations. Modeling the trajectories of externalizing problems from childhood to adulthood while taking into account the heterotypic continuity of externalizing behavior allowed us to better understand 1) the patterns of developmental change in externalizing problems over a long span of development, 2) the risk and protective factors that predict the development of externalizing problems, and 3) the ways risk factors in different developmental periods contribute to the development of externalizing. We examined the effects of 40 different risk or protective factors from 9 domains across 6 time frames for a total of 66 risk or protective factors.

We expected that we would be able to model individual differences in trajectories across the long span of development based on findings from previous studies examining shorter spans. However, given the span of development covered and the few previous studies covering as much development, we could not be certain that our model would succeed. We also expected that some of the risk factors that have been associated with externalizing problems in prior studies would be associated with the trajectories of externalizing problems in the present study. Because risk factors may account for common variance in externalizing problems, we expected that some of the risk factors would not have unique effects when controlling for other risk factors. However, again studies with such large arrays of predictors over such lengths of development are few, and theoretical models cannot yet yield definitive hypotheses in such a set of variables, so we could not predict which variables would have the strongest links to externalizing behavior trajectories.

To further probe the meanings of the trajectories of externalizing behavior problems, we used individuals' predicted values of externalizing problems as predictors of several, particular, socially important illegal behavior outcomes. We also modeled the combinations of risk factors that were most predictive of one key outcome, having been arrested.

Method

Participants

Children (N = 585) were recruited for the Child Development Project (Dodge, Bates, & Pettit, 1990) from two cohorts in 1987 and 1988 from three sites: Nashville, Tennessee;

Knoxville, Tennessee; and Bloomington, Indiana. Children's parents were approached at random during kindergarten preregistration, on the first day of class, and by phone or mail. About 75% of parents approached agreed to participate. The schools and the composite sample reflected a broad range of socioeconomic status groups that were representative of the populations at the respective sites. The Hollingshead index of SES (M = 39.53, SD = 14.01) ranged from 8 to 66 for the original sample (reflecting a broad range), which was 52% male, 81% European American, 17% African American, and 2% of "other" ethnicity.

Measures

For a table of measures and at what ages they were collected, see Table 1. Rates of missingness for each variable are presented in Table 2.

Externalizing problems—Externalizing problems were measured by the Externalizing subscales on the developmentally-relevant Achenbach scales annually from ages 5–27 (except ages 18, 25, and 26): Mothers' scores came from the relevant factor of the Child Behavior Checklist (CBCL, 33 items; Achenbach, 1991a) from ages 5–17; teachers' scores from the Teacher Report Form (TRF, 34 items; Achenbach, 1991b) from ages 5–13; fathers' from the CBCL from ages 5–9 (only for cohort 1 at age 9); self-reports from the Youth Self-Report (YSR, 30 items; Achenbach, 1991c) at ages 12, 15–17, and 19, and the Young Adult Self-Report (YASR, 28 items; Achenbach, 1997) from ages 20–24 and 27; peer reports from the YASR at age 27. In each of the scales, reporters rated whether a given behavior was "not true," "somewhat or sometimes true," or "very or often true" (scored 0, 1, and 2, respectively).

Although the different Externalizing scales were obtained (from different raters at different ages) and the scales included different numbers of items, the Externalizing subscales shared many overlapping items across the various Achenbach forms. As a result of the different numbers of items, the Externalizing scores were adjusted across scale as a function of the number of items included in each version of the scale. We calculated a proportion of maximum (POM) score for each participant at each age on the different scales available. For example, the CBCL Externalizing subscale includes 33 items with a total possible score of 66. To calculate a POM on the CBCL, a participant's score was summed across items and then was divided by the total possible score (66), and finally multiplied by 100 for interpretability. The resulting metric corresponds to a person's proportion out of the highest possible score on a 0–100 metric. The POM scores were then averaged within year across raters to form a composite score of externalizing problems at each age for the growth curve analysis. The percentage of participants with scores on externalizing problems at different numbers of time points is in Appendix S1.

Cronbach's alpha of externalizing problems ranged from .88 to .92 for mothers on the CBCL, .83 to .89 for fathers on the CBCL, .94 to .96 for the teachers on the TRF, .84 to .92 for self-reports on the YSR, and .83 to .91 for self-reports on the YASR, depending on the year, and was .89 for peers on the YASR. Correlations across raters within-year ranged from .24 to .71 (M = .40, $ps < .001$), depending on the year measured. Correlations within type of rater across years ranged from .34 to .77 for mothers (M = .61), .29 to .63 for

teachers (M = .50), .58 to .79 for fathers (M = .68), and .28 to .81 for self-reports (M = .56; ps < .001) depending on the year. Correlations, means, and standard deviations of the POM-rescaled externalizing problems are in Table 3. Averages of externalizing problems over time by time frame are depicted in Figure 1.

In addition to the ratings of externalizing problems, illegal behavior was examined to verify the model predictions against particularly important and costly societal outcomes, including arrests, injecting illegal drugs, illegal drug use, and drunk driving. At age 27, adults reported whether they had ever been arrested, and 138 people (30%) reported that they had been arrested. At age 27, adults also reported whether they had ever injected an illegal drug (2%), used recreational drugs for non-medical purposes in the past 6 months (excluding alcohol and nicotine) (26%), and driven after drinking 5 or more alcoholic drinks in a row in the past year (15%).

Predictors

Time frames: Some of the predictors were measured at only one time point, whereas other predictors were measured at multiple time points. In order to make inferences about the developmental timing of various risk factors, we split up predictors according to the following time frames: a) time-invariant, b) early childhood (0–5 years of age), c) middle childhood (6–10), d) earlier adolescence (11–14), e) later adolescence (15–18), and f) adulthood (19–27). Thus, if a predictor was measured at multiple time points, we either: 1) computed an average score for the predictor within the time frame for whichever time frames data were available (e.g., peer deviance in earlier adolescence versus later adolescence), or 2) used a categorical index of whether the risk factor occurred during the time frame of interest (e.g., divorce in early childhood versus middle childhood). For predictors with multiple raters, the information was first combined across raters within year and then consolidated within time frame. For a correlation matrix of the predictors and their means and standard deviations, see Appendix S2.

Demographics: Child *sex* was measured at age 5 as male=0, female=1. Children's *ethnicity* was reported at age 5 as European American, African American, or "other." Ethnicity was dummy coded into two variables (variable name in italics): 1) *African American*=1, European American=0, other ethnicity=0 and 2) *other ethnicity*=1, European American=0, African American=0. *Socioeconomic status* (SES) was measured by the Hollingshead four-factor index (Hollingshead, 1975) when the children were 5 years old, based on parents' education and occupational status. Because prior studies have shown that the effect of SES on externalizing problems is accounted for by more proximal aspects of socialization (Dodge et al., 1994), we partialed out the variance of specific early childhood risk factors, e.g., the child's exposure to violence, harsh discipline, and positive parenting, from SES in order to determine how (i.e., by which processes) SES affects risk for externalizing problems. SES was residualized for the collective risk models by regressing SES on the other risk factors at age 5 that were significant predictors of SES and saving the residuals to render the residualized SES term independent of the other early childhood risk factors.

Educational attainment was measured as the target individual's highest level of education attained as of age 27 on a scale of 1–10 (1=8th grade or lower … 10=more than four years of school beyond high school). *Length of unemployment* was measured at age 27 as the total length of unemployment, in months, since high school (while not in school or at home having a baby). The length of unemployment also included time in jail or prison if he or she did not work.

Parenting: When the child was 5 years old, mothers rated their *attitudes toward aggression* (Culture Questionnaire; Dodge et al., 1994) on 15 items with a 7-point scale ranging from 1="definitely disagree" to 7="definitely agree." Items included "If my child gets into a fight with another child, I won't try to stop it because my child has to show that she/he can defend herself/himself," and "I let my child watch adventure television shows that have killing and violence in them." Five items were reverse-scored to reduce response bias. Internal consistency (α = .57) was low but acceptable for the present purpose.

Positive parenting was measured by maternal report in an interview of her discipline strategies in various hypothetical situations of child misbehavior (Concerns and Constraints Questionnaire; Pettit, Bates, & Dodge, 1997). Five child misbehavior vignettes were presented (e.g., aggression, name-calling, teasing others), and parents were asked how they could prevent their child from acting this way in the first place. Responses were coded on a 5-point scale (1="do nothing, is unpreventable" to 5="preventive, anticipatory, situation specific"). Internal consistency was .63.

At home when their child was 5 years old, interviewers asked parents about parental conflict, discipline practices, exposure to violence, and possible child abuse across two time frames: 1–4 and 4–5 years old. After discussing these issues, interviewers made ratings about physical harm, parental conflict, exposure to violence, and discipline (for information on interrater reliability, see Deater-Deckard et al., 1998). Interviewers rated the likelihood that the child had been *physically harmed* (with codes ranging from "definitely not" to "authorities involved"). Children were classified as physically harmed if the interviewer rated the likelihood of physical harm as probable or as having involved child welfare authorities in either time frame.

Interviewers rated the amount of *parental conflict* and violence during early childhood on a 5-point scale (1="rarely even shout" to 5="physical, more than once"). For both cohorts, interviewers rated the child's amount of *exposure to violence* on a 5-point scale (1="none" to 5="physical, more than once"). For cohort 2 only, interviewers rated the child's level of exposure to violence inside and outside the home separately. For cohort 2, the level of exposure was averaged across the ratings from within and outside the home. Exposure to violence was averaged across the two early-childhood time frames. Interviewers also rated the parents' use of harsh, punitive, and restrictive discipline on a 5-point scale (1="non-restrictive, mostly positive guidance" to 5="severe, strict, often physical discipline"). *Harsh discipline* was averaged across the two early-childhood time frames.

On the Conflict Tactics Scale (Straus, 1979) at age 5, mothers rated how often in the past year and in the first four years of the child's life they and their spouses 1) threatened to

spank their child, 2) spanked their child, and 3) spanked their child with something. Frequencies ranged from never (0) to almost every day (6). *Spanking* scores were averaged within parent then across parents, and then across time frames, and were only available for cohort 2. Alphas ranged from .76 to .80, depending on the year. Correlations between parents' spanking ranged from .69 to .75 ($p < .001$).

At ages 12, 14, 15, and 16, youths rated how well their parents monitored them by how well their parents knew where they went after school, who their friends were, how they spent their money, where they went at night, and how they spent their free time. Ratings were 1=the parents "do not know", 2="know a little", and 3="know a lot." At ages 12, 14, and 15, *parents' monitoring* ratings were made for parents as a unit. At age 16 ratings were made for each parent separately, and then were averaged together. Internal consistency ranged from .74 to .99, depending on the year, except at age 12, in which it was .42. Parental monitoring was averaged within time frame.

At ages 13 and 16, we collected measures of the *parents' involvement* with the youth in terms of the average amount of time in hours over the course of a week that the parents: 1) sit around and talk with their teen and 2) spend time with their teen doing things that the teen enjoys. Youths' ratings were made at age 16, and mothers' ratings were made at ages 13 and 16 (at age 16, mothers only reported on their involvement for cohort 1). Correlations between the two items ranged from .36 to .41 ($p < .001$), depending on the year and reporter. The correlation between youths' and mothers' ratings at age 16 was .28 ($p < .001$).

Parent Adjustment: *Mothers' and fathers' drug problems* were reported by each parent or, in cases where the child's mother or father was unavailable, by another primary caregiver that was knowledgeable about the parents' mental health. Drug problems were reported in the context of a Family History Epidemiologic interview (Lish, Weissman, Adams, Hoven, & Bird, 1995) at age 16. Each parent was scored as having an alcohol/drug problem if he or she had been reported by the self or other caregiver as having ever: 1) been hospitalized for drugs or alcohol, 2) drunk a lot, 3) had drinking problems, 4) used illegal drugs, or 5) had a drug problem. *Parents' arrest*s were also reported by mothers and fathers, or, in some cases, by another caregiver. The parent was considered as having been arrested if the mother, father, or other caregiver reported that he or she had ever 1) been put in jail, 2) arrested, or 3) convicted of any crime (other than drunk driving or traffic violations).

Peers: *Peer deviance* was measured by the child's report of friends' deviant behavior on a questionnaire at ages 11, 12, 14, 15, and 16. We selected the common items assessed across ages to include in the peer deviance composite, which included asking the youth how often his or her friends: 1) steal things (from stores), 2) get into fights with other kids, 3) smoke cigarettes, 4) lie to parents/teachers, 5) get into trouble at school, 6) suggest that he or she do something illegal, 7) use bad language, and 8) do things that make him or her scared or uncomfortable. At ages 12, 14, 15, and 16, ratings were made about the youth's friends in general. At age 11, the youth rated the deviance of the two children with whom they spent the most time. Ratings at age 11 were averaged across the two friends, and the scale of 1–3 at age 11 (1=never, 2=sometimes, 3=very often) was re-scored according to the corresponding levels of the 1–5 scale used at the other ages (1=never, 2=once in a while,

3=sometimes,4=fairly often, 5=very often) (i.e., 1 = 1, 2 = 3, 3 = 5). Internal consistency ranged from .76 to .83. Peer deviance was averaged within time frame.

Social preference was measured by peer sociometric interviews from ages 5–9 (Keiley et al., 2003). Children nominated up to three classmates they especially liked and up to three they especially disliked. The frequencies of liking and disliking were summed for the target child and standardized within their classroom. Social preference was calculated by subtracting the child's disliking score from his or her liking score, and scores were averaged within time frame. Social preference scores at age 9 were only available for cohort 1.

Child Characteristics: The child's *intelligence* was measured at age 13 by the average of his or her scaled scores ($M = 10$, $SD = 3$) on the Block Design and Vocabulary subtests of the Wechsler Intelligence Scale-Revised (Wechsler, 1974). The child's *social information processing* (SIP) was measured annually from ages 5–8 by responses to cartoon pictures and 24 video vignettes, which depicted child protagonists attempting unsuccessfully to enter peer groups and encountering provocation. After each video, children were instructed to pretend they were the protagonist, and responded to questions to assess their four steps of processing: 1) encoding, 2) attributions, 3) response generation, and 4) response evaluation, with higher values representing encoding deficits, hostile attributions, aggressive response generation, and aggressive response evaluation, respectively. The composite SIP variable for a given year represented the proportion of four SIP steps on which the child scored 1 *SD* above the mean or greater. Internal consistency for each of the four steps was strong at each age (Lansford, Malone, Dodge, Pettit, & Bates, 2010).

At age 5, mothers reported on their child's temperament retrospectively during infancy on the Retrospective Infant Characteristics Questionnaire (Bates & Bayles, 1984; Bates, Pettit, Dodge, & Ridge, 1998). Three dimensions of temperament were examined: 1) difficultness, 2) unadaptability, and 3) resistance to control. *Difficult temperament* ($\alpha = .86$) was measured by 9 items related to the child's negative emotionality (e.g., how easily upset, how often fussing/crying). *Unadaptable temperament* ($\alpha = .72$) was measured by 4 items related to negative reactions to novelty (new food, people, places, and adaptation in general). *Resistance to control* ($\alpha = .83$) was measured by how often the child 1) persists in playing with objects when told to leave them alone, 2) continues to go someplace even when told to stop, and 3) gets upset when removed from something he or she is interested in but should not be getting into. Ratings on each item ranged from 1–7, with higher values representing less optimal temperament traits (more difficultness, unadaptability, and resistance to control). Correlations between mothers' and fathers' ratings were .47, .34, and .37 ($p < .001$) for difficult, unadapatable, and resistance to control temperaments, respectively.

Reward sensitivity was measured at age 16 using a computerized card playing task (Goodnight et al., 2006), as adapted from J. P. Newman, Patterson, and Kosson (1987) and Siegel (1978). In the task, participants played cards to win money (gaining 25 cents for turning over face cards but losing 25 cents when turning over number cards). Participants could stop playing at any point, up to 100 cards, and keep their winnings. The probabilities were structured to set the optimal number of cards played at 50 to maximize winnings. After

playing 100 cards, however, all winnings were lost. Reward sensitivity was measured as the number of cards played.

Internalizing problems were measured by the Internalizing subscale of the Achenbach scales with the same raters and scales as were used for externalizing problems, and were POM-rescaled. Cronbach's alpha ranged from .81 to .90 for mothers, .82 to .84 for fathers, .85 to .91 for the teachers, .88 to .91 for self-reports on the YSR, and .89 to .92 for self-reports on the YASR, depending on the year, and was .89 for peers on the YASR. Correlations within year across raters ranged from .06 to .51 (p = .144 to < .001), depending on the year. Correlations within rater across years ranged from .34 to .79 for mothers, .29 to .63 for teachers, .58 to .79 for fathers, and .28 to .81 for self-reports (p < .001), depending on the year. Internalizing problems were averaged within developmental time frames (as listed in Table 1).

Language ability was measured as the child's percentile score on the composite language section of a nationally-normed standardized test, which was collected annually via official school records. A school records form with achievement test scores for the participants was completed by a school administrator. The school records were collected at the end of the school year in the summer, but the standardized tests were administered during the school year. School records from ages 7–10 were collected when the children were 10 years old, and school records from ages 11–13 were collected in the summer after each school year. The composite language ability score reflected two types of subtests, including language mechanics and language expression. Correlations between language mechanics and expression scores ranged from .59 to .71 (p < .001) depending on the year. Language ability was averaged within time frame.

Stress: *Individual stress* was reported by adolescents on the Changes and Adjustments Questionnaire (CAQ; Dodge et al., 1994) at ages 14, 15, and 17. The adolescent reported whether they experienced each of 27 possible stressful life events in the prior year, including events such as "moved," "serious illness or accident," "close family member died," and "money problems." The number of stressful life events was summed within year. Individual stress was also reported by the target adult at age 26 on a list of 18 stressful life events in the past year. Individual stress was averaged within time frame.

Family stress was reported by mothers on the CAQ at ages 5–17. At age 5, mothers reported on the family stressors in the following time frames of the child's life: 0–1, 1–4, and 4–5 years old. At age 5, the mother reported whether the family experienced each of 15 possible stressful life events in the prior year, including events such as legal problems, job loss, and financial instability. At ages 6–17, the list included 18 possible stressors. The number of stressful life events was summed within year and then averaged within time frame.

The child's *sleep problems* were measured at ages 5–17 as the average of mothers' reports on 3 CBCL items: 1) trouble sleeping, 2) sleeps less than others, and 3) overtired. Internal consistency was low (.34 to .58, depending on the year) but acceptable for the present purpose. Sleep problems were averaged within time frame.

Pregnancy: Mothers reported whether there were *medical complications* during pregnancy, birth, and the first few months of the child's life, and interviewers coded the response (1=healthy, 2=minor problems, 3=major problems). The mother's pregnancy was considered a *teenage pregnancy* if the mother gave birth to the target child when she was 18 years old or younger (coded as "1"). If the mother was 19 years old or older when she gave birth, it was coded as a "0." When the child was 5 years old, mothers reported whether the target child was born to an *unplanned pregnancy*, and the interviewer coded the response as "0" if the pregnancy was planned, discussed, or accepted, and "1" if unplanned.

Family Background: In the mother interview when the child was 5 years old, mothers were asked who the main caregivers of the child were, and how much time per week the child spent with each in two different time frames: ages 1–4 and 4–5. Responses were then coded as to how much time per week was spent with the father in each of the time frames on a 5-point scale (1="not in this type of care" to 5="major; more than 20 hours per week for more than 18 months"). *Low father caregiving* was scored as a "1" if the child was considered not in the care of the father. If the father exhibited brief, moderate, frequent, or major care with the child, the father's caregiving was not considered low (scored as a "0"). From ages 6–9, father caregiving was reported by the mother. The number of hours per week that the father spent with the child in the prior year was rated on a 5-point scale (1="occasionally or none" to 5="30 or more"). From ages 6–9, father caregiving was considered low if their caregiving was reported to be occasional or absent (i.e., less than 1 hour per week). Fathers' caregiving was considered low within the time frame if they met the criterion for low caregiving within any year of the given time frame.

The mother reported her marital status when the child was 5 years old. She was not considered a *single mother* (coded as "0") if she reported that she was married, living with a partner, or living with another adult, whereas she was coded as "1" if she reported that she was single and living alone. She was not considered *cohabiting* (coded as "0") if she reported that she was married or single and living alone, whereas she was coded as "1" if she reported that she was not married and was living with someone else.

Parents' *divorce* was rated as part of the CAQ. The mother reported whether she had divorced or separated from her partner in the prior year. At age 5, mothers reported whether a divorce occurred in the following time frames of the child's life: 0–1, 1–4, and 4–5 years old. From ages 5–17, mothers reported annually whether a divorce occurred in the prior year. At age 26, the target participant reported whether he or she had divorced from a spouse in the past year. If no new divorce or separation occurred within the time frame, it was scored as a "0," whereas it was scored as a "1" if the participant (age 26) or the participant's mother (age 0–17) divorced or separated within the time frame. The *child to adult ratio* was calculated as the number of children in the household divided by the number of caregivers when the child was 5 years old.

Child Activities: The child's *amount of television watched* at age 5 was calculated as the average number of hours of television watched alone per day 1) during the week and 2) on weekends on a three point scale: 1=1 hour, 2=2–3 hours, 3=4 or more hours. The correlation between weekday and weekend was .46 ($p < .001$). *Amount of non-parental childcare* from

birth to 1 year of age, 1–4, and 4–5 years of age was reported retrospectively by mothers in an interview when the child was 5 years old. Parents' responses concerning birth to 1 year of age were coded on a 7-point scale (0=none to 6=more than 30 hours per week for at least 7 months). From 1–4 and 4–5 years old, parents reported the amount of time that children spent in care outside the home in the following settings: a relative's residence, small group babysitter, group daycare, preschool, neighbors/friends, or other. Times were scored on a 5-point scale ranging from 0="not in this type of care" to 4="major care, more than 20 hours per week." Childcare scores were summed across the types of care within age range. Childcare scores from the different ages were standardized with a *z*-score before averaging across time frames. For coding reliability, see Bates et al. (1994).

Statistical Analysis: To model growth curves of externalizing problems from ages 5–27, we used the lme function of the nlme package (Pinheiro, Bates, DebRoy, & Sarkar, 2009) in R 3.0 (R Development Core Team, 2009) for hierarchical linear modeling (HLM). Because our aim was to predict risk for externalizing problems in adulthood, we set the intercept at age 27 rather than at age 5, consistent with other studies investigating externalizing trajectories (e.g., Owens & Shaw, 2003). Various curvilinear forms of growth were compared. After settling on a form of growth, we related the risk factors individually to the growth curves and then collectively (similar to the approach by Owens & Shaw, 2003), taking a best predictors approach.

After identifying all of the risk factors that were individually associated with the intercepts or slopes of externalizing trajectories, we combined the risk factors in one model. To avoid systematic bias in model parameter estimates and inferences, we used multiple imputation, which is preferable in developmental studies when there is missingness (Jeličić, Phelps, & Lerner, 2009). We multiply imputed 20 data sets using Amelia II version 1.6.3 (Honaker, King, & Blackwell, 2011) in R to have adequate power (i.e., power falloff of about 1% with respect to full information maximum likelihood estimates) when missingness is between 10–50% (most of the variables in the present study) (Graham, Olchowski, & Gilreath, 2007). Amelia uses an expectation-maximization with bootstrapping algorithm, and is well suited for longitudinal data (Honaker & King, 2010). For accurate imputations, we imputed the data with a cubic polynomial to account for the effects of time over a long time span (23 years). We examined imputation diagnostics, including 1) comparing the descriptives and distributions of observed and imputed data, 2) overimputation (sequentially removing and imputing observed values as if they had actually been missing values), 3) using overdispersed starting values (convergence in the imputations from different starting values), and 4) examining time series to ensure the imputed data fell within the participants' general trends. Diagnostics suggested that the imputed data were acceptable. The conditional multilevel models were run on each imputed data set separately, and then the results were combined using the mitools (Lumley, 2010) and mix (Schafer, 1997) packages in R, which use Rubin's (1987) rules for combining results of analyses on multiply imputed data sets.

The risk factors were examined collectively via forward selection in HLM growth curves. We used forward selection because it tends to be more accurate and conservative than backward elimination in selecting predictors (Derksen & Keselman, 1992). The stepAIC function of the MASS (Venables & Ripley, 2002) package in R determined the best set of

predictors by selecting iteratively only those predictors that incrementally improved model fit, as measured by Akaike's Information Criterion (AIC). The AIC balances the goodness of fit with the complexity of the model, by penalizing models with more predictors. The typical penalty for AIC is 2 times the number of parameters (Sheather, 2009), whereas we set the penalty to 4 times the number of parameters for a more conservative threshold for selecting predictors (Venables & Ripley, 2002). We kept a predictor if it was selected by forward selection in at least half of the imputed data sets (10/20). First, we selected predictors of the intercepts. Second, in a separate model, we selected predictors of the slopes. Third, we combined the predictors of the intercepts and slopes to select the best set of predictors. Finally, non-significant predictors of the intercepts and slopes were removed to retain only significant predictors in the final model.

The predictions from the final model were then tested on illegal behavior, including arrests. Because the illegal behaviors were reported as binary (i.e., whether a behavior did or did not occur), we examined them in the context of receiver operating characteristic (ROC) curves, which examine the diagnostic utility of a given assessment tool by evaluating the tradeoff between its sensitivity and specificity to predict the outcome. ROC curves were estimated using the ROCR package (Sing, Sander, Beerenwinkel, & Lengauer, 2005) in R. All of the descriptive statistics (means, standard deviations, and Pearson correlations) and the unconditional multilevel models are from the raw, non-imputed data set.

Analysis of Missingness: The number of time points that participants had scores for externalizing problems was positively associated with SES (r[568], = .21, $p < .001$, two-tailed). The number of time points that participants had scores for externalizing problems was not significantly related to their ending values of externalizing problems ($B = -0.08$, $p = .112$), but was related to their linear slopes of externalizing problems ($B = 0.01$, $p = .006$). Compared to participants with more time points of externalizing problems, participants with fewer time points had higher initial values of externalizing, yet smaller increases in externalizing problems over time. In other words, participants who dropped out of the study had more externalizing problems at the earliest ages than non-dropouts. The relation of missingness to SES and externalizing problems highlights the importance of conducting multiple imputation with these variables to help explain the pattern of missing data. The two cohorts did not significantly differ in terms of SES (t[558.82] = 1.60, $p = .110$) or in ending values ($B = 1.25$, $p = .216$) or linear slopes ($B = 0.04$, $p = .535$) of externalizing problems.

Results

Construct Validity Invariance

We examined whether externalizing problems showed construct validity invariance over time by examining the convergent validity of externalizing and its discriminant validity with respect to internalizing problems. Ratings of externalizing problems and internalizing problems were divided into 3 different blocks according to the primary raters within the era: block 1 = ages 5–13 (parent- and teacher-report), block 2 = ages 14–17 (parent- and self-report), block 3 = ages 19–27 (self- and peer-report). POM-rescaled ratings were averaged

across years within a given block. We tested whether externalizing problems predicted later externalizing problems more strongly than later internalizing problems.

Externalizing problems showed convergent validity across time from block 1 to block 2 (r = .65, p < .001) and from block 2 to block 3 (r = .62, p < .001). Although externalizing problems predicted later internalizing problems from block 1 to block 2 (r = .32, p < .001) and from block 2 to block 3 (r = .37, p < .001), the associations were stronger from externalizing problems to later externalizing problems than to later internalizing problems from block 1 to 2 (Fisher's r-to-z = 7.00, p < .001) and from block 2 to 3 (z = 5.04, p < .001). Thus, there was cross-time convergent and discriminant validity for externalizing and internalizing problems across all three blocks.

Describing Growth Curves of Externalizing

An unconditional means model with random intercepts was fit to the trajectories of externalizing problems and showed considerable within-person (σ^2_ε = 61.26, SD = 7.83) and between-person (σ^2_0 = 56.33, SD = 7.51) variance, suggesting that the average person varies over time, and that the cross-time means of externalizing problems differ between individuals.[1] Moreover, the intraclass correlation was ρ = .48, suggesting that about half of the variability in externalizing problems is between individuals, and that externalizing problems have a high residual autocorrelation over time.

To account for the change in externalizing problems over time, an unconditional growth model was fit with random intercepts and a linear random slope for time (a random intercepts and slopes model). The unconditional growth model was a better fitting model than the unconditional means model (χ^2[3] = 1382.63, p < .001), suggesting that externalizing problems change over time. Moreover, the model with a random effect of time fit better than a model with a fixed effect of time (χ^2[2] = 1300.16, p < .001), suggesting that trajectories of externalizing problems differed between individuals.

Curvilinear forms of change were examined. Quadratic forms of change were significantly better fitting than a linear model (χ^2[4] = 208.85, p < .001). Despite the modest variance in the quadratic curvature (σ^2_2 = 0.002, SD = 0.05), the model that allowed the curvature to vary across individuals fit better than the model with a fixed quadratic curvature (χ^2[3] = 161.88, p < .001). There was not adequate variability across individuals in the cubic curves for cubic models to converge, so we examined subsequent polynomials with fixed effects. Models with a fixed cubic term fit better than models without the cubic term (χ^2[1] = 21.26, p < .001). Moreover, models with a fixed quartic term fit better than models without the quartic term (χ^2[1] = 120.22, p < .001). Models with a quintic term did not fit significantly better than models without a quintic term (χ^2[1] = 0.64, p = .423), so we chose the simpler quartic model for parsimony. To prevent over-fitting, we split the sample into two random subsets of cases and examined the quartic model with each subset. The quartic model was the best fitting model for each subset. Thus, subsequent growth models examined

[1]Notation for variance components follows the convention in the HLM literature (Snijders & Bosker, 2011).

trajectories of externalizing problems with random intercepts, linear slopes, and quadratic curvatures, along with fixed cubic and quartic effects.

For a plot of the average quartic trajectory of externalizing problems overlaid with the means of externalizing problems over time, see Figure 1. The means of externalizing problems showed decreases from ages 5–11, followed by increases from 11–16, and decreases from 16–27. For a plot of individuals' quartic trajectories of externalizing problems, see Figure 2. Although the average trajectory is fairly flat, the individual trajectories show considerable variability, both in intercepts, slopes, and curvatures, suggesting that the development of externalizing problems differs between people.

After the addition of the curvilinear effects of time, the proportional reduction in intercept variance (PRV, similar to ΔR^2; Peugh, 2010) was .49, suggesting that about half of the between-person differences in ending values of externalizing problems at age 27 was accounted for by the effects of time (i.e., linear, quadratic, cubic, and quartic slopes). Moreover, the positive correlations between the intercept and linear slope (r = .44) and the intercept and quadratic curvature (r = .09) suggest that the higher a person's slope and curvature, the higher his or her ending value of externalizing problems at age 27.

Predicting Growth Curves of Externalizing

Although we fit a quartic model, for easier interpretability, we only examined whether risk factors predicted the intercepts and linear slopes of externalizing problems because the interpretation of predictors of polynomial terms is notoriously difficult (Grimm, Ram, & Hamagami, 2011). Each predictor was tested separately in two models: 1) predicting the intercepts, and 2) predicting the intercepts and slopes. Parameter estimates for the predictors of the intercepts in model set 1 and predictors of the slopes in model set 2 are in Table 4. Any variables that significantly predicted the intercepts or slopes were included in the multiple imputation. The significant terms were then examined collectively by forward selection.

After imputation, SES was residualized by regressing SES on the other risk factors at age 5 that were significant predictors of SES, including positive parenting, single mother, divorce, and child to adult ratio, and saving the residuals. In separate models, residualized SES did not significantly predict the intercepts (B = −0.18, p = .145) or slopes (B = 0.01, p = .875) of externalizing problems. Then we examined the significant predictors of the intercepts and slopes collectively by forward selection.

First, forward selection was used to select predictors of the intercepts. Second, in a separate model, forward selection was used to select predictors of the slopes. Third, the predictors of the intercepts and slopes were then combined, and forward selection was used to select the set of best predictors. After forward selection of the combined set of predictors, in the final model we included variables that were, in the separate analyses, significant predictors (p < .05) of the intercepts or slopes in the final model. The parameter estimates from the final model are in Table 5. We present the analyses with externalizing problems in their raw metric of POM scores. The POM scores were not very normally distributed, so we tested the

model with externalizing scores square-root transformed. The predictors remained essentially the same.

Six variables predicted individuals' intercepts of externalizing problems. The following groups/predictors were associated with higher ending values of externalizing problems at age 27: males, peer deviance (earlier and later adolescence), individual stress (later adolescence), and internalizing problems (later adolescence and adulthood).

Nine variables predicted individuals' slopes of externalizing problems over time. To understand how each of these 9 risk factors was associated with changes in externalizing problems over time, we probed the effects with plots. We created separate plots to examine the effect of low (1 *SD* below the mean) and high (1 *SD* above the mean) values of each risk factor on the slopes of externalizing problems over time. Examining the plots showed that some predictors of the slopes were related to the initial values of externalizing problems. In these instances, the slopes converged over time for low and high levels of the risk factor. Because the slopes converged from different starting points, the effects of the risk factors on the slopes were detected as significant. The observation that the risk factors related to the initial values of externalizing problems was confirmed empirically by setting the intercepts to the initial rather than ending values in a separate model. The following predictors were characterized by higher initial values of externalizing, yet smaller increases or greater decreases in externalizing problems over time: resistant to control temperament (early childhood), spanking (early childhood), harsh discipline (early childhood), low father caregiving (early childhood), lower peer social preference (early and middle childhood), internalizing problems (middle childhood), and poorer language ability (earlier adolescence).

One interpretation for these effects on the slopes is that the risk factors did not have enduring effects. Alternatively, the slopes could reflect a self-righting characteristic of development (Kohlberg, LaCrosse, & Ricks, 1972), which could be due to efforts families and individuals make to rein in uncomfortably high levels of behavior problems, whatever their source, as in the study concerning parental campaigns of increased involvement and control of the child (Goodnight, Bates, Pettit, & Dodge, 2008). Or, more simply it could reflect a statistical law, regression to the mean, where elevations resulting from risk factors eventually returned to typical levels. Another risk factor was associated with the slopes in different ways. Higher peer deviance in later adolescence was associated with greater increases in externalizing problems over time compared to lower peer deviance. Re-centering the intercepts to different points in development (e.g., Muthén & Muthén, 2000) showed that the effect of peer deviance in later adolescence on slopes of externalizing problems became significant in later adolescence (age 15) and remained significant in adulthood.

Several patterns are worth noting. First, although SES was individually associated with the intercepts and slopes of externalizing problems (see Table 4), residualized SES was not a significant predictor of the ending values or slopes of externalizing problems when controlling for more proximal risk factors (positive parenting, single mother, divorce, and child to adult ratio). Second, males had higher ending values of externalizing problems than

females, but males and females did not differ in their slopes (see Table 4). Third, in the individual models, African Americans tended to show greater increases in externalizing problems over time than European Americans, but African Americans no longer had greater increases than European Americans in externalizing problems over time when other risk factors, such as stress and peer deviance, were controlled.

A pseudo-R^2 of the final model was calculated by examining the squared correlation between the model's fitted and observed values (Singer & Willett, 2003). The pseudo-R^2 for the final model was .70, suggesting that the model fit the data well and accounted for 70% of the variability in externalizing problems over time. Moreover, the proportional reduction in intercept variance with the addition of the risk and protective factors to the baseline quartic model was .06. Thus, the specific risk and protective factors accounted for an additional 6% of variability in the ending values above the effects of the linear, quadratic, cubic, and quartic terms.

Applying the Model to Predict Illegal Behavior

The model predictions were then used to predict illegal behavior including arrests. The fitted values of the final model were averaged across time and then across the multiple imputations. Therefore, the final fitted values represented the average level of predicted externalizing problems from ages 5–27. We only examined the predictions in relation to the observed values of the outcomes (the illegal behaviors were not imputed) to avoid overestimating the model's predictive ability. In ROC curves, the area under the curve (AUC) represents the probability that a randomly selected person meeting the diagnostic threshold (i.e., having been arrested) will have a higher test result (i.e., more externalizing problems) than a randomly selected person who does not meet the cutoff. The AUC represents the tradeoff between a test's sensitivity and specificity. Sensitivity is the likelihood of correctly identifying individuals meeting the diagnostic threshold (true positive rate or hits). Specificity is the likelihood of correctly identifying individuals not meeting the diagnostic threshold (true negative rate or correct rejections). In general, a higher AUC, sensitivity, and specificity represent a better performing diagnostic test (range: 0–1, chance = 0.5).

Predicting arrests, the predicted externalizing problems had an AUC of .78 (see Figure 3), indicating that the prediction was moderately accurate (Akobeng, 2007). The optimal cutoff was defined as the number of externalizing problems which maximized the sum of the test's sensitivity and specificity. The optimal cutoff for arrests was 12.6 externalizing problems, at which point the sensitivity was .73 and the specificity was .70. A POM score of 12.6 corresponds approximately to a sum score of 9 externalizing problems on the CBCL and TRF, 8 problems on the YSR, and 7 problems on the YASR (where every rating of 1 counts as 1 problem and ratings of 2 count as 2 problems).

See Table 6 for the accuracy of the model's predictions for the other illegal behaviors. Predictions of whether a person had ever been arrested and had injected illegal drugs came from the average level of predicted externalizing problems from ages 5–27. Model predictions for the other outcomes (clinical level of externalizing, illegal drug use, and drunk driving) were from the model's predicted values of externalizing problems at age 27 because

the outcomes occurred within the prior 6–12 months of reporting at age 27. The predictions had high accuracy for clinical levels of externalizing problems at age 27 (AUC = .99) and moderate accuracy for illegal drug use (AUC = .72), injecting illegal drugs (AUC = .82), and drunk driving (AUC = .71).

Risk Profiles Associated with High Risk of Arrest

Because arrests were fairly prevalent in the sample (30%) and the final model's predictions were fairly sensitive and specific in predicting arrests, we examined the combination of risk factors that resulted in the greatest risk of arrest. We used a conditional inference tree with the ctree function of the party package (Hothorn, Hornik, & Zeileis, 2006) in R to determine the most common risk profiles among those who had been arrested. Using a conditional inference tree to identify the risk profiles associated with arrest may improve classification of those at greatest risk of arrest, which may lead to targeted, cost-effective interventions.

The conditional inference tree recursively estimated the association between risk factors from the final model and risk of arrest. First, the model selected the risk factor with the strongest association with arrest. Second, the model used a binary split on this risk factor at the cutpoint that maximized the discrepancy between the risk of arrest among the two subsamples (above and below the cutpoint). The model recursively repeated these steps with the next strongest predictor until the stop criterion, based on Bonferroni-adjusted *p*-values, was met to prevent overfitting. The results of the conditional inference tree are depicted in Figure 4.

The risk factor with the strongest association with arrest was peer deviance in later adolescence. The next strongest predictor of arrest was sex. Among females, their risk of arrest depended on individual stress in later adolescence. High individual stress, however, was not associated with particularly high risk of arrest among females who did not have high peer deviance (25% or less were arrested regardless of their levels of individual stress). Among males, on the other hand, risk of arrest was strongly conditional on individual stress in later adolescence. Two risk profiles associated with particularly high risk of arrest were: 1) high peer deviance in later adolescence (above the 89th percentile) and 2) males with high individual stress during later adolescence (above the 65th percentile). The first risk profile, high peer deviance, included 51 individuals of whom 67% were arrested. There were 59 males who had high individual stress in later adolescence, of whom 54% were arrested.

Discussion

The present study sought to describe and predict developmental profiles of externalizing problems longitudinally from childhood to adulthood using a developmentally-informed actuarial approach. Findings suggested that, on average, externalizing problems decreased from early childhood to preadolescence (ages 5–11), increased during adolescence (11–16), and decreased again from late adolescence to adulthood (16–27). There was considerable variability in the developmental trajectories of externalizing problems. We were best able to account for individuals' trajectories with a quartic function. Further, many individual risk factors predicted the ending values at age 27 (intercepts) or the change over time (slopes) in externalizing problems. This affirms the theoretical and empirical basis for our selection of

the risk variables. In a very broad sense, the findings replicate the prior research on factors in externalizing behavior problems. However, of course it would be expected that these risk factors have some degree of overlap. And in fact, when we combined the predictors into one model to test collective risk, fewer predictors remained associated with the ending values or slopes of externalizing problems, suggesting that the risk and protective factors accounted for overlapping variance in externalizing problems.

Modeling the risk variables together allowed us to examine the unique contributions of individual risks when taking into account numerous other risk factors. The specific variables uniquely associated with higher ending values of externalizing problems at age 27 included male sex, peer deviance in early adolescence and later adolescence, individual stress in later adolescence, and internalizing problems in later adolescence and adulthood. Other variables were uniquely associated with the initial (age 5) levels of problems and with the slopes from ages 5 to 27. Temperamental resistance to control in early childhood, parents' spanking in early childhood, parents' harsh discipline in early childhood, low father caregiving in early childhood, lower peer social preference in early and middle childhood, internalizing problems in middle childhood, and lower language ability in earlier adolescence were characterized by higher initial values of externalizing yet smaller increases or greater decreases in externalizing problems over time, resulting in slopes for low and high levels of the risk factors that converged over time. Higher peer deviance in later adolescence was associated with greater increases in externalizing problems over time compared to lower peer deviance. Thus, the risk and protective factors provided incremental prediction across a wide range of ages and domains.

There were also notable non-predictors of the development of externalizing problems. First, although males had higher ending values than did females, males and females did not differ in their slopes. Second, SES did not predict the ending values or slopes of externalizing problems when controlling for more proximal risk factors (positive parenting, single mother, divorce, child to adult ratio), suggesting that we were able to account for the commonly observed effect of SES with more proximal risk variables. This study thus provides new evidence of the operative mechanisms in the association between family SES and children's development of adjustment. Third, although African Americans showed greater average increases in externalizing problems over time compared to European Americans, African Americans did not have greater increases when controlling for other risk factors, suggesting that we were able to account for ethnic differences in trajectories with other risk factors.

We examined the associations between risk factors and externalizing problems in a) bivariate models that did not include other control variables and in b) multivariate models that controlled for other variables to identify the independent effects of the risk factors. There were three different patterns of associations and non-associations of the risk factors with the intercepts and slopes of externalizing problems in the bivariate models and multivariate models. Of the 66 variables, eight were not associated with the intercepts or slopes of externalizing problems in the bivariate or multivariate models: other ethnicity, parental involvement in earlier or later adolescence, parental monitoring in later adolescence, divorce in adulthood, and unadaptable temperament, medical complications, and teenage pregnancy in early childhood. These risk factors have been found to be related

to externalizing problems in some prior studies. We did not observe an association, however, and we are not sure if these risk factors have had consistent associations with externalizing problems in all prior studies. Nevertheless, differences in findings may owe to methodological differences when examining growth curves from ages 5 to 27.

Of the remaining 58, there were also 44 variables that were associated with either the intercepts or slopes of externalizing problems in the bivariate models, but not in the multivariate models. In these cases, the independent associations of the risk factors were too weak to be detected and their effects were statistically accounted for by other variables. Observing significant predictors in the bivariate but not multivariate models reflects collinearity among risk factors, and may reflect either a) mediation or cascade where the effects of a risk factor can be explained by more proximal causes or b) an artifact where some variables had somewhat stronger associations with externalizing problems than others, and the stronger predictors were retained over weaker predictors. Whether for mediational or artifactual reasons, some variables likely could be grouped together to reduce collinearity. For example, some non-significant parenting predictors of externalizing problems (e.g., physical harm, positive parenting) could be subsumed under other parenting variables (e.g., spanking, harsh discipline). In other cases, some risk factors may be more salient to the individual (e.g., individual stress) than others (e.g., family stress). Another possibility is that some risk factors may have different effects at different developmental periods. For example, stress may be experienced particularly acutely in later adolescence because of the lagging development of the PFC relative to the earlier developing limbic areas (Petersen et al., 2012).

Fourteen variables, on the other hand, were associated with either the intercepts or slopes of externalizing problems in both the bivariate and multivariate models, i.e., had independent associations even after controlling for the other risk factors. Although the independent associations of these risk factors do not demonstrate causality, the robust associations of these risk factors with externalizing problems in this study and many prior studies point to candidate mechanisms for future research to examine. The incremental prediction of these risk factors may prove even more useful when specifying the developmental process linking them to the development of externalizing problems.

It is also encouraging that the predictions in the present study were somewhat accurate: The model accounted for more than two-thirds of the variability in externalizing problems over time. The specific risk and protective factors accounted for approximately 6% of the variability in the ending values of externalizing problems at age 27 above the effects of time (linear, quadratic, etc.). The fact that considerable variance in externalizing problems appears to be explained by continuity (i.e., the effects of time) is consistent with the notion that past behavior is the best predictor of future behavior. However, continuity does not necessarily indicate stability of individuals' levels of externalizing problems over time—there were heterogeneous patterns of change within individuals across time. Moreover, even the continuity of externalizing behavior followed a nonlinear pattern of change across time. In general, the model was somewhat accurate in predicting within-individual changes in externalizing problems from ages 5 to 27. Nevertheless, we were able to augment our predictions using specific risk factors that explained variance in externalizing problems

above and beyond the strong continuity of externalizing behavior over more than 20 years. In other words, taking into account family process, peer process, stress, and child characteristics like temperament and language ability can improve our predictive accuracy of the development of externalizing problems.

The model's predictions were tested on illegal behavior in an attempt to validate the actuarial model's predictive utility. The model's predicted values were a fairly good predictor of the person having been arrested, used or injected illegal drugs, and driven while drunk. Moreover, the risk factors for externalizing identified two risk profiles associated with high risk of arrest: 1) high peer deviance in later adolescence (predicting a 67% risk) and 2) males with high individual stress during later adolescence (predicting a 54% risk). These findings could reflect two possibilities. First, the risk profiles could reflect causal pathways involving deviant peers and, particularly for males, high stress. Second, the risk profiles could reflect markers of other, unmeasured causal processes. Even if the risk profiles represent markers rather than causal processes *per se*, they may still be useful in prediction, as was the case in the present study in which the risk profiles were fairly discriminating in terms of risk for arrest. Thus, evidence suggests that the externalizing profiles and their associated risk factors were meaningful for predicting important and costly societal outcomes. The risk profiles for arrest may lead to targeted, cost-effective interventions that take into account both risks and developmental stage (i.e., peer deviance and individual stress during later adolescence). For example, preventive interventions might target adolescents with deviant peers or adolescent males who are at risk of experiencing high levels of stress.

Given that we used changing measures of externalizing problems over time, it was necessary to consider whether the measures showed construct validity invariance. Earlier, we described five necessary conditions for construct validity invariance. First, the measures were chosen from subscales that were 1) theoretically- and developmentally-meaningful, and that were 2) derived empirically (Achenbach & Rescorla, 2001) with a similar factor structure across time (Reitz, Deković, & Meijer, 2005). In addition, the externalizing problems in our study showed 3) strong cross-time consistency and 4) strong convergent and discriminant validity over time with respect to internalizing problems. Also, the items showed 5) high internal consistency at each age. Finally, the patterns of trajectories showed construct validity. Consistent with our findings, previous studies have shown decreases in externalizing problems from early childhood to preadolescence (Leve et al., 2005), and we had found the same in the present sample (Keiley et al., 2000). In addition, studies show that rates of mental disorder increase from late childhood to adolescence, consistent with our findings of increasing rates of externalizing problems during the same time frame (D. L. Newman et al., 1996). Moreover, studies examining the age-crime curve show increases during adolescence and decreases during adulthood (Sampson & Laub, 2003), also consistent with our findings. Thus, we feel that there is theoretical and empirical support for the construct validity invariance of our measures of externalizing problems on a common metric, which permits examining the changes in externalizing problems over time.

Given changing measurement across time, we are unable to be completely certain that differences across time were reflective of actual change, and we therefore present our

descriptions of the trajectories with caution. Nevertheless, we feel the externalizing profiles in the present study reflected meaningful individual differences in the development of externalizing problems, and the predictors of these problems were meaningful, as well. Heterotypic continuity is a developmental complexity that arises in many different domains, and we feel that seeking to understand and predict changes across important developmental periods is better than ignoring the phenotypic complexities associated with meaningful developmental change.

The individual components of the approach in the present study are not new. There are precedents in the literature of modeling actuarial predictions from growth curves (Deater-Deckard et al., 1998; Lussier & Davies, 2011), of modeling growth curve trajectories derived from different raters (Odgers et al., 2008) and different measures/scales (Owens & Shaw, 2003; Pettit et al., 2007) from childhood to adulthood (Curran et al., 2008), and of rendering measures more equivalent with proportional scoring metrics in the context of growth curves (McArdle et al., 2000). What is novel in the present study is the assembling of these approaches to predict risk for developing externalizing problems from childhood to adulthood. We believe this is a methodological and conceptual advance toward understanding development, because using different measures over time is necessary for describing development across long spans characterized by changes in how the same construct is manifested over time, or heterotypic continuity. Following Rutter and Sroufe's (2000) argument that developmental psychopathology research should strive to understand development over the lifespan, we were able to chart the development of externalizing behavior over years in one piece by using developmentally-appropriate, changing measures over time. An alternative approach for future studies might be to examine change over time in subdimensions of externalizing problems (e.g., physical aggression) in an attempt to focus on more homotypic patterns of change (e.g., Olson et al., 2013). Nevertheless, we feel there is utility in examining the construct of general externalizing behavior because 1) it is an efficient summary of many cases of psychopathology, 2) the subdimensions tend to co-occur, and 3) similar developmental processes appear to be involved with the different subdimensions (Olson, Bates, Sandy, & Lanthier, 2000).

Strengths and Limitations

The present study had several strengths. First, the measurement of externalizing problems was theoretically- and empirically-based on developmentally-relevant constructs. Second, it incorporated measurement of externalizing problems from multiple sources to reduce source bias. Third, it described and predicted developmental profiles from childhood to adulthood with numerous measurement occasions. Fourth, it considered many different domains of risk and developmental time frames. Finally, it applied its predictions to illegal behavior in order to demonstrate the robustness of the model's predictive utility.

The present study also had several limitations. First, the use of changing measures in externalizing problems may limit our ability to draw conclusive inferences regarding developmental change. POM scores across different measures may not be comparable if their items have different severity. For instance, if items on one scale reflect a more severe level of psychopathology (e.g., uses drugs and sets fires) than items on another scale (e.g.,

argues and brags), the proportion scores may not be equivalent across the two scales. This may not be as much of an issue for the externalizing problem questionnaires in the present study because their items overlap substantially. For other research using more disparate scales, item response theory models or other advanced measurement models may be necessary for calculating trait or scale scores by linking items (e.g., Curran et al., 2008). Even the same items could have different severities for different informants (e.g., self versus mother), however, so the importance of ensuring that measures are conceptually and empirically equivalent is not specific to studies using changing measures. In any case, keeping the measures identical over time would not resolve the issue of measuring developmental change, because a static measure would likely not have construct validity invariance across the time frame in the present study due to the heterotypic continuity of externalizing problems. Changes in constructs over time require changes in measurement (Eddy et al., 1998); failure to accommodate changes in the form of externalizing problems over time may make differences across age meaningless. Achenbach (2005) emphasized the need for measures to reflect the changing nature of externalizing problems, forming the theoretical foundation for the changes in items across development in the Achenbach scales according to developmentally-relevant forms of externalizing behavior. Because of the developmental relevance of the scales, Owens and Shaw (2003) also modeled externalizing trajectories with different Achenbach scales over time. In any case, we have attempted to show that the trajectories are meaningful insofar as they map onto other important externalizing problems. Moreover, we have shown evidence for the construct validity invariance of the externalizing problems over time, and there is prior support for the trajectories we identified.

Another limitation comes from the fact that, because of the correlational nature of the present study, we cannot determine causality from any of the risk or protective factors that we examined. Nevertheless, the risk factors were chosen because of their theoretical importance for the development of behavior problems. Additionally, there is an elevated likelihood of type II error because of shared variance between the risk factors. In other words, we may have failed to detect meaningful associations because of overlapping variance and conservative cutoffs. There is an increase in power to detect associations in studies with repeated measures, however, providing further confidence in our findings (Muthén & Curran, 1997). Finally, we did not consider how risk factors may interact to influence externalizing problems. Emerging findings suggest, for example, that temperament-by-parenting interactions augment the prediction of children's externalizing behavior (Bates & Pettit, in press; Bates, Schermerhorn, & Petersen, 2012). Thus, future studies could extend these findings by testing the interactions among risk factors. Future studies might also consider the effects that risk factors may have on each other in successive developmental periods, such as models of developmental cascades (e.g., Cox, Mills-Koonce, Propper, & Gariépy, 2010; Dodge, Greenberg, Malone, & Conduct Problems Prevention Research Group, 2008; Dodge et al., 2009; Lansford et al., 2010; Masten et al., 2005; Sitnick, Shaw, & Hyde, in press).

Conclusion

In summary, the present study considered the development of externalizing problems as a function of early risk factors along with successive risk and protective factors from early childhood to adulthood. The development of externalizing problems can be described in terms of multiple domains of risk—both from their momentum of adjustment (i.e., continuity) and from other risk across multiple developmental eras. The specific risk factors explained variance in the development of externalizing problems above and beyond the strong continuity of externalizing problems. Moreover, the continuity of externalizing problems was nonlinear across time. The findings support a model that simultaneously takes into account numerous characteristics of children and their living situations, and predicts trajectories of externalizing problems with a moderately high degree of accuracy. The study also suggests that the modeled trajectories and their predictions are also meaningful for important societal outcomes including arrests, illegal drug use, and drunk driving. Nevertheless, there remains much room for improvement in terms of predictive precision. We expect to see improvements from: 1) further specification of the causal mechanisms, 2) the consideration of additional risk factors from other domains (e.g., genetics), and 3) modeling interactions among risk factors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge Thomas Achenbach and his associates for their research and development of the Achenbach System of Empirically Based Assessment. The Child Development Project has been funded by grants MH42498, MH56961, MH57024, and MH57095 from the National Institute of Mental Health; HD30572 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development; and DA016903 from the National Institute on Drug Abuse. Isaac T. Petersen was supported by Grant Number TL1 TR000162 (A. Shekhar, PI) from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award, and by a Ruth L. Kirschstein National Research Service Award grant from the National Institute of Mental Health (1 F31 MH100814-01A1). Kenneth A. Dodge was supported by the National Institute of Drug Abuse grant 2K05 DA015226. This research was supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute, and in part by the Indiana METACyt Initiative. The Indiana METACyt Initiative at IU is also supported in part by Lilly Endowment, Inc. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health or Indiana University.

References

Achenbach, TM. Manual for the Child Behavior Checklist and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry; 1991a.

Achenbach, TM. Manual for the Teacher’s Report Form and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry; 1991b.

Achenbach, TM. Manual for the Youth Self-Report and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry; 1991c.

Achenbach, TM. Manual for the Young Adult Self-Report and Young Adult Behavior Checklist. Burlington, VT: University of Vermont, Department of Psychiatry; 1997.

Achenbach TM. Advancing assessment of children and adolescents: Commentary on evidence-based assessment of child and adolescent disorders. Journal of Clinical Child and Adolescent Psychology. 2005; 34:541–547.10.1207/s15374424jccp3403_9 [PubMed: 16026217]

Achenbach TM, Howell CT, McConaughy SH, Stanger C. Six-year predictors of problems in a national sample: III. Transitions to young adult syndromes. Journal of the American Academy of Child & Adolescent Psychiatry. 1995; 34:658–669.10.1097/00004583-199505000-00018 [PubMed: 7775361]

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

Ackerman BP, D'Eramo KS, Umylny L, Schultz D, Izard CE. Family structure and the externalizing behavior of children from economically disadvantaged families. Journal of Family Psychology. 2001; 15:288–300.10.1037/0893-3200.15.2.288 [PubMed: 11458634]

Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatrica. 2007; 96:644–647.10.1111/j.1651-2227.2006.00178.x [PubMed: 17376185]

Appleyard K, Egeland B, Van Dulmen MHM, Sroufe LA. When more is not better: the role of cumulative risk in child behavior outcomes. Journal of Child Psychology & Psychiatry. 2005; 46:235–245.10.1111/j.1469-7610.2004.00351.x [PubMed: 15755300]

Bates JE, Bayles K. Objective and subjective components in mothers' perceptions of their children from age 6 months to 3 years. Merrill-Palmer Quarterly: Journal of Developmental Psychology. 1984; 30:111–130.

Bates JE, Marvinney D, Kelly T, Dodge KA, Bennett DS, Pettit GS. Child care history and kindergarten adjustment. Developmental Psychology. 1994; 30:690–700.10.1037/0012-1649.30.5.690

Bates, JE.; Pettit, GS. Temperament, parenting, and social development. In: Grusec, J.; Hastings, PD., editors. Handbook of socialization. New York, NY, US: Guildford Press; in press

Bates JE, Pettit GS, Dodge KA, Ridge B. Interaction of temperamental resistance to control and restrictive parenting in the development of externalizing behavior. Developmental Psychology. 1998; 34:982–995. [PubMed: 9779744]

Bates, JE.; Schermerhorn, AC.; Petersen, IT. Temperament and parenting in developmental perspective. In: Zentner, M.; Shiner, R., editors. Handbook of Temperament. New York: The Guilford Press; 2012. p. 425-441.

Bauer DJ. Observations on the use of growth mixture models in psychological research. Multivariate Behavioral Research. 2007; 42:757–786.

Bauer DJ, Curran PJ. Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. Psychological Methods. 2003; 8:338–363.10.1037/1082-989x.8.3.338 [PubMed: 14596495]

Bauer DJ, Reyes HLM. Modeling variability in individual development: Differences of degree or kind? Child Development Perspectives. 2010; 4:114–122.10.1111/j.1750-8606.2010.00129.x [PubMed: 24910711]

Beyers JM, Bates JE, Pettit GS, Dodge KA. Neighborhood structure, parenting processes, and the development of youths' externalizing behaviors: A multilevel analysis. American Journal of Community Psychology. 2003; 31:35–53.10.1023/a:1023018502759 [PubMed: 12741688]

Bongers IL, Koot HM, Ende Jvd, Verhulst FC. Developmental trajectories of externalizing behaviors in childhood and adolescence. Child Development. 2004; 75:1523–1537. [PubMed: 15369529]

Brame B, Nagin DS, Tremblay RE. Developmental trajectories of physical aggression from school entry to late adolescence. Journal of Child Psychology and Psychiatry. 2001; 42:503–512.10.1111/1469-7610.00744 [PubMed: 11383966]

Broidy LM, Nagin DS, Tremblay RE, Bates JE, Brame B, Dodge KA, Vitaro F. Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. Developmental Psychology. 2003; 39:222–245.10.1037/0012-1649.39.2.222 [PubMed: 12661883]

Buehler C, Anthony C, Krishnakumar A, Stone G, Gerard J, Pemberton S. Interparental conflict and youth problem behaviors: A meta-analysis. Journal of Child & Family Studies. 1997; 6:233–247.

Burt SA, Donnellan MB, Iacono WG, McGue M. Age-of-onset or behavioral sub-types? A prospective comparison of two approaches to characterizing the heterogeneity within antisocial behavior. Journal of Abnormal Child Psychology. 2011; 39:633–644.10.1007/s10802-011-9491-9 [PubMed: 21298333]

Caspi, A. Personality development across the life course. In: Damon, W., editor. Handbook of child psychology: Social, emotional, and personality development. 5. Vol. 3. Hoboken, NJ, US: John Wiley & Sons Inc; 1998. p. 311-388.

Coghill D, Sonuga-Barke EJS. Annual Research Review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders: implications of recent empirical study. Journal of Child Psychology and Psychiatry. 2012; 53:469–489.10.1111/j.1469-7610.2011.02511.x [PubMed: 22288576]

Connell AM, Frye AA. Growth mixture modelling in developmental psychology: overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behaviour. Infant & Child Development. 2006; 15:609–621.10.1002/icd.481

Connell AM, Goodman SH. The association between psychopathology in fathers versus mothers and children's internalizing and externalizing behavior problems: A meta-analysis. Psychological Bulletin. 2002; 128:746–773.10.1037/0033-2909.128.5.746 [PubMed: 12206193]

Cox MJ, Mills-Koonce R, Propper C, Gariépy JL. Systems theory and cascades in developmental psychopathology. Development and Psychopathology. 2010; 22:497–506.10.1017/s0954579410000234 [PubMed: 20576174]

Curran PJ, Hussong AM, Cai L, Huang W, Chassin L, Sher KJ, Zucker RA. Pooling data from multiple longitudinal studies: The role of item response theory in integrative data analysis. Developmental Psychology. 2008; 44:365.10.1037/0012-1649.44.2.365 [PubMed: 18331129]

Dallaire D, Wilson L. The relation of exposure to parental criminal activity, arrest, and sentencing to children's maladjustment. Journal of Child & Family Studies. 2010; 19:404–418.10.1007/s10826-009-9311-9

Dawes RM, Faust D, Meehl PE. Clinical versus actuarial judgment. Science. 1989; 243:1668–1674.10.1126/science.2648573 [PubMed: 2648573]

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; 10:469–493.10.1017/S0954579498001709 [PubMed: 9741678]

Deegan J. On the occurrence of standardized regression coefficients greater than one. Educational and Psychological Measurement. 1978; 38:873–888.10.1177/001316447803800404

Derksen S, Keselman HJ. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. British Journal of Mathematical and Statistical Psychology. 1992; 45:265–282.10.1111/j.2044-8317.1992.tb00992.x

Dodge KA, Bates JE, Pettit GS. Mechanisms in the cycle of violence. Science. 1990; 250:1678–1683.10.1126/science.2270481 [PubMed: 2270481]

Dodge, KA.; Coie, JD.; Lynam, D. Aggression and antisocial behavior in youth. In: Damon, W.; Lerner, RM.; Eisenberg, N., editors. Handbook of child psychology 6th ed., Vol. 3: Social, emotional, and personality development. Hoboken, NJ, US: John Wiley & Sons Inc; 2006. p. 719-788.

Dodge KA, Greenberg MT, Malone PS. Conduct Problems Prevention Research Group. Testing an idealized dynamic cascade model of the development of serious violence in adolescence. Child Development. 2008; 79:1907–1927.10.1111/j.1467-8624.2008.01233.x [PubMed: 19037957]

Dodge, KA.; Malone, PS.; Lansford, JE.; Miller, S.; Bates, JE.; Pettit, GS. A dynamic cascade model of the development of substance-use onset. Vol. 74. Boston, MA, US: Wiley-Blackwell; 2009.

Dodge KA, Pettit GS, Bates JE. Socialization mediators of the relation between socioeconomic status and child conduct problems. Child Development Special Issue: Children and poverty. 1994; 65:649–665.10.1111/j.1467-8624.1994.tb00774.x

Dodge KA, Pettit GS, Bates JE, Valente E. Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. Journal of Abnormal Psychology. 1995; 104:632–643. [PubMed: 8530766]

Eddy JM, Dishion TJ, Stoolmiller M. The analysis of intervention change in children and families: Methodological and conceptual issues embedded in intervention studies. Journal of Abnormal Child Psychology. 1998; 26:53–69.10.1023/a:1022634807098 [PubMed: 9566546]

Farrington, DP. Key results from the first forty years of the Cambridge Study in Delinquent Development. In: Thornberry, TP.; Krohn, MD., editors. Taking stock of felinquency: An

overview of findings from contemporary longitudinal studies. New York, NY, US: Springer; 2003. p. 137-183.

Fergusson DM. Stability and change in externalising behaviours. European Archives of Psychiatry & Clinical Neuroscience. 1998; 248:4. [PubMed: 9561348]

Gershoff ET, Lansford JE, Sexton HR, Davis-Kean P, Sameroff AJ. Longitudinal links between spanking and children's externalizing behaviors in a national sample of white, black, hispanic, and asian american families. Child Development. 2012:838–843.10.1111/j.1467-8624.2011.01732.x [PubMed: 22304526]

Goodnight JA, Bates JE, Newman JP, Dodge KA, Pettit GS. The interactive influences of friend deviance and reward dominance on the development of externalizing behavior during middle adolescence. Journal of Abnormal Child Psychology. 2006; 34:573–583. [PubMed: 16823636]

Goodnight JA, Bates JE, Pettit GS, Dodge KA. Parents' campaigns to reduce their children's conduct problems: Interactions with temperamental resistance to control. International Journal of Developmental Science. 2008; 2:100–119.10.3233/DEV-2008-21207

Goodnight JA, Bates JE, Staples AD, Pettit GS, Dodge KA. Temperamental resistance to control increases the association between sleep problems and externalizing behavior development. Journal of Family Psychology. 2007; 21:39–48.10.1037/0893-3200.21.1.39 [PubMed: 17371108]

Graham J, Olchowski A, Gilreath T. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prevention Science. 2007; 8:206–213.10.1007/s11121-007-0070-9 [PubMed: 17549635]

Grimm KJ, Ram N, Hamagami F. Nonlinear growth curves in developmental research. Child Development. 2011; 82:1357–1371.10.1111/j.1467-8624.2011.01630.x [PubMed: 21824131]

Hayatbakhsh MR, Najman JM, Khatun M, Al Mamun A, Bor W, Clavarino A. A longitudinal study of child mental health and problem behaviours at 14 years of age following unplanned pregnancy. Psychiatry Research. 2011; 185:200–204.10.1016/j.psychres.2010.05.019 [PubMed: 20576294]

Honaker J, King G. What to do about missing values in time-series cross-section data. American Journal of Political Science. 2010; 54:561–581.10.1111/j.1540-5907.2010.00447.x

Honaker J, King G, Blackwell M. Amelia II: A program for missing data. Journal of Statistical Software. 2011; 45:1–47.

Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: A conditional inference framework. Journal of Computational and Graphical Statistics. 2006; 15:651–674.10.1198/106186006X133933

Huesmann LR, Eron LD, Lefkowitz MM, Walder LO. Stability of aggression over time and generations. Developmental Psychology. 1984; 20:1120–1134.10.1037/0012-1649.20.6.1120

Jeličić H, Phelps E, Lerner RM. Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. Developmental Psychology. 2009; 45:1195–1199.10.1037/a0015665 [PubMed: 19586189]

Keijsers L, Loeber R, Branje S, Meeus W. Parent-child relationships of boys in different offending trajectories. A developmental perspective. Journal of Child Psychology and Psychiatry. 2012:1222–1232.10.1111/j.1469-7610.2012.02585.x [PubMed: 22816682]

Keiley MK, Bates JE, Dodge KA, Pettit GS. A cross-domain growth analysis: externalizing and internalizing behaviors during 8 years of childhood. Journal of Abnormal Child Psychology. 2000; 28:161–179.10.1023/A:1005122814723 [PubMed: 10834768]

Keiley, MK.; Bates, JE.; Dodge, KA.; Pettit, GS. Effects of temperament on the development of externalizing and internalizing behaviors over 9 years. Hauppauge, NY, US: Nova Science Publishers; 2001.

Keiley MK, Howe TR, Dodge KA, Bates JE, Pettit GS. The timing of child physical maltreatment: A cross-domain growth analysis of impact on adolescent externalizing and interalizing problems. Development and Psychopathology. 2001; 13:891–912. [PubMed: 11771913]

Keiley MK, Lofthouse N, Bates JE, Dodge KA, Pettit GS. Differential risks of covarying and pure components in mother and teacher reports of externalizing and internalizing behavior across ages 5 to 14. Journal of Abnormal Child Psychology. 2003; 31:267–283. [PubMed: 12774860]

Kim KJ, Conger RD, Elder GHJ, Lorenz FO. Reciprocal influences between stressful life events and adolescent internalizing and externalizing problems. Child Development. 2003; 74:127–143.10.1111/1467-8624.00525 [PubMed: 12625440]

Knight GP, Zerr AA. Informed theory and measurement equivalence in child development research. Child Development Perspectives. 2010; 4:25–30.10.1111/j.1750-8606.2009.00112.x

Kohlberg, L.; LaCrosse, J.; Ricks, D. The predictability of adult mental health from childhood behavior. In: Wolman, B., editor. Manual of child psychopathology. New York, NY, US: McGraw-Hill; 1972. p. 1217-1284.

Kokko K, Pulkkinen L, Huesmann LR, Dubow EF, Boxer P. Intensity of aggression in childhood as a predictor of different forms of adult aggression: A two-country (Finland and the United States) analysis. Journal of Research on Adolescence (Blackwell Publishing Limited). 2009; 19:9–34.10.1111/j.1532-7795.2009.00579.x

Krueger RF, Markon KE, Patrick CJ, Iacono WG. Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implications for DSM-V. Journal of Abnormal Psychology. 2005; 114:537–550.10.1037/0021-843x.114.4.537 [PubMed: 16351376]

Lansford JE, Malone PS, Castellino DR, Dodge KA, Pettit GS, Bates JE. Trajectories of internalizing, externalizing, and grades for children who have and have not experienced their parents' divorce or separation. Journal of Family Psychology. 2006; 20:292–301.10.1037/0893-3200.20.2.292 [PubMed: 16756405]

Lansford JE, Malone PS, Dodge KA, Pettit GS, Bates JE. Developmental cascades of peer rejection, social information processing biases, and aggression during middle childhood. Development and Psychopathology. 2010; 22:593–602.10.1017/S0954579410000301 [PubMed: 20576181]

Leve LD, Kim HK, Pears KC. Childhood temperament and family environment as predictors of internalizing and externalizing trajectories from ages 5 to 17. Journal of Abnormal Child Psychology. 2005; 33:505–520.10.1007/s10802-005-6734-7 [PubMed: 16195947]

Lish JD, Weissman MM, Adams PB, Hoven CW, Bird H. Family psychiatric screening instrument for epidemiologic studies: pilot testing and validation. Psychiatry Research. 1995; 57:169–180.10.1016/0165-1781(95)02632-7 [PubMed: 7480383]

Little, TD. Longitudinal structural equation modeling. New York, NY, US: The Guilford Press; 2013.

Little, TD.; Card, NA.; Preacher, KJ.; McConnell, E. Modeling longitudinal data from research on adolescence. In: Lerner, RM.; Steinberg, L., editors. Handbook of adolescent psychology: Individual bases of adolescent development. Hoboken, NJ: John Wiley & Sons, Inc; 2009.

Lumley, T. mitools: Tools for multiple imputation of missing data. R package version 2. 2010. Retrieved from http://cran.r-project.org/web/packages/mitools/

Lussier P, Davies G. A person-oriented perspective on sexual offenders, offending trajectories, and risk of recidivism: A new challenge for policymakers, risk assessors, and actuarial prediction? Psychology, Public Policy, and Law. 2011; 17:530–561.10.1037/a0024388

Manganello JA, Taylor CA. Television exposure as a risk factor for aggressive behavior among 3-year-old children. Archives of Pediatrics & Adolescent Medicine. 2009; 163:1037–1045.10.1001/archpediatrics.2009.193 [PubMed: 19884595]

Markon KE, Krueger RF. Categorical and continuous models of liability to externalizing disorders: A direct comparison in NESARC. Archives of General Psychiatry. 2005; 62:1352–1359.10.1001/archpsyc.62.12.1352 [PubMed: 16330723]

Masten AS, Roisman GI, Long JD, Burt KB, Obradović J, Riley JR, Tellegen A. Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. Developmental Psychology. 2005; 41:733–746.10.1037/0012-1649.41.5.733 [PubMed: 16173871]

McArdle JJ, Grimm KJ, Hamagami F, Bowles RP, Meredith W. Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. Psychological Methods. 2009; 14:126–149.10.1037/a0015857 [PubMed: 19485625]

McArdle JJ, Hamagami F, Meredith W, Bradway KP. Modeling the dynamic hypotheses of Gf-Gc theory using longitudinal life-span data. Learning & Individual Differences. 2000; 12:53.10.1016/S1041-6080(00)00036-4

Miller JL, Vaillancourt T, Boyle MH. Examining the heterotypic continuity of aggression using teacher reports: Results from a national Canadian study. Social Development. 2009; 18:164–180.10.1111/j.1467-9507.2008.00480.x

Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. Psychological Review. 1993; 100:674–701.10.1037/0033-295X.100.4.674 [PubMed: 8255953]

Mott FL, Kowaleski-Jones L, Menaghan EG. Paternal absence and child behavior: Does a child's gender make a difference? Journal of Marriage and Family. 1997; 59:103–118.

Muthén BO, Curran PJ. General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. Psychological Methods. 1997; 2:371–402.10.1037/1082-989x.2.4.371

Muthén BO, Muthén LK. The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a US national sample. Journal of Studies on Alcohol and Drugs. 2000; 61:290.

Newman DL, Moffitt TE, Caspi A, Magdol L, Silva PA, Stanton WR. Psychiatric disorder in a birth cohort of young adults: Prevalence, comorbidity, clinical significance, and new case incidence from ages 11–21. Journal of Consulting and Clinical Psychology. 1996; 64:552–562.10.1037/0022-006x.64.3.552 [PubMed: 8698949]

Newman JP, Patterson CM, Kosson DS. Response perseveration in psychopaths. Journal of Abnormal Psychology. 1987; 96:145–148.10.1037/0021-843x.96.2.145 [PubMed: 3584663]

Nigg, JT.; Huang-Pollock, CL. An early-onset model of the role of executive functions and intelligence in conduct disorder/delinquency. In: Lahey, BB.; Moffitt, TE.; Caspi, A., editors. Causes of conduct disorder and juvenile delinquency. New York, NY, US: Guilford Press; 2003. p. 227-253.

Odgers CL, Moffitt TE, Broadbent JM, Dickson N, Hancox RJ, Harrington H, Caspi A. Female and male antisocial trajectories: From childhood origins to adult outcomes. Development and Psychopathology. 2008; 20:673–716.10.1017/S0954579408000333 [PubMed: 18423100]

Olson SL, Bates JE, Sandy JM, Lanthier R. Early developmental precursors of externalizing behavior in middle childhood and adolescence. Journal of Abnormal Child Psychology. 2000; 28:119–133.10.1023/A:1005166629744 [PubMed: 10834765]

Olson SL, Sameroff AJ, Davis-Kean P, Lansford JE, Sexton HR, Bates JE, Dodge KA. Deconstructing the externalizing spectrum: Growth patterns of overt aggression, covert aggression, oppositional behavior, impulsivity/inattention and emotion dysregulation between school entry and early adolescence. Development and Psychopathology. 2013; 25:817–842. [PubMed: 23880394]

Olweus D. Stability of aggressive reaction patterns in males: A review. Psychological Bulletin. 1979; 86:852–875.10.1037/0033-2909.86.4.852 [PubMed: 482487]

Owens EB, Shaw DS. Predicting growth curves of externalizing behavior across the preschool years. Journal of Abnormal Child Psychology. 2003; 31:575–590.10.1023/a:1026254005632 [PubMed: 14658739]

Patterson GR. Orderly change in a stable world: The antisocial trait as a chimera. Journal of Consulting and Clinical Psychology. 1993; 61:911–919. [PubMed: 8113492]

Petersen IT, Bates JE, D'Onofrio BM, Coyne CA, Lansford JE, Dodge KA, Van Hulle CA. Language ability predicts the development of behavior problems in children. Journal of Abnormal Psychology. 2013; 122:542–557.10.1037/a0031963 [PubMed: 23713507]

Petersen IT, Bates JE, Goodnight JA, Dodge KA, Lansford JE, Pettit GS, Dick DM. Interaction between serotonin transporter polymorphism (5-HTTLPR) and stressful life events in adolescents' trajectories of anxious/depressed symptoms. Developmental Psychology. 2012; 48:1463–1475.10.1037/a0027471 [PubMed: 22390669]

Petersen IT, Bates JE, Staples AD. The role of language ability and self-regulation in the development of inattentive-hyperactive behavior problems. Development and Psychopathology. in press.

Pettit GS, Bates JE, Dodge KA. Supportive parenting, ecological context, and children's adjustment: A seven-year longitudinal study. Child Development. 1997; 68:908–923.10.1111/j.1467-8624.1997.tb01970.x

Pettit GS, Keiley MK, Laird RD, Bates JE, Dodge KA. Predicting the developmental course of mother-reported monitoring across childhood and adolescence from early proactive parenting,

child temperament, and parents' worries. Journal of Family Psychology. 2007; 21:206–217.10.1037/0893-3200.21.2.206 [PubMed: 17605543]

Peugh JL. A practical guide to multilevel modeling. Journal of School Psychology. 2010; 48:85–112.10.1016/j.jsp.2009.09.002 [PubMed: 20006989]

Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D. and the R Core team. nlme: Linear and nonlinear mixed effects models. R package version 3.1-93. 2009. Retrieved from http://cran.r-project.org/web/packages/nlme/index.html

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2009. Retrieved from http://www.R-project.org

Reitz E, Deković M, Meijer A. The structure and stability of externalizing and internalizing problem behavior during early adolescence. Journal of Youth and Adolescence. 2005; 34:577–588.10.1007/s10964-005-8947-z

Rubin, DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons, Inc; 1987.

Rutter M, Sroufe LA. Developmental psychopathology: Concepts and challenges. Development and Psychopathology. 2000; 12:265–296.10.1017/S0954579400003023 [PubMed: 11014739]

Sampson RJ, Laub JH. Life-course desisters? Trajectories of crime among delinquent boys followed to age 70*. Criminology. 2003; 41:555–592.10.1111/j.1745-9125.2003.tb00997.x

Schafer, JL. Analysis of incomplete multivariate data. Vol. 72. New York, NY: Chapman & Hall/CRC; 1997.

Schulenberg JE, Maslowsky J. Taking substance use and development seriously: Developmentally distal and proximal influences on adolescence drug use. Monographs of the Society for Research in Child Development. 2009; 74:121–130.10.1111/j.1540-5834.2009.00544.x [PubMed: 19930536]

Sheather, SJ. A modern approach to regression with R. New York, NY US: Springer Science +Business Media; 2009.

Siegel RA. Probability of punishment and suppression of behavior in psychopathic and nonpsychopathic offenders. Journal of Abnormal Psychology. 1978; 87:514–522.10.1037/0021-843x.87.5.514 [PubMed: 701604]

Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics. 2005; 21:3940–3941. [PubMed: 16096348]

Singer, JD.; Willett, JB. Applied longitudinal data analysis: Modeling change and event occurrence. Oxford University Press; USA: 2003.

Sitnick SL, Shaw DS, Hyde LW. Precursors of adolescent substance use from early childhood and early adolescence: Testing a developmental cascade model. Development and Psychopathology. 26:125–140. in press. 10.1017/S0954579413000539 [PubMed: 24029248]

Snijders, TAB.; Bosker, RJ. Multilevel analysis: An introduction to basic and advanced multilevel modeling. London: SAGE Publications; 2011.

Stoolmiller, M. Using latent growth curve models to study developmental processes. In: Gottman, JM., editor. The analysis of change. Hillsdale, NJ: Erlbaum; 1995. p. 103-138.

Straus MA. Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. Journal of Marriage and Family. 1979; 41:75–88.10.2307/351733

Sutton S. Predicting and explaining intentions and behavior: How well are we doing? Journal of Applied Social Psychology. 1998; 28:1317–1338.10.1111/j.1559-1816.1998.tb01679.x

Underwood BD. Law and the crystal ball: Predicting behavior with statistical inference and individualized judgment. The Yale Law Journal. 1979; 88:1408–1448.

Venables, WN.; Ripley, BD. Modern applied statistics with S. 4. New York: Springer; 2002.

Wakschlag LS, Gordon RA, Lahey BB, Loeber R, Green SM, Leventhal BL. Maternal age at first birth and boys' risk for conduct disorder. Journal of Research on Adolescence. 2000; 10:417–441.

Walters GD. The latent structure of life-course-persistent antisocial behavior: Is Moffitt's developmental taxonomy a true taxonomy? Journal of Consulting and Clinical Psychology. 2011; 79:96–105.10.1037/a0021519 [PubMed: 21171739]

Walters GD. Developmental trajectories of delinquent behavior: One pattern or several? Criminal Justice and Behavior. 2012; 39:1192–1203.10.1177/0093854812443645

Walters GD, Ruscio J. Trajectories of youthful antisocial behavior: Categories or continua? Journal of Abnormal Child Psychology. 2013; 41:653–666.10.1007/s10802-012-9700-1 [PubMed: 23188257]

Walton K, Ormel J, Krueger R. The dimensional nature of externalizing behaviors in adolescence: Evidence from a direct comparison of categorical, dimensional, and hybrid models. Journal of Abnormal Child Psychology. 2011; 39:553–561.10.1007/s10802-010-9478-y [PubMed: 21210204]

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation; 1974.

Widaman KF, Ferrer E, Conger RD. Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. Child Development Perspectives. 2010; 4:10–18.10.1111/j.1750-8606.2009.00110.x [PubMed: 20369028]

Willett JB, Singer JD, Martin NC. The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. Development and Psychopathology. 1998; 10:395–426.10.1017/s0954579498001667 [PubMed: 9635230]

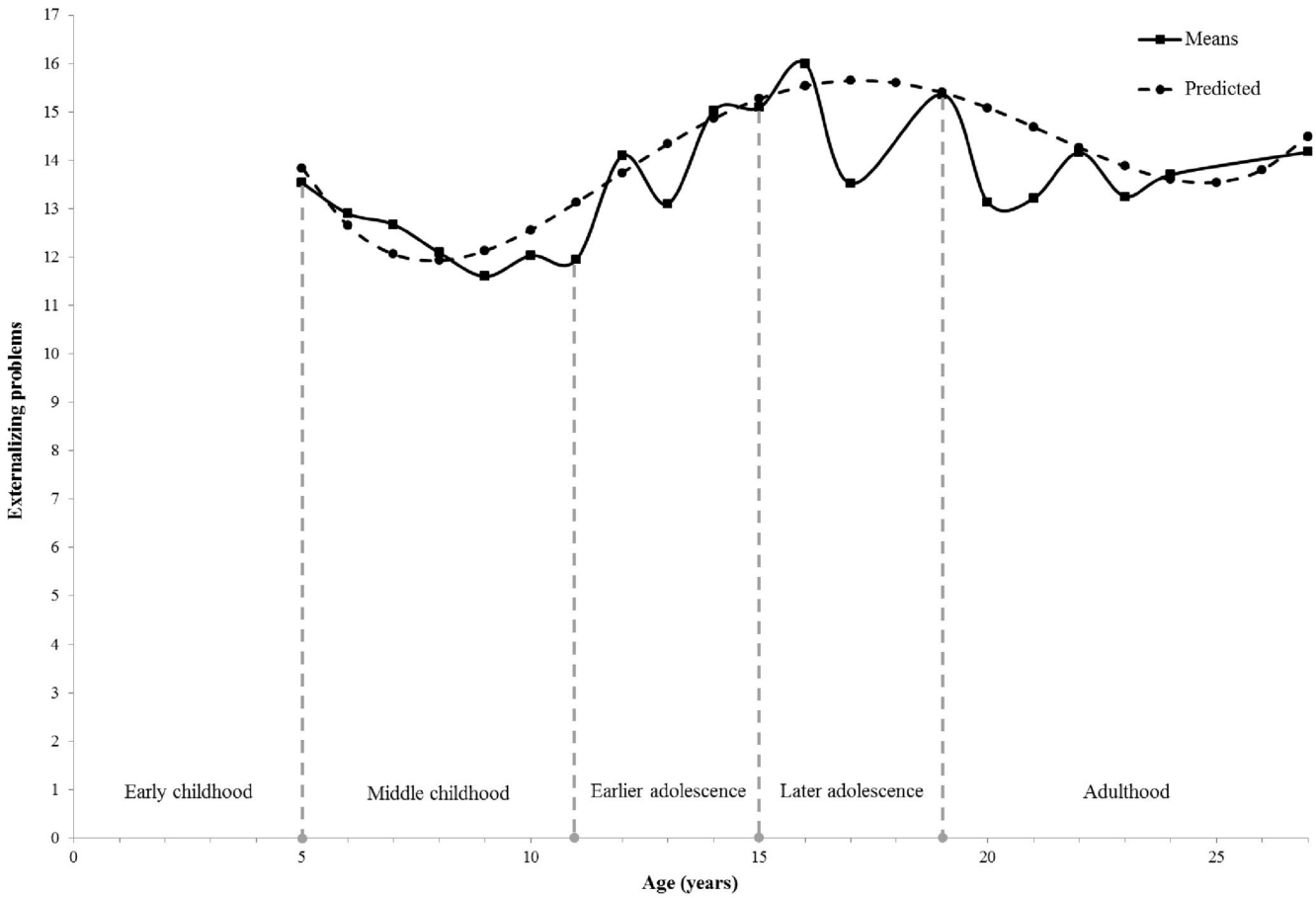

Figure 1.

Prototypical quartic trajectory of externalizing problems over time by time frame (overlaid with averages of externalizing problems over time).

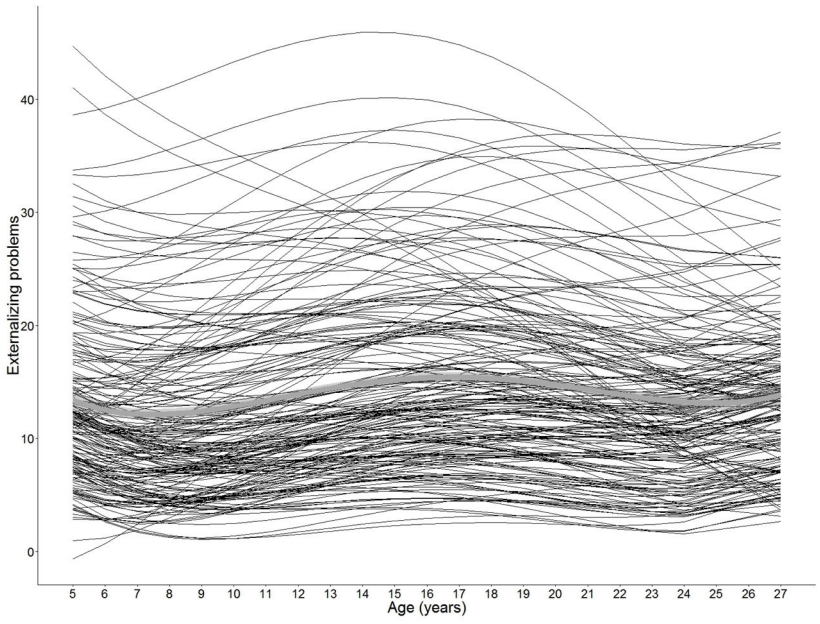

Figure 2.

Random subsample of 250 individuals' predicted quartic trajectories of externalizing problems in black. Average trajectory in gray. The subsample is depicted rather than the whole sample for the sake of graphic clarity.

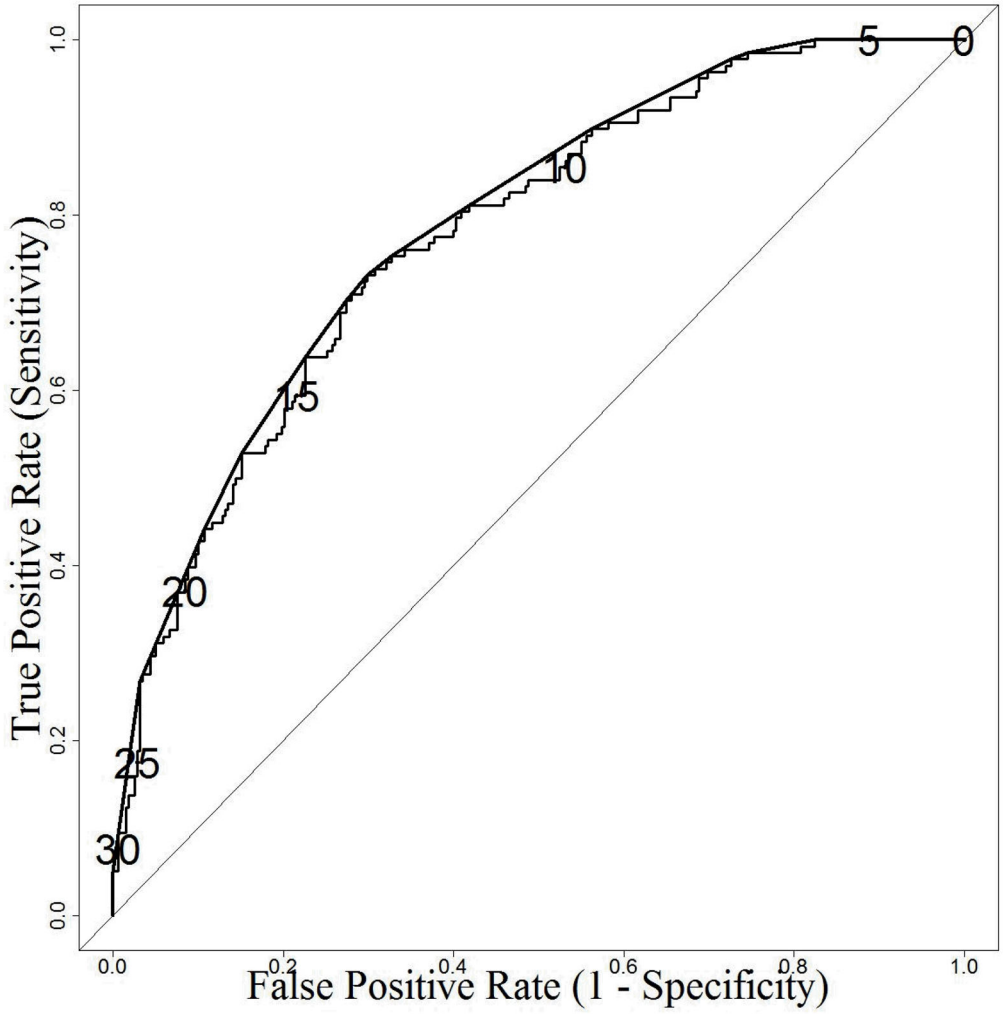

Figure 3.

Empirical ROC curve of average predicted externalizing problems from the final model predicting arrests, overlaid with ROC convex hull and cutoff values for externalizing problems at various thresholds.

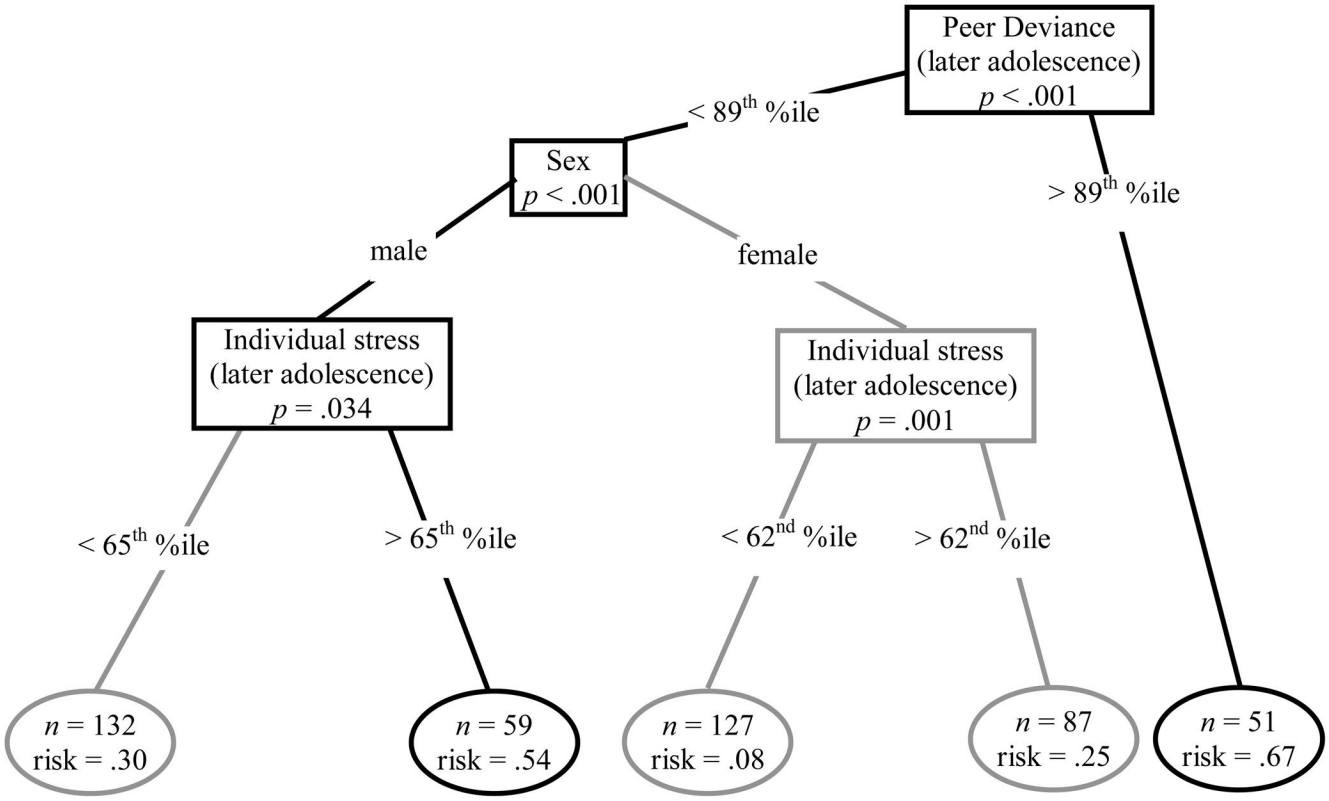

Figure 4.

Conditional inference tree predicting arrest. Boxes represent binary splits at the cutpoint (corresponding levels or percentiles are specified) that maximized the discrepancy in the two subsamples' risk for arrest. Ovals represent subsamples with different combinations of values on the risk factors. Black lines, boxes, and ovals represent the high risk combinations of risk for arrest (risk of arrest is greater than or equal to .54). Gray lines, boxes, and ovals represent the low risk combinations of risk for arrest (risk of arrest is less than or equal to . 30).

Table 1

Table of Measures (and when measured)

		Early childhood						Middle childhood					Earlier adolescence				Later adolescence			Adulthood							
	Variable -- Age (years)	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**19**	**20**	**21**	**22**	**23**	**24**	**26**	**27**
Predictors	Sex	x																				
	Ethnicity	x																				
	SES	x																				
	Values aggression	x																				
	Positive parenting	x																				
	Parental conflict	M																				
	Exposure to violence	M																				
	Harsh discipline	M																				
	Spanks child	M (2)																				
	Physical harm	1																				
	Difficult temperament	x																				
	Unadaptable temperament	x																				
	Resistance to control temperament	x																				
	Medical complications in pregnancy	x																				
	Teenage pregnancy	1																				
	Unplanned pregnancy	1																				
	Child:adult ratio	x																				
	Single mother	1																				
	Cohabiting	1																				
	Amount of television watched	x																				
	Non-maternal childcare	M																				
	Social information processing	x	M	M	M																	
	Low father caregiving	1	1	1	1	1																
	Social preference	x	M	M	M	M (1)																
	Intelligence									x												
	Peer deviance							M	M		M	M	M									

	Early childhood	Middle childhood					Earlier adolescence				Later adolescence			Adulthood							
Variable -- Age (years)	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**19**	**20**	**21**	**22**	**23**	**24**	**26**	**27**
Parental monitoring								M		M	M	M									
Parental involvement									x			x									
Father arrest												1									
Mother arrest												1									
Father alcohol/drug												1									
Mother alcohol/drug												1									
Reward sensitivity												x									
Sleep problems	M	M	M	M	M	M	M	M	M	M	M	M	M								
Family stress	M	M	M	M	M	M	M	M	M	M	M	M	M								
Language ability		M (2)	M	M	M	M	M	M	M	M	M	M	M								
Divorce	1	1	1	1	1	1	1	1	1	1	1	1	1							1	
Individual stress										x	M		M							x	
Length of unemployment																					x
Educational attainment																					x
Internalizing problems	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M		M
Outcomes: Externalizing problems	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x
Outcomes: Achenbach - Mother	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL	CBCL								
Outcomes: Achenbach - Teacher	TRF	TRF	TRF	TRF	TRF	TRF	TRF	TRF	TRF												
Outcomes: Achenbach - Father	CBCL	CBCL	CBCL	CBCL	CBCL (1)																
Outcomes: Achenbach - Self								YSR		YSR	YSR	YSR	YSR	YSR	YASR	YASR	YASR	YASR	YASR		YASR
Outcomes: Achenbach - Peer																					YASR
Illegal Behavior: Ever arrested																					1
Illegal Behavior: Ever injected illegal drugs																					1
Illegal Behavior: Used Recreational drugs																					1
Illegal Behavior: Drunk Driving																					1

Note: Times of measurement refers to when the data were *collected*, not necessarily the time frame to which the measure refers. No variables in the present study were collected at ages 18 or 25.

x = measured; M = mean across time (within time frame); 1 = scored 1 if value is present at any measurement occasion within time frame, scored 0 if never present within time frame; (1) denotes only in cohort 1; (2) denotes only in cohort 2; CBCL = Child Behavior Checklist; TRF = Teacher Report Form; YSR = Youth Self Report; YASR = Young Adult Self Report

Table 2

Rates of missingness (in percentages) for all predictors (by time frame) and outcomes

Predictors				Outcomes (Externalizing problems)			
Class	**Predictor**	**Time**	**Missingness**	**Age**	**Rater**	**Missingness**	**Overall Missingness**
Demographics	female	TI	0	5	Mother	3	
	SES	EC	3	5	Teacher	2	1
	African American	TI	0	5	Father	34	
	Other ethnicity	TI	0	6	Mother	16	
	Educational attainment	AD	22	6	Teacher	8	6
	Length of unemployment	AD	22	6	Father	43	
Parenting	Values aggression	EC	4	7	Mother	21	
	Positive parenting	EC	1	7	Teacher	12	9
	Parental conflict	EC	2	7	Father	51	
	Exposure to violence	EC	3	8	Mother	19	
	Harsh discipline	EC	1	8	Teacher	15	11
	Parental involvement	EA	25	8	Father	46	
	Parental involvement	LA	20	9	Mother	28	
	Parental monitoring	EA	20	9	Teacher	20	15
	Parental monitoring	LA	18	9	Father	72	
Parent Adjustment	Spank	EC	54	10	Mother	32	18
	Physical harm	EC	1	10	Teacher	23	
	Mother drug use	LA	23	11	Mother	23	18
	Father drug use	LA	29	11	Teacher	24	
	Mother arrest	LA	23	12	Mother	22	
	Father arrest	LA	32	12	Teacher	27	18
Peers	Peer deviance	EA	17	12	Self	30	
	Peer deviance	LA	18	13	Mother	27	22

Predictors				Outcomes (Externalizing problems)			
Class	**Predictor**	**Time**	**Missingness**	**Age**	**Rater**	**Missingness**	**Overall Missingness**
Child Characteristics	Social preference	EC	3	13	Teacher	31	28
	Social preference	MC	8	14	Mother	30	
	Intelligence	EA	27	14	Self	30	
	Social information processing	EC	0	15	Mother	30	29
	Social information processing	MC	10	15	Self	31	
	Difficult temperament	EC	5	16	Mother	23	22
	Resistance to control temperament	EC	5	16	Self	23	
	Unadaptable temperament	EC	5	17	Mother	26	25
	Reward sensitivity	LA	32	17	Self	27	
	Internalizing problems	EC	1	19	Self	21	21
	Internalizing problems	MC	3	20	Self	18	18
	Internalizing problems	EA	14	21	Self	21	21
	Internalizing problems	LA	16	22	Self	20	20
	Internalizing problems	AD	8	23	Self	17	17
Stress	Language ability	MC	21	24	Self	21	21
	Language ability	EA	23	27	Self	21	21
	Language ability	LA	88	27	Peer	37	
	Family stress	EC	1				
	Family stress	MC	8				
	Family stress	EA	15				
	Family stress	LA	17				
	Individual stress	EA	28				
	Individual stress	LA	21				

Illegal Behavior

Age	Variable	Missingness
27	Arrest	22
27	Inject drugs	22
27	Recreational drugs	22

Predictors				Outcomes (Externalizing problems)			
Class	**Predictor**	**Time**	**Missingness**	**Age**	**Rater**	**Missingness**	**Overall Missingness**
	Individual stress	AD	29	27	Drunk driving	23	
	Sleep problems	EC	29				
	Sleep problems	MC	10				
	Sleep problems	EA	17				
	Sleep problems	LA	18				
Pregnancy	Medical complications	EC	2				
	Teenage pregnancy	EC	2				
	Unplanned pregnancy	EC	6				
Family Background	Low father caregiving	EC	1				
	Low father caregiving	MC	10				
	Single mother	EC	6				
	Cohabiting	EC	6				
	Divorce	EC	1				
	Divorce	MC	8				
	Divorce	EA	15				
	Divorce	LA	16				
	Divorce	AD	56				
	Child:adult ratio	EC	1				
Activities	Amount of television watched	EC	4				
	Non-maternal childcare	EC	1				

Note: High rates of missingness were observed for spanking and father-reported externalizing problems at age 9 because data were only collected for one of the two cohorts. TI = time-invariant, EC = early childhood, MC = middle childhood, EA = earlier adolescence, LA = later adolescence, AD = adulthood. Ethnicity was dummy coded into two variables (variable name in italics): 1) *African American*=1, European American=0, other ethnicity=0 and 2) *other ethnicity*=1, European American=0, African American=0.

Table 3

Correlation matrix of externalizing problems (and means and standard deviations)

Age	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
1. 5 yrs	1																			
2. 6 yrs	.69	1																		
3. 7 yrs	.61	.67	1																	
4. 8 yrs	.60	.65	.70	1																
5. 9 yrs	.57	.64	.63	.71	1															
6. 10 yrs	.51	.56	.63	.60	.66	1														
7. 11 yrs	.50	.55	.55	.60	.62	.67	1													
8. 12 yrs	.47	.49	.49	.61	.62	.62	.65	1												
9. 13 yrs	.43	.51	.50	.61	.64	.53	.56	.68	1											
10. 14 yrs	.39	.44	.45	.57	.51	.50	.51	.63	.62	1										
11. 15 yrs	.38	.44	.45	.52	.50	.55	.54	.63	.56	.78	1									
12. 16 yrs	.38	.45	.45	.52	.50	.50	.51	.64	.63	.75	.79	1								
13. 17 yrs	.34	.39	.41	.46	.44	.43	.46	.53	.49	.70	.70	.76	1							
14. 19 yrs	.16	.25	.21	.23	.24	.21	.25	.34	.26	.48	.51	.49	.61	1						
15. 20 yrs	.21	.28	.25	.25	.27	.30	.30	.37	.29	.47	.51	.51	.60	.72	1					
16. 21 yrs	.23	.21	.27	.26	.29	.29	.28	.34	.24	.42	.49	.48	.54	.62	.70	1				
17. 22 yrs	.20	.23	.25	.29	.26	.25	.34	.32	.28	.46	.47	.47	.56	.61	.68	.68	1			
18. 23 yrs	.21	.26	.25	.28	.28	.29	.34	.39	.34	.45	.49	.49	.55	.61	.72	.70	.73	1		
19. 24 yrs	.21	.21	.29	.18	.23	.32	.31	.30	.28	.38	.41	.43	.50	.59	.68	.67	.71	.81	1	
20. 27 yrs	.28	.29	.29	.32	.34	.36	.41	.39	.41	.46	.43	.49	.53	.52	.58	.58	.55	.64	.67	1
M	13.53	12.90	12.67	12.09	11.60	12.03	11.95	14.09	13.10	15.02	15.09	16.00	13.52	15.36	13.13	13.21	14.16	13.25	13.70	14.17
SD	8.74	10.13	10.96	10.48	11.10	11.88	10.75	9.32	12.61	10.40	10.36	10.45	10.01	10.00	10.71	10.90	10.95	10.89	10.94	10.61

Note: All correlations are significant at $p < .001$ level.

Table 4

Individual Predictor Models of Risk Factors Predicting Intercepts and Linear Slopes of Externalizing Problems

		Intercept					Linear Slope				
Predictor	**Time frame**	***B***	***SE***	***df***	***t***	***p***	***B***	***SE***	***df***	***t***	***p***
female	time-invariant	−3.19	0.61	582	−5.23	.00	−0.03	0.05	9004	−0.50	.62
SES	early childhood	−0.16	0.02	568	−7.78	.00	0.01	0.00	8793	3.34	.00
African American	time-invariant	1.46	0.84	582	1.74	.08	−0.23	0.07	9004	−3.32	.00
Other ethnicity	time-invariant	0.12	2.33	582	0.05	.96	0.11	0.20	9004	0.54	.59
Educational attainment	adulthood	−1.15	0.13	455	−8.73	.00	0.01	0.01	7648	0.52	.61
Length of unemployment	adulthood	0.14	0.02	424	6.86	.00	0.00	0.00	7169	1.42	.16
Values aggression	early childhood	2.78	0.45	562	6.19	.00	−0.01	0.04	8725	−0.28	.78
Positive parenting	early childhood	−1.44	0.48	575	−3.02	.00	0.02	0.04	8883	0.37	.71
Parental conflict	early childhood	1.19	0.34	571	3.47	.00	0.04	0.03	8821	1.23	.22
Exposure to violence	early childhood	1.54	0.34	567	4.60	.00	−0.07	0.03	8771	−2.33	.02
Harsh discipline	early childhood	2.67	0.33	578	8.07	.00	−0.09	0.03	8933	−2.90	.00
Parental involvement	earlier adolescence	0.04	0.05	435	0.71	.48	0.00	0.00	7616	−0.08	.94
Parental involvement	later adolescence	−0.11	0.07	464	−1.60	.11	−0.01	0.01	8073	−0.85	.39
Parental monitoring	earlier adolescence	−5.75	1.13	467	−5.07	.00	−0.06	0.10	8142	−0.54	.59
Parental monitoring	later adolescence	−0.38	0.45	479	−0.85	.40	−0.04	0.04	8301	−1.00	.32
Spank	early childhood	2.01	0.36	267	5.67	.00	−0.08	0.03	4135	−2.75	.01
Physical harm	early childhood	3.97	1.60	578	2.48	.01	−0.18	0.14	8933	−1.37	.17
Mother drug use	later adolescence	3.16	0.70	450	4.50	.00	0.12	0.06	7839	1.93	.05
Father drug use	later adolescence	2.89	0.63	415	4.60	.00	0.06	0.06	7237	1.09	.27
Mother arrest	later adolescence	2.97	1.40	448	2.13	.03	0.17	0.12	7802	1.35	.18
Father arrest	later adolescence	3.81	0.80	395	4.74	.00	−0.11	0.07	6917	−1.47	.14
Peer deviance	earlier adolescence	5.64	0.72	485	7.83	.00	0.06	0.07	8361	0.83	.40
Peer deviance	later adolescence	5.25	0.47	480	11.08	.00	0.22	0.05	8312	4.80	.00
Social preference	early childhood	−2.98	0.31	564	−9.60	.00	0.11	0.03	8731	4.21	.00
Social preference	middle childhood	−4.79	0.35	537	−13.51	.00	0.15	0.04	8491	4.30	.00

		Intercept					Linear Slope				
Predictor	**Time frame**	***B***	***SE***	***df***	***t***	***p***	***B***	***SE***	***df***	***t***	***p***
Intelligence	earlier adolescence	−0.47	0.11	426	−4.47	.00	0.03	0.01	7466	3.01	.00
SIP	early childhood	6.58	1.47	581	4.47	.00	−0.43	0.13	9002	−3.19	.00
SIP	middle childhood	8.17	2.37	527	3.44	.00	−0.59	0.20	8578	−3.03	.00
Difficult temperament	early childhood	1.44	0.37	564	3.92	.00	−0.10	0.03	8762	−3.13	.00
Resistance to control temperament	early childhood	2.34	0.30	565	7.69	.00	−0.08	0.03	8769	−2.74	.01
Unadaptable temperament	early childhood	−0.24	0.32	565	−0.75	.46	−0.01	0.03	8769	−0.26	.79
Reward sensitivity	later adolescence	0.03	0.01	393	1.92	.06	0.00	0.00	6885	−2.44	.01
Internalizing problems	early childhood	0.25	0.06	580	4.12	.00	−0.02	0.01	8990	−3.03	.00
Internalizing problems	middle childhood	0.54	0.06	563	9.57	.00	−0.03	0.00	8945	−6.83	.00
Internalizing problems	earlier adolescence	0.41	0.04	503	9.12	.00	−0.01	0.00	8514	−3.01	.00
Internalizing problems	later adolescence	0.36	0.03	491	10.52	.00	0.01	0.00	8426	2.11	.04
Internalizing problems	adulthood	0.30	0.02	535	14.21	.00	0.02	0.00	8771	9.78	.00
Language ability	middle childhood	−0.11	0.01	462	−7.85	.00	0.00	0.00	7903	3.82	.00
Language ability	earlier adolescence	−0.10	0.01	447	−7.86	.00	0.00	0.00	7773	4.13	.00
Language ability	later adolescence	−0.08	0.03	71	−2.92	.00	0.00	0.00	1325	0.34	.73
Family stress	early childhood	1.18	0.20	578	5.89	.00	−0.04	0.02	8933	−2.24	.03
Family stress	middle childhood	1.49	0.19	537	7.93	.00	−0.05	0.02	8733	−3.26	.00
Family stress	earlier adolescence	1.48	0.19	495	7.90	.00	−0.02	0.02	8440	−1.36	.17
Family stress	later adolescence	1.09	0.15	485	7.35	.00	−0.02	0.01	8349	−1.75	.08
Individual stress	earlier adolescence	0.66	0.08	420	8.00	.00	0.01	0.01	7471	0.78	.44
Individual stress	later adolescence	1.22	0.09	463	13.00	.00	0.03	0.01	8074	3.37	.00
Individual stress	adulthood	1.15	0.16	413	7.01	.00	0.03	0.02	7067	1.84	.07
Sleep problems	early childhood	3.15	0.88	566	3.59	.00	−0.14	0.08	8769	−1.83	.07
Sleep problems	middle childhood	7.51	1.30	527	5.77	.00	−0.33	0.11	8572	−2.90	.00
Sleep problems	earlier adolescence	8.73	1.34	484	6.51	.00	−0.25	0.12	8255	−2.12	.03
Sleep problems	later adolescence	6.41	1.07	475	5.97	.00	−0.12	0.10	8177	−1.29	.20
Medical complications	early childhood	0.93	0.56	569	1.65	.10	−0.04	0.05	8812	−0.74	.46
Teenage pregnancy	early childhood	−0.94	1.93	159	−0.49	.63	−0.26	0.17	2334	−1.51	.13
Unplanned pregnancy	early childhood	2.89	0.95	570	3.03	.00	−0.19	0.08	8822	−2.30	.02

		Intercept					Linear Slope				
Predictor	**Time frame**	***B***	***SE***	***df***	***t***	***p***	***B***	***SE***	***df***	***t***	***p***
Low father caregiving	early childhood	3.20	0.74	576	4.30	.00	−0.15	0.07	8902	−2.36	.02
Low father caregiving	middle childhood	1.57	0.65	524	2.42	.02	−0.09	0.06	8524	−1.55	.12
Single mother	early childhood	3.84	0.73	548	5.27	.00	−0.10	0.06	8501	−1.70	.09
Cohabiting	early childhood	4.72	1.89	548	2.49	.01	−0.14	0.16	8501	−0.89	.38
Divorce	early childhood	2.13	0.75	578	2.86	.00	−0.04	0.07	8933	−0.56	.57
Divorce	middle childhood	1.85	0.78	536	2.35	.02	−0.03	0.06	8723	−0.51	.61
Divorce	earlier adolescence	2.48	0.84	497	2.95	.00	−0.08	0.07	8471	−1.05	.30
Divorce	later adolescence	1.63	0.77	488	2.12	.03	−0.03	0.07	8391	−0.44	.66
Divorce	adulthood	−0.14	2.11	258	−0.06	.95	0.11	0.18	4372	0.58	.56
Child:adult ratio	early childhood	1.10	0.37	578	3.00	.00	−0.05	0.03	8949	−1.45	.15
Television	early childhood	1.58	0.55	559	2.87	.00	−0.06	0.05	8671	−1.25	.21
Childcare	early childhood	0.82	0.41	577	2.00	.05	−0.01	0.04	8921	−0.40	.69

Note: significant predictors in bold. SIP = social information processing. Ethnicity was dummy coded into two variables (variable name in italics): 1) *African American*=1, European American=0, other ethnicity=0 and 2) *other ethnicity*=1, European American=0, African American=0.

Table 5

Final model

Variable	Time frame	*B*	*β*	*SE*	lower	upper	*p*
intercept		−3.42	0.00	2.47	−8.33	1.49	.170
time (linear)	time-varying	1.05	0.70	0.29	0.47	1.63	< .001
time (quadratic)	time-varying	0.29	3.86	0.04	0.21	0.37	< .001
time (cubic)	time-varying	0.02	6.31	0.00	0.02	0.03	< .001
time (quartic)	time-varying	0.00	3.12	0.00	0.00	0.00	< .001
Predicting intercepts							
Female	time-invariant	−2.39	−0.15	0.50	−3.38	−1.40	< .001
Peer deviance	earlier adolescence	1.87	0.09	0.65	0.60	3.15	.004
Peer deviance	later adolescence	5.87	0.09	1.33	3.19	8.54	< .001
Individual stress	later adolescence	0.37	0.13	0.10	0.17	0.57	< .001
Internalizing problems	later adolescence	0.09	0.11	0.04	0.01	0.18	.027
Internalizing problems	adulthood	0.13	0.16	0.03	0.06	0.20	< .001
Predicting linear slopes							
Resistant to control temperament	early childhood	−0.05	−0.02	0.02	−0.08	−0.02	.002
Spanking	early childhood	−0.05	−0.03	0.02	−0.09	−0.02	.001
Harsh discipline	early childhood	−0.05	−0.01	0.02	−0.08	−0.01	.008
Low father caregiving	early childhood	−0.09	−0.02	0.04	−0.16	−0.02	.013
Social preference	early childhood	0.04	0.03	0.02	0.00	0.07	.043
Social preference	middle childhood	0.13	0.05	0.02	0.09	0.18	< .001
Internalizing problems	middle childhood	−0.01	−0.07	0.00	−0.02	−0.01	< .001
Language ability	earlier adolescence	0.00	0.03	0.00	0.00	0.00	.007
Peer deviance	later adolescence	0.32	0.10	0.07	0.17	0.46	< .001

Note: “Upper” and “lower” represent the upper and lower bounds, respectively, of 95% confidence intervals. Some regression coefficients were too small to be visible with two decimal places, including time (quartic): $B = 0.0005$ ($SE = 0.0001$), and language ability (earlier adolescence): $B = 0.0019$ ($SE = 0.0007$). The standardized coefficients for the quadratic, cubic, and quartic effects of time were greater than 1 because these predictors were perfectly correlated with the predictor representing linear time. Strong correlations among predictors make it likely for the predictors’ standardized regression coefficients to exceed (−1, 1) (Deegan, 1978).

Table 6

Sensitivity and specificity of model's predictions for clinical level of externalizing problems and illegal behavior.

Variable	Prevalence of Outcome	Sensitivity	Specificity	AUC	Optimal Cutoff
Arrest	30.3%	0.73	0.70	0.78	12.6
Clinical level of EXT	3.9%	1.00	0.92	0.99	28.2
Illegal drug use	26.0%	0.67	0.68	0.72	15.7
Inject illegal drugs	2.4%	0.91	0.61	0.82	13.4
Drunk driving	15.3%	0.81	0.53	0.71	12.9

Note: The cutoff for clinical levels of externalizing problems was selected as the number of externalizing problems at age 27 reflecting a *T*-score of 67 or above on the Young Adult Self-Report (i.e., 19 externalizing problems or a POM score of 33.9 for females and 22 externalizing problems or a POM score of 39.3 for males), as suggested by Achenbach (1997). AUC values above 0.9 indicate high accuracy, 0.7–0.9 indicates moderate accuracy, and 0.5–0.7 indicates low accuracy (Akobeng, 2007). Prevalence of outcome refers to the percentage of the sample that exceeded threshold for the clinical outcome. AUC = area under the curve. EXT = externalizing problems.