LIFE OUTCOMES AS A FUNCTION OF DEVELOPMENTAL PATTERNS OF PROBLEM BEHAVIORS

By

Allison Elizabeth Gornik

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Doctor of Philosophy

2020

ABSTRACT

LIFE OUTCOMES AS A FUNCTION OF
DEVELOPMENTAL PATTERNS OF PROBLEM BEHAVIORS

By

Allison Elizabeth Gornik

The aim of this study was to examine how the magnitude and change in youth internalizing and externalizing problem behaviors across childhood and adolescence were connected to theoretically meaningful outcomes in adulthood. Two central questions were addressed: First, do the effects of externalizing and internalizing problem behaviors sustain or decay over time in their relation to later adulthood outcomes? Second, if so, which informants in children's lives (mothers, fathers, teachers, and the children themselves) provide the strongest predictive ability? Utilizing a longitudinal, high-risk sample that followed children (N=1,069) and their families for several decades, the current study employed structural equation modeling to estimate growth models for each informant, test the latent growth parameters for their prediction of later outcomes, and examine whether significant associations held above and beyond a series of covariates. Findings suggest that overall, children's internalizing problem behaviors carry less weight towards long-term functioning than the presence of externalizing problem behaviors, while externalizing problem behaviors carry implications even above relevant familial and child-based covariates for subsequent timing of transitions, romantic relationship quality, educational attainment, income, and legal infractions. Most of these outcomes were predicted by two or more informants, and the pattern of results suggest that parent-reports possess reasonably good predictive validity, but that teacher-reports in particular possess uniquely strong associations with outcomes. Additional clinical and research implications are discussed.

ACKNOWLEDGEMENTS

First and foremost, thank you to my adviser, Emily Durbin, without whom this dissertation (or the long winding road before it) would not have been possible. Thank you for your good ideas, your humor, your passion for causes outside of academia, and your unwavering confidence in me. Thank you to my committee and to the Clinical Science program at large, from whom I've gained knowledge, thoughtfulness, and creativity. Thank you to the children and families I've worked with over the years, who gave me the questions I sought to answer in this dissertation.

Thank you to my grad school friend group over the past six years, who let me be me in all my various ways of being, from endless work dates to nights dancing on sometimes sticky floors and sometimes both in the same day. From themed parties (soup, wine, sushi) to quiet nights of reflection to group adventures in St. Louis and DC, I am thankful for the laughs and the love that continue even though we're now spread far and wide. Thank you especially to my bookclub friends, who have kept me grounded and remind me it's okay to admit when it gets hard, and to my cohort, who are some of the most brilliant and competent people I know.

Thank you to Eli, to Chris, to Sharon, to Duncan, to Sara, to Sean, to Marc, and to New College, for each of the different ways you have taught me to think a bit differently about the world.

Thank you to my Ohio family for your excellent sense of humor, weird quirkiness, frequent Snapchats of cats, and general brilliance mixed with unwavering support.

Thank you to Jonathan. I am better for knowing you, and I look forward to many more times of windows down, singing with the radio to a song we can't name.

Thank you to Derek. You are my favorite person in this world. We've found flowers twice the height what I thought we could find. I hope to never stop adventuring with you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Internalizing and externalizing problem behaviors can develop beginning in early childhood, and elevations in either of these domains are associated with a host of negative concurrent and future youth outcomes (e.g., Broidy et al., 2003; Keiley, Lofthouse, Bates, Dodge, & Pettit, 2003; Kovacs & Devlin, 1998; Moffitt, Caspi, Harrington, & Milne, 2002). Internalizing behaviors are characterized by withdrawal, anxiety, depression, fearfulness, and somatic complaints, while externalizing behaviors are characterized by aggression, defiance, hyperactivity, and destructive behaviors (e.g., Achenbach, 1991). Models of mean levels of internalizing and externalizing symptomatology over development have highlighted the importance of early and later emergence of problems, as well as the consequence of chronicity over time. Some models theorize that children's dispositions and behaviors elicit reactions from their social environment, creating a transactional response that reinforces their behavioral presentation as time passes (Caspi, Bem, & Elder, 1989; Hinde, 1992; Sameroff, 1983). Thus, problems that emerge earlier in development are believed to be, on average, more consequential than problems that emerge later; while later-emerging problems can also self-sustain, problems during early developmental periods in which key psychological processes are developing are theorized to be more damaging than those first appearing after those psychological skills and processes are already in place. Specifically, pathology that emerges earlier on is associated with greater impairment, a more intractable course, and increased difficulty achieving success in developmental tasks (e.g, Barlow, 1988; Cicchetti & Toth, 1995).

Onset of disorders prior to adulthood is also associated with several environmental, family, and child-centered risk factors. One birth cohort study found that an onset of major depressive disorder in childhood compared to adulthood was associated with distinct genetic and

childhood risk factors, including greater parental psychopathology and criminality, parent instability, and more teacher-reported peer and behavior problems at age 9 (Jaffee et al., 2002). Those experiencing an onset in emerging adulthood had similar earlier presentations to those who had never been depressed. In another birth cohort study, van Os, Jones, Lewis, Wadsworth, and Murray (1997) identified several neurodevelopmental risk factors that separated childhood-onset internalizing disorders from adult-onset internalizing disorders, and risk factors preceded the onset of childhood internalizing disorders. Following the development of problem behaviors, subsequent developmental stages may then be affected; for example, Essau (2003) highlighted how earlier-onset, chronic trajectories of externalizing behaviors in particular prompt a developmental trajectory that negatively impacts later outcomes across a variety of domains. Similarly, Moffitt et al. (2002) found that youth with elevated antisocial symptoms only in adolescence were more successful during the transition to adulthood than youth with elevated antisocial behaviors earlier in life.

While a modest but compelling body of literature has identified connections between children's problem behaviors and emerging adulthood functioning and outcomes, these studies have rarely extended past age 21. Many notable adulthood outcomes, however, are better established after this period, when assessments of educational attainment, income, legal infractions, timing of transitions, and difficulties in serious romantic relationships may yield more stable estimates. Thus, the current investigation can answer questions with greater confidence as to the extent to which the effects of problem behaviors sustain or decay with greater temporal distance from childhood.

The predictive validity of child psychopathology for later outcomes cannot be understood without considering how psychopathology was measured. One of the most robust findings in

child clinical psychology is that informants reporting on children's psychopathology will provide reports that are only weakly to moderately intercorrelated (Achenbach, 2006). In a recent meta-analysis analyzing informant agreement on internalizing and externalizing behaviors over the last 25 years, De Los Reyes et al. (2015) found the overall average correlation of agreement to be r=.25 for internalizing problems and r=.30 for externalizing problems. In light of this low-to-moderate agreement between reporters (generally termed "informant discrepancies"), researchers and clinicians are faced with the challenge of how to best integrate (or not) conflicting sources of information, particularly when trying to predict outcomes. Unfortunately, there is no definitive way to identify the most "accurate" informant (e.g., De Los Reyes et al., 2015), and diverging reports can lead to inconsistent conclusions across research studies. Most work estimating the longitudinal trajectories of problem behaviors and their correlates has not included multiple reporters, or has simply averaged them across informants (e.g., Leve, Kim, & Pears, 2005). However, in a review of studies using a variety of designs (e.g., randomized controlled trials, laboratory, longitudinal), De Los Reyes (2011) noted that informant discrepancies often exhibit stability over time, predict outcomes in a way that individual informants do not, and reveal important information about the role of context in exhibiting behaviors.

The current investigation targets two central questions: Given the critical importance of understanding predictors of adulthood outcomes, in what way do individual differences in problem behaviors across child and adolescent development predict these outcomes? Secondly, which informants in children's lives (mothers, fathers, teachers, and the children themselves) provide the strongest predictive ability?

Problem Behavior Trajectories

Estimates of continuity in children's problem behaviors vary by assessment method, but most studies have found that rank-order stability between persons is typically preserved, even while there are mean level changes in problem behaviors across development (e.g., Broidy et al., 2003; Reitz, Deković, & Meijer, 2005; Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007). By utilizing a dimensional approach to problem behaviors, we can capture children's mean-level trajectories across development and explore whether individual variations around trajectories predict later outcomes.

Externalizing and internalizing problem behaviors. For externalizing problem behaviors, mean-levels appear to decrease across early childhood through late adolescence (e.g., Bongers, Koot, van der Ende, & Velhurst, 2004; Leve et al., 2005). However, the particular presentations of externalizing behaviors may change over time, with aggression more characteristic of childhood-onset antisocial behavior (Burt, Donnellan, Iacono, & McGue, 2011). An earlier onset of pronounced externalizing behaviors has been linked to a more persistent and impairing course, with more co-occurring adjustment difficulties (e.g., Moffitt, 1993; Barker, Oliver, & Maughan, 2010). In the sample used in this dissertation (the Michigan Longitudinal Study, MLS), previous work has identified a linear decline in boys' aggression from preschool to early adolescence (Loukas, Zucker, Fitzgerald, & Krull, 2003).

For internalizing problem behaviors, most studies have identified relative stability across childhood and increases in adolescence, particularly for girls (Bongers, Koot, van der Ende, & Velhurst, 2003; Twenge & Nolen-Hoeksema, 2002). Most of these studies have focused on middle childhood and adolescence; a meta-analysis of the Children's Depression Inventory (Kovacs, 1985) indicated that girls display relative stable average trajectories from ages 8-11,

followed by a linear increase in depressive symptoms, while boys display a relatively stable average level from age 8-16 with a brief elevation at age 12. Studies that have examined internalizing symptom trajectories earlier in development have yielded conflicting and often opposite findings (e.g., Colder, Mott, & Berman, 2002; Gazelle & Ladd, 2003; Keiley et al., 2003). Taken together, the existing literature suggests a fairly consistent average decrease in externalizing behaviors across early childhood through late adolescence, while internalizing problem behaviors may be relatively stable in middle childhood and increase in adolescence, particularly for girls.

Individual differences in problem behaviors. While a sizeable literature has focused on trajectories of mean-level problem behavior as well as identifying latent classes of problems over time, little work has focused on describing the actual magnitude of average change that occurs. In an exception to this trend, Hicks et al. (2007) looked at emerging adults to determine their changes in externalizing disorders through mean-level change, individual-level change, and rank-order stability. However, similar explorations have not occurred for children, and thus less is known about change across the full span of development. Some research has targeted specific developmental periods, extending a maximum of four years, and has found that teachers rating very young children tend to show greater stability in externalizing than internalizing problems, and that parent reports of child behavior problems have greater rank-order stability than adolescents' and young adults' self-report (Ferdinand, Verhulst, & Wiznitzer, 1995; Verhulst & van der Ende, 1991, 1992; Youngstrom et al., 2000). However, these investigations have not encompassed early childhood through late adolescence, and thus far have relied on different informants to determine stability. The current study will assess rank-order stability over time,

and importantly, will also estimate the magnitude of absolute change in children's problem behaviors by each informant.

Distal Outcomes of Youth Problem Behaviors

Understandably, the vast majority of research examining how problem behaviors relate to other aspects of functioning has focused on concurrent concerns and proximal outcomes (i.e., outcomes in later childhood), while very few have examined outcomes past emerging adulthood, missing the window in which outcomes achieve greater stability. However, there are many ways in which individual variation around the average trajectories of problem behaviors across childhood might predict some aspect of functioning in adulthood. These early behavior problems may have a causal influence on major life transitions, roles, and accomplishments, even if the problem behaviors resolve, as children's early environmental experiences interact with their biology to coactively shape development and subsequent behavior, a process typically referred to as experiential canalization (Blair & Raver, 2012; Gottlieb, 1991, 1997). This theory suggests that there is iterative and selective optimization of behavior to maximize functioning within one's environmental experiences, and the influence of earlier optimization initiates a cascade effect on later optimizations even as behaviors and experiences change. For the current investigation, the outcome variables of interest are broad markers of psychosocial adjustment that could plausibly be influenced by earlier problem behaviors, including the timing of normative transitions, intimate partner relationship quality, educational attainment, household income, legal infractions, and alcohol use concerns. Thus, these outcomes are theoretically impacted by experiential canalization as well as sensitive periods of development, such as adolescence. The evidence on each of these as a distal outcome of youth problem behavior, most extending only to youths' early 20s, are reviewed below.

Transition timing. Many normative transitions typically occur across individuals' teenage years through their twenties, such as moving out of their parents' or caregivers' home, establishing romantic relationships that can include cohabitation and marriage, and becoming a parent. In general, the timing of these transitions is an important outcome because it provides insight into adjustment, and also may signal "faster" or "slower" life course trajectories (e.g. Chisholm, Quinlivan, Petersen, & Coall, 2005; Griskevicius et al., 2013). While living away from home carries different connotations if one considers moving for higher education versus more permanent moves, the current study explores three transition ages: age at first cohabitation with a long-term romantic partner, marital age, and age at first child.

In 2011, the average age of first marriage in the United States was 26.5 for women and age 29 for men, and the average age of cohabitation was 22 (Child Trends, 2006; U.S. Census Bureau, 2011). While 21% of women experience their first birth by age 20, nearly 10% of women experience their first birth after age 35, indicating wide variation in timing (Mathews & Hamilton, 2009). Earlier marriage and earlier age at first child have been linked to poorer functioning, including lower educational attainment, lower income, higher rates of lifetime poverty, and poorer health outcomes (Amato et al., 2008; Carroll et al., 2007; Coyne & D'Onofrio, 2012; Gaughan, 2002; Raley, Crissey, & Muller, 2007; Uecker & Stokes, 2008). Greater educational attainment is related to slower timing; women with college degrees are more likely to marry later and have fewer children. They are also less likely to divorce and report greater happiness in their marriages than women without college degrees (Isen & Stevenson, 2010).

Internalizing and externalizing behavior problems in adolescence appear related to earlier transition times in emerging adulthood. Girls who were depressed at age 15 were two and a half

times more likely to be cohabitating with a partner and somewhat more likely to have given birth by age 21 than girls who were not depressed at age 15 (Bardone, Moffitt, Caspi, Dickson, & Silva, 1996). These associations were magnified for girls with conduct problems, where girls with elevated externalizing concerns at age 15 left home significantly earlier, were five times more likely to have cohabitated with a partner, and four and a half times more likely to have given birth by age 21 than girls who did not have conduct problems at age 15 (Bardone et al., 1996). Other work has also identified a link between early antisocial behaviors and subsequent increased pregnancy rates, finding that girls in the top 10% of their cohort for antisocial behavior problems at age 8 were five times more likely to have become pregnant by age 18 than girls in the lower 50% of the cohort (Woodward & Fergusson, 1999). This was explained in part due to social and familial disadvantage, and in part due to a cascade of early-onset conduct problems prompting increased risk-taking in adolescence, which in turn predicted teenage pregnancy, as might be suggested with experiential canalization models.

Intimate partner relationship quality. In addition to the timing of relationship transitions, the quality of the bond with a long-term partner is also an important outcome. Girls with depression as well as girls with conduct problems at age 15 were significantly more likely to be the victim of intimate partner violence by age 20 than girls who had neither depression nor conduct problems at age 15 (Keenan-Miller, Hammen, & Brennan, 2007). Externalizing behaviors are perhaps particularly important; when assessed from ages 8-21, the presence of greater antisocial behaviors predicted greater conflict and relationship ambivalence later on, even above and beyond family socioeconomic status, parent-child relationship, interparental conflict, and children's cognitive ability (Woodward, Fergusson, & Horwood, 2002). Timing mattered: Childhood-onset antisocial behaviors predicted worse relationship outcomes (partner violence,

conflict, and relationship ambivalence) at age 21 than adolescent-limited antisocial behaviors, which provides additional support for the taxonomy proposed by Moffitt (1993).

Externalizing problems do not necessarily resolve when romantic relationships are initiated, and instead tend to persist over successive relationships (Rhule-Louis & McMahon, 2007), which might be especially true for those with childhood-onset antisocial behaviors. Romantic partners tend to be fairly matched in terms of their overall level of antisocial behaviors and substance use (Rhule-Louis & McMahon, 2007), which has generally been explained by self-selection or assortative mating frameworks, postulating that individuals choose environments and others with whom they feel comfortable. Romantic partners may mutually reinforce each other's behaviors, supporting the continuity and exacerbation of antisocial problems and substance use. Adolescents' substance use and antisocial behaviors may prompt partnering with individuals with similar behaviors, which increases risk of maintained or increased involvement in substance use and antisocial behaviors. Thus, we might expect that childhood externalizing problem behaviors may have a sustained association with decreased romantic relationship quality later on in adulthood, while internalizing problem behaviors may have a slightly weaker association.

Educational attainment. Longitudinal studies have found evidence for the contribution of externalizing and internalizing problem behaviors to educational attainment, with stronger and more consistent effects for externalizing behaviors. Conduct disorder and depression diagnoses at age 15 predicted lower educational attainment at age 21 (Bardone et al., 1996; Miech, Caspi, Moffitt, Wright, & Silva, 1999), and children age 6-8 with more mother-reported externalizing as well as internalizing problems were less likely to later graduate from high school, above and beyond children's race, age, gender, birth weight, maternal delinquency, maternal self-esteem,

maternal age at child birth, maternal education, maternal marital status, and poverty status (McLeod & Kaiser, 2004). Other work has identified that while substance use in adolescence and early adulthood predicted lower college completion, higher rates of broad externalizing behaviors in early adolescence was the strongest independent negative predictor of college completion (King, Meehan, Trim, & Chassin, 2006). These findings suggest that both externalizing and internalizing problem behaviors may have later connections to educational attainment.

Income. Educational attainment is often a gateway to later occupational status and income (Day & Newburger, 2002), but few studies have explicitly examined the connection between youth problem behaviors and later income. The presence of age-15 conduct disorder or age-19 substance use disorder has been connected to greater financial difficulties later on than those without these disorders (Bardone et al., 1996; Rohde et al., 2007), while a similar connection was not observed for those with depression at age 15 (Bardone et al., 1996). Tentatively, these results indicate that externalizing but not internalizing youth problem behaviors may be linked to later income.

Legal infractions. A large literature has highlighted the effect of earlier externalizing problems on later legal difficulties (e.g., Darney, Reinke, Herman, Stormont, & Ialongo, 2013; Hämäläinen & Pulkkinen, 1996). Disruptive behaviors in adolescence are particularly strong predictors of trouble with law enforcement in emerging adulthood (Bardone et al., 1996; Grella, Stein, & Greenwell, 2005; Vander Stoep, Weiss, McKnight, Beresford, & Cohen, 2002). Internalizing behaviors, on the other hand, are typically not related to later legal infractions (e.g., Bardone et al., 1996). However, the majority of this work has been done with boys, given their

higher absolute levels of externalizing problems, and has not assessed problem behaviors in early childhood, which may provide additional insight into their trajectories.

Alcohol use. A large body of work has found strong effects of earlier externalizing behaviors on later substance use, particularly for boys (e.g., Bardone et al., 1996; Chassin, Pitts, & Prost, 2002). This has also been found using the current MLS sample in a study that extended to ages 15-17 (Jester et al., 2008), finding that aggressive, inattentive/hyperactive, and combined latent classes had significantly earlier age of onset of drinking and drunkenness than the latent healthy class, indicating that earlier externalizing behavior predicted increased substance use in late adolescence. In addition, early externalizing behaviors predicted early adult (age 18-23) diagnosis of alcohol use disorder (Zucker, Jester, Puttler, Buu, & Sheddon, 2007).

Also in the MLS sample, Foster, Hicks, and Zucker (2018) found that the shared variance of internalizing and externalizing problems (e.g., negative emotionality, distress) predicted *increased* alcohol use problems from ages 9-11 through ages 21-23, while the unique aspects of internalizing problems (e.g., social withdrawal, inhibition) predicted *decreased* alcohol use problems. In a mixture of retrospective and prospective-longitudinal reports of alcohol use, Jester, Buu, and Zucker (2016) identified that adolescent antisocial symptoms were linked to the most severe alcohol use trajectories for both men and women within this current sample.

Across transition age, intimate partner relationships, educational attainment, income, legal infractions, and alcohol use problems, prior research has generally found connections between problems emerging earlier in development and distal outcomes (supporting experiential canalization) as well as the relation of adolescent functioning to early adulthood outcomes (supporting adolescence as a sensitive period). These connections appear to be particularly

strong for externalizing problem behaviors. However, whether the effect of these problems decays or is sustained at later periods in development remains largely unexplored.

The Importance of Informant

The second primary goal of this study was to examine whether different conclusions would be reached based on the reporter of these problems, given the low-to-moderate agreement between informants regarding the presence of problem behaviors (De Los Reyes et al., 2015). Two robust findings in informant discrepancies continually and predictably emerge. The first line of evidence is context: Informants yield higher agreement when they interact with youth in similar contexts (e.g., two parents at home, two teachers at school), as setting-specific factors may elicit divergent sets of behaviors. The second line of evidence concerns observability: More observable behavior yields higher informant agreement (e.g., aggressive behaviors vs. anxiety). Thus, researchers can make and test predictions as to where informants may be more likely to diverge (Achenbach, McConaughy, & Howell, 1987; De Los Reyes et al., 2015; De Los Reyes, 2011; De Los Reyes, Henry, Tolan, and Wakschlag, 2009).

Mean-level differences between informants have been consistently noted. Overall, parents tend to report more problems than teachers report, and adolescents tend to report more problems than parents report (Rescorla et al., 2013; Rescorla et al., 2014). Within internalizing and externalizing domains, informants may have differing sensitivities: teachers may be most attuned to externalizing behaviors, while adolescents themselves may be most conscious of their internalizing concerns (e.g., Berg-Nielsen, Solheim, Belsky, & Wichstrom, 2012; De Los Reyes & Prinstein, 2004; Rescorla et al., 2013; Zahner & Daskalakis, 1998). In the current sample, prior work has already identified substantial differences between informants, at least in early childhood, in a manner consistent with past literature suggesting that parents in this sample have

higher agreement for more observable behaviors (Bingham, Loukas, Fitzgerald, and Zucker, 2003). Also within this sample, parents reported higher levels of aggression than teachers, while teachers report higher levels of inattention than parents (Jester et al., 2005).

Parents: The depression-distortion hypothesis and neuroticism. When individuals are depressed, they tend to perceive others' behaviors more negatively than do individuals who are not depressed (Ohrt, Sjödin, & Thorell, 1999). This tenet, the *depression-distortion hypothesis* (Richters & Pellegrini, 1989; Richters, 1992), has been studied extensively in parents' reports of their children's behaviors, particularly for mothers. Multiple studies have supported this hypothesis in clinically-referred children, such as in reporting symptoms of children with ADHD (Chi & Hinshaw, 2002) and in reports of internalizing behaviors for boys in active psychotherapy (Kroes, Veerman, & De Bruyn, 2003). This distortion may also extend to fathers: Treutler and Epkins (2003) found that discrepancies between mothers and fathers regarding their children's problem behaviors varied as a function of each parent's broad symptomology. While the mechanisms through which depression affects parents' ratings of their children continue to be debated, a reasonable body of work has emerged to suggest that parental depression may impact reports of their children due to predictable differences in perception when compared to non-depressed parents.

Parents' trait neuroticism may also be a characteristic impacting parent reporting. Neuroticism is a relatively stable personality trait that reflects individuals' tendencies to react with negative emotions to frustration, loss, or threat (Watson & Clark, 1984; Watson & Clark, 1992). Psychopathology is associated with elevated levels of neuroticism (Khan, Jacobson, Gardner, Prescott, & Kendler, 2005; Kotov, Gamez, Schmidt, & Watson, 2010; Saulsman & Page, 2004). While depressive symptomology may represent a relatively specific

indicator of bias with a lower base rate, the relatively wider variability in neuroticism may provide a broader marker affecting parents' reporting. Limited empirical work has examined how higher levels of neuroticism impact parents' assessment of their children's problem behaviors. Neuroticism has been associated with a negative memory bias, which may be one mechanism through which neuroticism impacts parents' recall of problem behaviors (Chan, Goodwin, & Harmer, 2007; Rusting, 1999).

For the current study, we can examine how parents' reports of children's problem behaviors across childhood and adolescence predict distal outcomes when parent depressive symptoms are and are not taken into account, and in a similar way we can examine if the strength of this prediction changes after accounting for trait neuroticism. If parents' reports are influenced by their own mental health, we can examine this as a potential confound that could be accounting for the associations between early problem behaviors and later life outcomes.

Teachers. Teachers have experience with a much larger number of children than parents, which provides a set of norms to guide their perception of typical child behavior. However, the school context may also select for increased attention to externalizing behaviors at the expense of internalizing problems, given that externalizing behaviors tend to be more visible and disruptive to the classroom (e.g., Zahner & Daskalakis, 1998; Berg-Nielsen et al., 2012). Of note, this finding has not been universal, with some indication that teachers may be more likely than parents to report internalizing symptoms in children when children themselves also report concerns (Mesman & Koot, 2000).

Children. Adolescents consistently report significantly more problems than their parents (Rescorla et al., 2013). Youth psychopathology appears to impact how youth report on their own concerns; youth with diagnoses of ADHD tend to underreport problems relative to other

informants, and youth with more depressive symptomology tend to overreport problems relative to other informants (De Los Reyes, Goodman, Kliewer, & Reid-Quinones, 2008, 2010; De Los Reyes & Prinstein, 2004; Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007). Thus, overall, adolescents' privileged access to their own internal experiences may be quite important in understanding subsequent outcomes and functioning, although this may be impacted by youth overreporting or underreporting relative to other informants based on their presenting concerns.

Hypotheses

The current study seeks to answer the following questions: Do children's internalizing and/or externalizing problem behaviors across childhood predict distal life functioning outcomes? If so, does the strength of these predictions vary by who is reporting the behaviors, including mothers, fathers, teachers, and children themselves?

To answer these questions, the current study will utilize internalizing and externalizing problem behavior measures, assessed by parents from early childhood through late adolescence (T1-T5); by teachers from middle childhood through late adolescence (T2-T5); and by children from pre-adolescence through late adolescence (T3-T5). The distal outcomes will reflect outcomes in adulthood from ages 24-32, including the timing of normative adult transitions (cohabitation age, marital age, age at first child), romantic relationship quality, legal infractions including arrest, educational attainment, income, and alcohol use problems, utilizing a high-risk sample that provides greater variability on these measures than a typical community-based sample of families.

Based on the literature reviewed thus far, the following hypotheses are proposed:

Hypothesis 1: Higher initial levels of children's externalizing problems and/or greater escalation (or less pronounced decrease) of externalizing problems over time will predict an

earlier age at all transitions (cohabitation age, marital age, age at first child), lower romantic relationship quality, lower levels of later educational attainment, lower income, more legal infractions, greater likelihood of arrest, and more alcohol use problems.

Hypothesis 2: Higher levels of children's internalizing problems and/or greater escalation (or less pronounced decrease) of internalizing behaviors over time will predict earlier age at cohabitation, lower romantic relationship quality, and lower educational attainment.

Hypothesis 3: I predict that teachers' reports of externalizing behaviors will be a stronger predictor of distal outcomes than parents' reports, while teachers' reports of internalizing behaviors will fail to predict distal outcomes. I tentatively hypothesize that adolescents' reports of internalizing concerns will be a stronger predictor of later outcomes than parents' reports. Given a lack of clear prior theoretical or empirical rationale, I do not make specific hypotheses regarding effect size comparisons of parents' and adolescents' externalizing reports to later outcomes. Similarly, I do not make specific hypotheses regarding mothers' and fathers' reports; while mothers are typically conceptualized as the primary caregiver to their children and may be more sensitive to problem identification, fathers might have a higher threshold for reporting problems or may be more likely to tolerate gender-stereotyped behaviors such as aggression, which in turn may have a stronger signal to long-term outcomes (e.g., Lindsey, Mize, & Pettit, 1997; Leve & Fagot, 1997; Kerr, Lunkenheimer, & Olson, 2007).

Hypothesis 4: Parents' depressive symptoms and/or neuroticism may impact their assessment of children's problem behaviors, which in turn affects their prediction of distal outcomes. Prior literature has shown that psychological difficulties for both mothers and fathers contribute to discrepancies between their partner and their child in reporting their children's problem behaviors (Treutler & Epkins, 2003). Thus, I hypothesize that when parents' depressive

symptoms or neuroticism are controlled for, reports of children's problem behaviors will become more strongly related to distal outcomes (i.e., effect sizes will be higher once depressive symptoms and neuroticism are factored out).

Hypothesis 5: Child sex and cognitive ability as well as parental education level and income will also explain a portion of the variance in later distal outcomes, and in some cases, carry greater weight towards the distal outcome than problem behaviors. Specifically, I predict that children with parents with a greater level of education and income as well as children with greater cognitive abilities will have higher levels of degree and grade as well as higher incomes themselves (e.g., Dubow, Boxer, & Huesmann, 2009). In addition, given the link between legal infractions and poverty (e.g., Hsieh & Pugh, 1993; Kingston & Webster, 2015), I predict that children with parents with higher income will have fewer later legal infractions and a lower likelihood of arrest. Finally, because base rates indicate that males on average are more likely to experience arrest, commit legal infractions, and have higher alcohol use as well as alcohol-related problems than females (e.g., Heimer, Lauritsen, & Lynch, 2009; Nolen-Hoeksema, 2004; Steffensmeier, Schwartz, Zhong, & Ackerman, 2005), I predict that child sex will explain a significant portion of the variance in these outcomes.

METHOD

Participants and Procedure

The present study is part of the multi-wave Michigan Longitudinal Study (Zucker, Ellis, Bingham, & Fitzgerald, 1996; Zucker et al., 2000). In the initial recruitment process, fathers of sons aged 3-5 were recruited based on drunk-driving convictions in a four-county area. Fathers were required to have a blood alcohol concentration (BAC) of at least .15% if the conviction was their first drinking-related legal problem or a .12% BAC if any prior drinking-related legal problems had occurred. In addition, all men were required to meet a diagnosis for probable or definite alcoholism, which was based on the Feighner criteria (Feighner et al., 1972); since initial recruitment, all participants have been re-diagnosed through DSM-IV alcohol-use disorder criteria. Fathers were required to live with their son's biological mother at the beginning of the study, and although mother alcohol status could be variable, presence of fetal alcohol syndrome in the child was exclusionary. While having a son between the ages of 3-5 years was a requirement for parent inclusion, there were no restrictions on parent age. Based on the receipt of additional grant funding, two additional samples were recruited; the third sample began when children were aged 6-8.

As a control/contrast lower-risk sample, families with age-matched children living in the same neighborhood were recruited. Control families were required to have no lifetime parental history of substance abuse or dependence. Through this door-to-door canvassing process, an intermediate risk group was identified through recruitment of neighborhood control families with fathers who met criteria for alcoholism. For additional details on study characteristics and sample, see Zucker et al. (2000).

After the first wave of data collection, dependent on wave and available funding, full biological siblings of the 3- to 5-year-old sons were also recruited if they were within 8 years of age of the son. Beginning at age 18, youths' romantic partners were invited to participate if they had been cohabitating the partner for at least 9 months.

The data on youth problem behavior includes the first five waves of the study, encompassing ages 3-17, as well as three waves in early adulthood. The childhood waves were captured when children were in early childhood (time 1, T1; age 2-5), middle childhood (time 2, T2; age 6-8), pre-adolescence (time 3, T3; age 9-11), early adolescence (time 4, T4; age 12-14), and mid-to-late adolescence (time 5, T5; age 15-17). Waves in adulthood were used to assess distal outcomes (times 8, 9, and 10, T8-T10; age 24-32). See Figure 1 for an assessment schematic.

Data collection typically occurred three years apart. In total, 1,069 children (749 boys, 70%) from 482 families were included in analyses. A total of 978 parents or stepparents (479 mothers, 499 fathers) reported on their children at least once (greater than the number of families due to stepparents). Of biological parents reporting at least once on their children, 479 mothers and 475 fathers participated. Children began reporting on their own problem behaviors at age 11; between the ages of 11 and 17, 993 children reported on their own problem behaviors at least once. For distal outcomes, assessed at ages 24-32, 565 participants completed at least one measure. While there was some participant drop-out, incomplete data is due in large part to participants entering the study at different ages and times, with many having not yet reached age 24.

Trained project staff was blind to diagnostic and recruitment status and conducted primarily in-home assessments every three years. Assessment batteries varied based on time

point but generally included questionnaire completion, diagnostic interviewing, and neuropsychological testing.

Measures: Predictor Variables

Children's internalizing and externalizing behaviors. Maternal and paternal ratings of children's internalizing and externalizing problems were measured by the Child Behavior Checklist (CBCL; Achenbach, 1991). Each item is rated on a 0 to 2 scale, with 0 indicating the item is not true of the child within the past 6 months, 1 indicating the item is sometimes or somewhat true, and a 2 indicating the item is very true or often true. The CBCL is a 113-item measure designed to identify externalizing problem behaviors (aggressive, rule-breaking, and inattention problems) as well as internalizing problem behaviors (anxious, depressed, withdrawn, and somatic problems). These scales were calculated separately, and have been shown to have high reliability and validity (Achenbach & Rescorla, 2001). Mothers and fathers completed the CBCL at the first five waves of data collection. Teachers completed the teacher-report version of the CBCL, the Teacher Report Form (TRF; Achenbach, 1991), at the second through fifth waves of data collection. Finally, children began completing the Youth Self-Report (YSR; Achenbach, 1991) at age 11, providing reports at the third, fourth, and fifth waves of data collection. Many youth also completed the YSR during annual mailings; when youth did *not* complete the three-year data collection wave, if annual YSR data was available, it was used rather than considering the data to be missing. Similarly, if teacher reports were not available for the three-year data collection wave, the earliest annual report within the age range was used when available.

Table 1 provides the *N*s, means, and standard deviations for the CBCL (N range=463-643 per time point), TRF (N=409-732 per time point), and YSR (N=428-894 per time point). For the

CBCL, Cronbach's alpha calculations revealed good to excellent internal consistency at the scale level (externalizing α=.88 (T2) to .92 (T5); internalizing α=.82 to (T1) to .87 (T3, T5)) and moderate to good internal consistency at the subscale level (withdrawn α=.62-.76, somatic α=.63-.71, anxious α=.73-.82, delinquent α=.57-.82, aggressive α=.87-.90). For the TRF, Cronbach's alpha was similarly good to excellent (externalizing α=.94 (T5) to .95 (T2-T4); internalizing α=.88 (T2) to .90 (T4-T5)) and good internal consistency at the subscale level (withdrawn α=.81-.86, somatic α=.63-.78, anxious α=.84-.89, delinquent α=.70-.80, aggressive α=.94-.95). Finally, for the YSR, scales had good internal consistency (externalizing α=.85 (T3) to .88 (T5); internalizing α=.89 (T5) to .90 (T3-T4)) while subscales had variable levels of internal consistency (withdrawn α=.65-.72, somatic α=.71-.75, anxious α=.85-.87, delinquent α=.51-.73, aggressive α=.84-.86). Overall, 967 children had at least one completed mother-reported CBCL; 917 children had at least one completed father-reported CBCL; 952 had at least one teacher-reported TRF; and 993 children had self-reported concerns at least once on the YSR.

Control variables. For each problem behavior growth model that significantly predicted a distal outcome, a series of control variables were tested to determine if the significant pathways were maintained above and beyond potentially relevant covariates. See below for details of each covariate.

***Children's cognitive ability*.** Children were administered a measure of cognitive functioning during the second, third, or fourth wave of data collection, as new children were recruited to the study. For children in Sample 1, the Wechsler Intelligence Scale for Children – Revised (WISC-R; Wechsler, 1974) was administered; for children in Samples 2 and 3, the Wechsler Intelligence Scale for Children – 3rd Edition (WISC-III; Wechsler, 1991) was administered. Some children received an abbreviated measure at T4 that included Information,

Block Design, Vocabulary, and Symbol Search. Each cognitive ability measure provided a normed full-scale IQ (FSIQ), which was used. While there were no FSIQ differences by time point when collapsing across samples and measures (N=619, M=104.79 at T2; N=106, M=106.52 at T3; and N=534, M=106.02 at T4), there were some differences in cognitive ability based on the cognitive assessment. That is, children who received the WISC-R obtained the highest average scores (N=972, M=106.82, SD=13.48), followed by the full WISC-III (N=287, M=100.83, SD=13.80), followed by the abbreviated WISC-III (N=275, M=98.36, SD=15.34). Counts provided are higher than the number of children because many children received a cognitive assessment at more than one time point; in these cases, an average FSIQ was taken on the basis that an average would provide the most accurate estimate of cognitive functioning. In total, 835 children had FSIQ data, M=104.70, SD=13.67.

Parental education level and household income. Maternal and paternal education level as well as household income was obtained from the earliest available wave of a demographic questionnaire. Education level was assessed by the number of years of education (mothers N=454 (of 979 children), M=13.34, SD=2.06; fathers N=439 (of 878 children), M=13.49, SD=2.25). Fathers had a slightly higher level of education, $t(857)=3.29$, $r=.52$, $p=.001$. In addition, household income was provided by both mothers and fathers, which yielded similar estimates ($r=.85$, $p<.001$), with 473 mothers (of 1017 children) and 481 fathers (of 951 children) providing income information. The average household income bracket reported by both mothers and fathers was \$20,000-\$30,000/year (N=82 families); 56% of the sample reported their annual income to fall within \$16,000 to \$50,000, and household incomes ranged from below \$4,000/year (N=11 families) to over \$100,000/year (N=4 families).

Parental depressive symptomology. Mothers and fathers reported on the Beck Depression Inventory at each of the five waves across their offsprings' childhood (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The Beck Depression Inventory (BDI) is a 21-item, self-report measure that assesses symptoms of depression on a 0-3 scale, and displays high internal consistency for both psychiatric and non-psychiatric populations (Beck, Steer, & Carbin, 1988). Total scores are determined through summing. Although most samples at most time points received the 21-item BDI, a 13-item abbreviated version was also utilized at some time points; because scoring occurs through sums, BDI scores were z-scored within informant at each time point (e.g., maternal BDI score at T1 was z-scored, paternal BDI score at T1 was z-scored, etc.) and then averaged across the first five time points to produce an overall average BDI score for each mother (N=473, for 1016 children) and father (N=483, for 954 children). Cronbach's alpha ranged by time point and sample, but overall demonstrated good internal consistency (α=.76-.90, mean α=.86, median α=.87). Paired-samples *t*-tests indicated that mothers reported greater depressive symptomology than fathers, $t(946)=-7.38$, $r=.36$, $p<.0001$.

Parental neuroticism. Mothers and fathers self-reported on their trait personality, including neuroticism, through the Neuroticism-Extraversion-Openness Inventory (Costa & McCrae, 1985) and the NEO Personality Inventory – Revised (NEO-PI-R; Costa & McCrae, 1992). The neuroticism scale included items assessing anxiety, hostility, depression, self-consciousness, impulsivity, and vulnerability to stress. The NEO was administered at the second and third waves of data collection while the NEO-PI-R was administered at the fourth and fifth waves of data collection. Due to measure and scale differences, neuroticism was z-scored within each time point by informant (e.g., maternal neuroticism at T1 was z-scored, paternal neuroticism at T1 was z-scored, etc.) and then averaged within each person across the five waves

to produce an overall average neuroticism for each person, N=436 fathers (for 901 children), N=447 mothers (for 989 children). Cronbach's alpha indicated good consistency at each time point (α=.84-.87). Paired-samples *t*-tests indicated that mothers reported higher neuroticism than fathers, $t(890)=-9.58$, $r=.14$, $p<.0001$.

Parental antisocial personality disorder symptoms and alcohol use disorder symptoms.[1] To provide information into parents' externalizing behaviors, two clinical measures were used. At each in-person assessment visit, assessors administered the Diagnostic Interview Schedule, Version III (DIS-III; Robins, Helzer, Croughan, & Ratcliff, 1981) to both mothers and fathers. This produced a scale for each disorder, with 0=negative for disorder, 1=recovery without symptoms in the past year, 2=partial recovery with symptoms in the past year, and 3=active disorder in the past year. Thus, the average of parents' scores across each time point was obtained to provide an average ASPD score as well as an average AUD score, with a higher score indicating greater symptomology in that area. Overall, 389 mothers (of 894 children) had an ASPD average, $M=.25$, $SD=.59$, range=0-3, while 383 mothers (of 883 children) had an AUD average, $M=.15$, $SD=.46$, range=0-3. Similarly, 383 fathers (of 818 children) had an ASPD average, $M=.33$, $SD=.65$, range=0-3, while 375 fathers (of 800 children) had an AUD average, $M=.47$, $SD=.86$, range=0-3. While paired-sample *t*-tests did not indicate a significant difference between parents' ASPD symptoms ($r=.36$, $p<.001$), fathers reported significantly more alcohol use concerns than mothers, $t(774)=9.97$, $r=.14$, $p<.001$.

[1] The externalizing scale of the Adult Self-Report Scale (Achenbach, 1997) was planned for use as a more broad-band measure of externalizing concerns. However, when conducting statistical analyses, the ASR did not provide sufficient data coverage (mother $N=144$, father $N=151$) and resulted in convergence errors, yielding decreased confidence in findings.

Parental marital quality. Mothers and fathers completed the Dyadic Adjustment Scale (DAS; Spanier, 1976), which is a 32-item measure of relationship quality. There are four subscales, including dyadic consensus (degree to which partners agree), dyadic satisfaction (degree to which the informant feels satisfied with their partner), dyadic cohesion (degree to which the informant and partner participate in activities together), and affectional expression (degree to which the informant agrees with their partner about emotional affection). The total score was determined by summing the four subscales, as is conventionally done, and then averaged across time points. Parents varied on how many times they completed the DAS (one time N=235, twice N=427, three times N=58, four or more times N=113). Overall, 391 mothers completed the DAS at least once (M=107.67, SD=18.49) as well as 380 fathers (M=110.18, SD=15.99). Cronbach's alpha was similar at each time point and had adequate to good internal consistency (α=.76-.82). Fathers reported slightly higher marital quality than mothers reported, $t(783)=5.09$, $r=.65$, $p<.0001$.

Measures: Outcome Variables

To determine outcome variables, information provided from the eighth through tenth waves of data collection (i.e., T8-T10, ages 24-32) was utilized. T8 encompasses ages 24-26; T9, ages 27-29; T10, ages 30-32. On years when young adults were not completing the larger in-person assessments, they were mailed questionnaires, up to age 27. Thus, for some measures, there were annual measures for age 24, 25, 26, and 27 (termed A24-A27) that were used to supplement missing data from the T-periods. The goal was to achieve the most comprehensive estimates of early adulthood functioning, with one value for each distal outcome. It was not uncommon for participants to participate in just one of these waves of data collection, precluding the ability to choose one specific age or time point without losing substantial valuable data.

However, as expected, this strategy also resulted in many participants providing multiple waves of data. While for some measures the variable value would not change (e.g., age at birth of first child), others would (e.g., legal infractions in the prior 3 years). The specific approach used to assess each of these variables has been detailed below. Correlations between distal outcomes are detailed in Table 10.

Timing of normative transitions. Each of these variables was assessed through a demographic questionnaire that was completed at T8, T9, and T10, as well as each intermediate annual time point. If there were discrepant reports across time points (e.g., at three time points, reported marital age to be 18, 18, and then 19), the age that was more consistent across time points was used (i.e., 18 in this case). If there were equal numbers of data points (e.g., two), the younger age was utilized on the assumption that the milestone had occurred more recently to the age at which it was first reported, resulting in more trustworthy recollection.

Cohabitation. In total, 231 respondents reported an age at which they first began living with a romantic partner for at least 9 months, M=22.78, SD=2.76, range=15-31. Males (M=23.06, SD=2.75) reported a slightly older age at first domestic partnership than females (M=22.21, SD=2.70), $t(229)=2.25$, $p<.05$.

Marital age. In total, 201 respondents reported a marital age, M=23.94, SD=2.42, range=18-29. There was no significant marital age difference between males and females. Few participants identified experiencing divorce, precluding examination of this potentially important experience.

Age at first child. In total, 167 respondents reported their age at the time of the birth of their first child, M=22.62, SD=3.28, range=14-30. There was no significant age difference between males and females.

Intimate partner relationship quality. Partners living with the target adult offspring for at least 9 months were recruited to participate in the study and completed the Dyadic Adjustment Scale (DAS; Spanier, 1976), which is a 32-item measure of relationship quality. As was used with offsprings' parents, the total summed DAS score was used as an overall measure of agreement, satisfaction, cohesion, and emotional connectedness, with a higher DAS score indicating better functioning and quality. The lowest total sum provided by an informant at any time point was utilized to provide a maximized estimate of risk; in addition, DAS scores were not eligible to be averaged due to some respondents changing partners, crossing informants. This measure was completed at T8-T10 as well as A24-27, for a total of 299 informant reports, *M*=110.90, *SD*=18.26, range=42-148. Cronbach's alpha demonstrated adequate to good internal consistency at each time point (α=.75-.85). Informant-reported relationship quality did not differ based on whether the target was male or female.

Educational attainment in years and degrees. Educational attainment was defined through total number of years of education as well as by degree obtainment (*r*=.84 between these two measures). For both measures, the highest value (i.e., most education) was utilized, unless the highest value provided was not consistent with later-provided data. For example, if time point values were 14, 14, 17, 14, and 14, then 14 years of education would be used. However, if time point values were 14, 14, 14, 14, and 17, then 17 years of education would be used on the theory that the participant obtained three more years of education between these waves of assessment. This same formula was applied to determining the highest degree achievement. In total, 530 participants reported their number of years of education (*M*=14.39, *SD*=2.26, range=7-20).

Second, participants (N=514) indicated the highest degree they had achieved, with categories including "no degree, or graduated high school/completed GED," "obtained technical, vocational, or associate's degree," "obtained bachelor's degree," and "obtained MA/MS," and "obtained PhD/MD/JD." In total, 256 participants (49.8%) obtained either no degree, graduated high school, or obtained a GED; 56 participants (10.9%) obtained a technical, vocational, or associate's degree; 161 participants (31.3%) obtained a bachelor's degree; 30 participants (5.8%) obtained a master's degree; and 11 participants (2.1%) obtained a PhD, MD, or JD. Degree attainment was treated as a continuous variable due to its ordered nature with several categories and more interpretable fit statistics, although all analyses were also run with degree attainment specified as a categorical variable, which yielded identical patterns of significance.

Paired-sample *t*-tests were conducted to determine if increased educational attainment was due to participation at later waves. For degree obtainment, no significant differences were observed between T8 (which was the earliest time point as well as the time point supplying the most data) and other T-periods and annuals. However, for number of years of education, participants who had T8 and T9 data had slightly more average years of education at T9 (M=14.14, SD=2.26 at T8; M=14.40, SD=2.54 at T9), $t(307)=-3.67$, $p<.001$. Similarly, participants who had T8 and T10 data had slightly more average years at T10; however, participants with T9 and T10 data did not differ significantly. There were no differences between T8 and any of the annual reports.

Individual and family income. Income was operationalized as individual income as well as the combined income of the individual and their partner. A series of income categories were provided (ranging from "under $4,000" to "over $100,000"), and respondents indicated substantial spread in their income level that was relatively normally distributed. For both

measures, the highest value (i.e., highest income) was utilized to maximize assessment of earning potential. Both income values (individual and family) were utilized due to the possibility of a homemaker occupation (individual income under $4,000) despite an elevated socioeconomic position (e.g., family income over $100,000). However, while family income may be a more comprehensive assessment of socioeconomic standing than individual income, fewer participants (N=340) provided information about family income relative to individual income (N=526), as many participants had perhaps not yet become partnered; therefore, both variables were retained. As with education, due to the ordered nature and number of categories, these variables were treated continuously. Males and females did not report significant differences in individual or family income.

The vast majority of income data (both individual and family) came from T8 and T9, particularly T8; paired sample *t*-tests were conducted to determine if participants had higher incomes as they aged, which would suggest that participants who participated in later time points may have elevated incomes relative to participants who did not have these time points. However, this test was nonsignificant for both individual and family incomes. Paired-sample *t*-tests were also conducted between T8 and T10 as well as T8 with A25 and A26. The only significant *t*-tests suggested that earnings (both individual and family) at T8 were significantly higher than earnings at T10, reducing potential concern that participants who had "aged in" to later periods would contribute higher incomes due to age.

For individual income, the average earnings fell in the $16,000-$20,000 category, while the median earning category was $20,000-$30,000. Overall, 58.0% of respondents (*N*=305) reported an individual income spanning $16,000-$50,000/year. For family income, the average earnings fell in the $30,000-$50,000/year category, while the median earning category was

$50,000-$75,000. Overall, 59.1% of respondents (*N*=201) reported an family income spanning $20,000-$75,000/year.

Legal infractions. Two measures of legal infractions were obtained. First, a demographic questionnaire asked one yes/no question regarding experiencing arrest in the past three years, and this was administered at T8-T10 as well as A24-A27. Respondents (*N*=532) were divided based on consistently denied prior arrest in the last three years vs. one affirmative response indicating prior arrest; in total, 16.5% of respondents had experienced arrest (*N*=88) and 83.5% had not (*N*=444). Significantly more males than females experienced arrest, $t(530)=3.39$, $p=.001$.

Secondly, crime and legal infractions were assessed through items on the Antisocial Behavior Checklist (ASB; Zucker & Noll, 1980). This questionnaire asked each respondent to rate whether they had "never," "rarely" (once or twice), "sometimes" (three to nine times), or "often" (more than ten times) done each behavior in the prior three years. Specific questions assessing direct interaction with law enforcement were used, with each response added together (0=never, 1=rarely, 2=sometimes, 3=often). Items included "Received a speeding ticket," "Been arrested for a felony," "Resisted arrest," "Been arrested for any other non-traffic police offenses (except fighting or a felony)," and "Been convicted of any non-traffic police offense." Thus, each participant's score could range from zero (endorsed no items) to 15 (endorsed all items as "often"). This measure was completed at T8-T10, and the highest ASB value from any time point was used; paired sample *t*-tests indicated no significant differences across time. Overall, 540 respondents completed this measure, $M=.78$, $SD=1.18$, range=0-10. As with prior arrest, males reported significantly more legal infractions than females, $t(538)=2.85$, $p=.005$.

Alcohol use. Amount of alcohol use as well as alcohol-use problems were utilized as two measures of involvement with alcohol, and were rated at T8-T10 as well as A24-A27. For both measures, in order to maximize risk, the greatest amount of alcohol use and greatest number of drinking problems reported were utilized in analyses, regardless of time point (r=.61 between these measures). Alcohol use was assessed with a single item from the Drinking and Drug History Form that asked respondents to report the maximum number of drinks they had ingested within a 24 hour period within the last three years (Zucker, Fitzgerald, & Noll, 1990). Participants who reported 30 or more drinks within a 24 hour period were considered to have ingested 30 (i.e., range=0-30). In total, 540 respondents provided their maximum number of alcoholic drinks within 24 hours, M=12.70, SD=7.92. Males reported consuming significantly more drinks than females, $t(538)=8.30$, $p<.001$.

Next, alcohol-use problems were assessed from a scale in the Drinking and Drug History Form that listed 31 potential problems as a result of alcohol use in the prior three years (e.g., "Had days where I drank much more than I expected when I began," "Restricted my drinking to certain times of day or week in order to control it or cut down," "Been arrested for a drinking related offense."). Problems endorsed (regardless of frequency or severity) were added together to create a continuous variable. In total, 542 respondents reported alcohol-related problems, M=3.91, SD=4.57, range=0-23. Cronbach's alpha demonstrated good to excellent internal consistency at each time point (α=.85-.94). Males reported experiencing significantly more alcohol-use problems than females, $t(538)=4.80$, $p<.001$.

Data Analytic Strategy

Following the examination of descriptive statistics, correlations between variables, and intraclass correlation coefficients, latent growth curve modeling (LGCM) was used to evaluate

intraindividual changes as well as interindividual variations contained within intraindividual change patterns (Baltes & Nesselroade, 1979; Grimm & Ram, 2012; Singer & Willet, 2003). All analyses were conducted using Mplus Version 7.4 (Muthén & Muthén, 2015) with full information maximum likelihood estimation for handling missing data. To evaluate overall model fit, several fit statistics were examined including the comparative fit index (CFI; Bentler, 1990), the standardized root mean square residual (SRMR; Hu & Bentler, 1999), the root mean square error of approximation (RMSEA; Kline, 2011), and the χ^2 index. CFI values above .90 and RMSEA and SRMR values below .08 traditionally indicate acceptable model fit (Hu & Bentler, 1999). To account for the non-independence of data (due to siblings living in the same family), which can artificially inflate standard error estimates, the CLUSTER command was used, which adjusts the standard errors for non-independence using the Huber-White correction (e.g., as done in Martel et al., 2009). Children were included if they had at least one wave of childhood data from one informant (i.e., participants with a distal outcome but no childhood problem behavior data were *not* included, while participants with problem behavior data but no distal outcomes *were* included).

The present analyses were conducted in three stages. First, the optimal change models were established independently for mother-, father-, teacher-, and self-reported child externalizing problem behaviors, and separately for child internalizing problems. That is, separate no-change and change models were fit for maternal child externalizing behaviors, maternal child internalizing behaviors, paternal child externalizing behaviors, paternal child internalizing behaviors, teacher child externalizing behaviors, teacher child internalizing behaviors, self-reported externalizing behaviors, and self-reported internalizing behaviors; these were done separately in order to determine if informant source may provide different change

patterns. As a reminder, measures of problem behaviors were collected over five waves of data collection separated by three years for mothers and fathers. These waves corresponded to children's developmental periods (early childhood, middle childhood, pre-adolescence, early adolescence, and mid-to-late adolescence). Teachers began reporting in middle childhood, resulting in four waves of data collection (middle childhood, pre-adolescence, early adolescence, and mid-to-late adolescence). Children began self-reporting on their problem behaviors in pre-adolescence for three waves of data (pre-adolescence, early adolescence, and mid-to-late adolescence). Change models tested varied based on the number of time points present. In addition, in order to provide standardized, matching data estimation of parents and teachers to the 3-wave structure of children, models spanning the pre-adolescent through mid-to-late adolescent periods were also fit for mothers, fathers, and teachers, resulting in full models (e.g., a mother-reported T1-T5 externalizing model) and abbreviated models (e.g., a mother-reported T3-T5 externalizing model) for parents and teachers.

As an example, consider the full model for maternal child externalizing behaviors, which had five time points, where a no-growth intercept-only model was first fit (intercept mean, intercept variance, and residual variance for each time point). Next, a linear growth model was fit (intercept and slope means; intercept and slope variances; their covariance; and the residual variances). To reduce multicollinearity, the median time point was set to be 0 with surrounding time points set at equal distance (e.g., for the five time points for maternal child externalizing behavior, time points were set to -2, -1, 0, 1, and 2). A quadratic growth model was tested next, with estimated intercept, linear, and quadratic means; intercept, linear, and quadratic variances; the covariances between the intercept and linear slope, intercept and quadratic slope, and linear and quadratic slopes; and the residual variances. Finally, a cubic growth model was tested,

including estimations of the intercept, linear, quadratic, and cubic means; intercept, linear, quadratic, and cubic variances; the covariances between the intercept and linear slope, intercept and quadratic slope, intercept and cubic slope, linear and quadratic slopes, linear and cubic slopes, and quadratic and cubic slopes; and the residual variances. Autoregressive structures were not included, as there is currently debate about whether including autoregressive paths can bias estimates when growth parameters have been modeled (Grimm & Widaman, 2010; Sivo, Fan, & Witta, 2005). In total, the following data were each fit to separate growth models: (1) mother-reported child externalizing problems for T1-T5; (2) mother-reported child externalizing problems for T3-T5; (3) father-reported child externalizing problems for T1-T5; (4) father-reported child externalizing problems for T3-T5; (5) teacher-reported child externalizing problems for T2-T5; (6) teacher-reported child externalizing problems for T3-T5; (7) child-reported child externalizing problems for T3-T5; (8) mother-reported child internalizing problems for T1-T5; (9) mother-reported child internalizing problems for T3-T5; (10) father-reported child internalizing problems for T1-T5; (11) father-reported child internalizing problems for T3-T5; (12) teacher-reported child internalizing problems for T2-T5; (13) teacher-reported child internalizing problems for T3-T5; and (14) child-reported child internalizing problems for T3-T5.

For the second stage, after deciding on the best-fitting models of change over time for each informant for each problem behavior type, the resulting 14 conditional latent growth curve models were tested with each of the 12 distal outcomes. As a reminder, distal outcomes included (1) cohabitation age; (2) marital age; (3) age at first child; (4) intimate partner relationship quality; (5) years of education; (6) educational degree; (7) individual income; (8) family income; (9) prior arrest; (10) legal infractions; (11) maximum alcoholic drinks consumed within 24

hours; and (12) alcohol-use problems. Each outcome was separately tested to determine whether the intercept and growth parameters of each problem behavior model predict that outcome. As an example, because mother-reported externalizing problems (T1-T5) emerged as quadratic, the intercept, slope, and quadratic growth parameters were tested as predictors of each distal outcome (e.g., age at first child, prior arrest, etc.). In this way, analyses shed light on whether (1) initial level or change in internalizing and/or externalizing behavior problems across childhood held predictive power for adulthood outcomes, and (2) if certain informants' reports were stronger indicators of these distal outcomes than other informants' reports through comparing amount of the variance accounted for in the outcome. Because prior arrest was a binary outcome, a Monte Carlo integration was specified; all other outcomes were treated as continuous. Of note, the covariance coverage was adjusted below the MPLUS default of .10 for a few outcomes, most notably the timing of transition variables (marital age, cohabitation age, and age at first child).

One question that emerged was whether certain time periods were especially important in predicting the distal outcome. That is, if predicting future legal infractions from childhood externalizing problems as reported by mothers, are elevations at particular times especially predictive? For example, are externalizing behaviors in early adolescence particularly predictive? Or late adolescence? To answer these questions, when at least one growth parameter significantly predicted an outcome, each T-period was tested in turn as a direct predictor of the outcome over and above the latent growth parameters.

Finally, in the third stage of analyses, for models that significantly predicted children's distal outcomes, a series of time-invariant covariates were included. For all predictions relying on maternal reports, the series included; (1) child cognitive ability; (2) child sex; (3) maternal

socioeconomic status (maternal income, maternal education); (4) maternal internalizing concerns (average depressive symptomology, average neuroticism); (5) maternal externalizing problems (average antisocial personality disorder symptoms, average alcohol use disorder symptoms); and (6) maternal report of marital quality. Models relying on paternal report followed a parallel structure (e.g., using paternal socioeconomic status, paternal internalizing concerns, paternal externalizing concerns, etc.). Predictions relying on child and teacher reports included (1) child cognitive ability; (2) child sex; and (3) family socioeconomic status (maternal education, paternal education, maternal income, paternal income). Covariates were specified as predicting the outcome and allowed to correlate with growth parameters, with the exception of child sex, which was specified as a predictor for both the outcome and growth parameters. As an example, if the latent intercept of child externalizing behaviors as reported by mothers predicted later drinking problems, covariates would be entered to determine if the significant relation between the intercept and drinking problems was maintained.

RESULTS

First, bivariate correlations (Tables 2-10) and intraclass correlation coefficients were examined. Next, the optimally-fitting latent growth curve models for each problem behavior and informant were determined, and are detailed in Tables 11-12 and displayed graphically in Figures 1-5. Distal outcomes were subsequently tested alone and with covariates with results included in Tables 13-25, while Tables 26-27 provide summary statistics and Figures 6-8 display select distal outcomes by informant reports of problem behaviors.

Bivariate Correlations and Comparisons

Associations between problem behaviors. As expected, there were mostly moderate to high correlations between problem behaviors across time within the same informant, with assessment periods occurring at closer points in time yielding higher correlations; see Tables 2-3. Correlations between time points were compared using Fisher's r-to-Z transformation (independent testing for non-overlapping time point comparisons, dependent testing for overlapping time point comparisons; Steiger, 1980; Lee & Preacher, 2013). The difference between both parent-reported T1-T5 problems (the time points furthest separated) was significantly greater than the difference between any two adjacent time points for both externalizing and internalizing problems, z-score range= 2.93-7.46, $p<.001$. There were no significant mother-father correlation differences between any of the same two time points for internalizing as well as externalizing problems.

For children, the difference between child-reported T3-T5 problems (the time points furthest separated) was significantly greater than the difference between any two adjacent time points for externalizing problems, z-score range= 3.49-4.10, $p<.001$, but only partially for internalizing problems; while the correlation between T3-T4 internalizing problems was

significantly lower than T3-T5, T4-T5 was not significantly different from T3-T5, z=1.38, *ns*. For teachers, the difference between T2-T5 problems (the time points furthest separated) was not significantly greater than the difference between any two adjacent time points for internalizing problems as well as most externalizing problems, with the exception that the T2-T3 externalizing correlation was significantly higher than the T2-T5 correlation, z-score=2.90, p<.01. Teacher-reported problem behavior correlations tended to be relatively lower than parent- or child-reported problems, which may be expected due to different teachers completing measures at each time point.

Thus, while there is notable rank-order stability at time points that were closer together in time, rank-order stability was significantly less when considered across multiple developmental periods, particularly for parent-reported problems and child-reported externalizing problems. This stability appeared less consistently for teachers, whose reports differed across a larger span of time almost as much as they did between time points that were closer together. Overall, consistent with prior literature, rank-order stability of problem behaviors were highest for parent reports (for mothers, T1-T5 mean *r*=.51, T3-T5 *r*=.58; for fathers, T1-T5 mean *r*=.52, T3-T5 *r*=.57), followed by children's self-reports (T3-T5 mean *r*=.40), and finally teacher-reports (T2-T5 mean *r*=.27, T3-T5 *r*=.28), averaged across both internalizing and externalizing domains.

To provide an overall estimate of within-person change over time, intraclass correlation coefficients (ICCs) were computed to quantify the proportion of the total variance attributable to between person (as opposed to within-person) differences. In all intercept-only models, the variance component was significant. For each informant, ICCs were larger for externalizing problems than internalizing problems. Externalizing problem behavior ICCs were as follows: mothers=.60, fathers=.60, teachers=.41, and children=.48. For internalizing behaviors, ICCs

were as follows: mothers=.48, fathers=.47, teachers=.19, and children=.37. The lower ICC value for teachers is expected given that a different teacher is reporting on the child at each time point, allowing for greater within-child differences potentially attributable to rater effects across teachers. Using father-reported externalizing problems as an example, the ICC value of .60 indicates that 60% of the variance in externalizing problems occurs within children while the remaining 40% of the variance is between children. Alternatively, for teacher-reported internalizing problems (ICC=.19), 19% of the variance in internalizing problems is explained within children and 81% of the variance is between children.

Across informants, there were varying levels of agreement at the same time point; see Tables 4-7. Between mothers and fathers, there were moderate correlations for child externalizing problems and weak to moderate correlations for internalizing problems at the same time point; at each point, mother-father externalizing correlations were significantly higher than mother-father internalizing correlations, z-score range=3.15-6.35, $p<.001$, indicating significantly greater agreement for externalizing concerns than internalizing concerns across time.

Mother-teacher and father-teacher correlations were roughly similar, with externalizing problem correlations ranging from weak to moderate and internalizing child problems falling typically in the weak range. At each time point, parent-teacher agreement for externalizing concerns was significantly higher than parent-teacher agreement for internalizing concerns (z-score range=2.30-4.17, $p<.05$), with the one exception that father-teacher agreement of internalizing vs. externalizing concerns at T2 did not differ significantly (z-score=1.58, *ns*). Overall, mother-father agreement of child externalizing behaviors was significantly higher at each time point than parent-teacher agreement, z-score range=2.65-7.27, $p<.01$, indicating that

parents, who share a context in observing children, agree more about children's externalizing behaviors than teachers, who observe children in an academic context. However, there was a less consistent picture for agreement of child internalizing concerns. Mother-father agreement of child internalizing problems was often but not always higher than parent-teacher agreement: At T2, mother-father agreement was significantly higher than parent-teacher agreement; at T3, mother-father agreement was higher than father-teacher agreement but not mother-teacher agreement; at T4 and T5, mother-father agreement was higher than mother-teacher agreement but not father-teacher agreement.

Mother-child and father-child correlations for externalizing problems were also roughly similar, and ranged modest to moderate depending on time point, while mother-child and father-child correlations for internalizing problems were typically weak. While mother-father agreement was significantly higher at each time point than parent-child agreement for externalizing behavior concerns (z-score range=3.97-6.11, $p<.001$), this was largely not the case for internalizing behavior concerns, which indicated that mother-father agreement did not differ from parent-child agreement (one exception of T3 mother-father agreement, which was higher than father-child agreement, z=2.80, $p<.01$). Finally, teacher-child correlations for both externalizing and internalizing problems were fairly weak, although higher for child externalizing problems than internalizing problems. At T4 and T5 (although not T3), adult-child (including both parents and teacher) externalizing correlations were significantly higher than adult-child internalizing correlations, z-score=2.15-4.09, $p<.05$, indicating consistently greater agreement between informants for externalizing concerns over internalizing concerns in adolescence, but not pre-adolescence, where agreement between externalizing and internalizing concerns were similar between children and each adult informant. At each time point for both

child externalizing problems and internalizing problems, mother-father agreement was significantly higher than child-teacher agreement, z-score range=2.38-6.73, $p<.05$.

Associations between problem behaviors and distal outcomes. Overall, many distal outcomes yielded significant bivariate correlations with children's average externalizing problem behaviors, with fewer associations between distal outcomes and children's average internalizing problems; see Table 8.

Externalizing problems. For all informants, higher average externalizing problem behaviors were weakly to moderately correlated with lower relationship quality, lower educational grade achievement, lower educational degree obtained, lower individual income, lower family income, greater likelihood of experiencing arrest, more legal infractions, and a greater maximum number of drinks in 24 hours. Externalizing problem behaviors were not correlated with cohabitation age or marital age, but later age at first child was correlated with parent and teacher report of greater problems. More alcohol-use problems were correlated with higher mother-, teacher-, and child-report, but not father-report of externalizing problems.

Internalizing problems. Marital age, age at first child, maximum number of drinks, and alcohol-use problems were not significantly correlated with any informant report of internalizing problems. Mother and teacher reports of internalizing problems were weakly correlated with lower relationship quality, lower educational obtainment, and greater likelihood of experiencing arrest; father and child reports of internalizing behaviors were not correlated with these outcomes. Mother, teacher, and child reports (but not father reports) of greater internalizing problems were correlated with a greater number of legal infractions. Teacher and child reports (but not parent reports) of greater internalizing problems were associated with lower individual income, while teacher reports alone were associated with lower family income. Child report of

greater internalizing problems was the only informant report that was linked to an earlier age of cohabitation. Significant correlations tended to be fairly weak (highest r=-.27 for decreased family income).

Overall, average externalizing problems were consistently related to distal outcomes across multiple informants, with correlations ranging from weak to moderate. Average internalizing problems, on the other hand, tended to be weaker and less uniform across informants.

Associations between problem behaviors and childhood covariates. Similar to links between problem behaviors and distal outcomes, a large number of childhood covariates were significantly associated with externalizing behavior problems, with some links to internalizing behavior problems. See Table 9.

Externalizing problems. Parent and teacher reports of greater externalizing problems were associated with lower cognitive ability, r=-.19 to -.28, p<.001. In addition, both father and mother higher educational degree, higher educational grade, greater income, and better relationship quality were associated with fewer externalizing problem behaviors across informants. Greater levels of both mother and father depressive symptomology and greater neuroticism were associated with increased externalizing behavior problems across informants, suggesting that children who have parents with these concerns are observed to have more externalizing problems regardless of who is asked; that is, higher child concerns were not the sole product of biased parent reporting.

Internalizing problems. Only teacher reports of greater internalizing problems were correlated with lower cognitive ability, r=-.20, p<.001; parent and child reports of internalizing problems were not related to child cognitive ability. Overall, greater mother as well as father

depressive symptomology in addition to greater neuroticism were associated with more internalizing problems across most reporters, suggesting, as above was the case with externalizing problems, that children who have parents with these concerns were reported to have greater internalizing problems regardless of informant. Mother, father, and child (but not teacher) reports of greater internalizing problems were linked to both mother and father lower marital quality. Of note, average teacher report of internalizing problems was significantly correlated with every childhood covariate except parents' martial quality. Father, teacher, and child reports (but not mother) of greater internalizing problems were associated with lower mother and father income. While mother-reported internalizing problems were not significantly related to parent education, father-reports of greater internalizing concerns were weakly but significantly related to lower mother and father education.

Shapes of Problem Behaviors over Time

Externalizing problem behaviors. For both mothers and fathers, externalizing behaviors from ages 3-17 (T1-T5) fit a quadratic model. Across the entire sample, consistent with past literature (Bongers et al., 2003; Leve et al., 2005), externalizing behaviors decreased over time and then slowed slightly in their decrease (N=917-967, χ^2=4.31-6.15 (*ns*), RMSEA=.000-.005, CFI=1.000, TLI=1.000-1.004, SRMR=.017-.022); in both models, the variance components of the latent intercept, slope, and quadratic terms were significant, indicating individual differences in children's trajectories. Child-report models, which spanned from T3-T5 (ages 11-17), best fit a linear model that demonstrated increasing externalizing behaviors over ages 11-17 (N=994, χ^2=.006 (*ns*), RMSEA=.000, CFI=1.000, TLI=1.014, SRMR=.001), with significant variance around the intercept and slope; this increasing model differs from past literature. Because these models spanned different time frames (parents from

T1-T5, children from T3-T5), abbreviated parent-report models were also run from T3-T5 in order to capture the same developmental span as children. For both abbreviated parent models (T3-T5), trajectories fell in an opposite direction to child-report models, with data fitting a linear model that suggested externalizing behaviors decreased over the course of adolescence (N=857-742, χ^2=.000-.246 (*ns*), RMSEA=.000, CFI=1.000, TLI=1.009-1.014, SRMR=.000-.004), a direction that is in line with prior research. Both parent models yielded significant variance around the intercept, while father-report models also had significant variance in the slope term (mother-report models slope variance approached significance). Teacher models were run with their full data (T2-T5) as well as the abbreviated timespan in order to match the T3-T5 period; for both, intercept-only models fit the data best (N=918-953, χ^2=4.39-13.38 (*ns*), RMSEA=.010-.027, CFI=.958-.994, TLI=.969-.996, SRMR=.031-.051), with significant variance around the intercept.

Internalizing problem behaviors. Mother- and father-reported models that assessed children aged 3-17 (T1-T5) yielded quadratic change patterns (N=742-967, χ^2=7.04-32.35 (range: *ns* to $p<.001$), RMSEA=.030-.069, CFI=.923-.990, TLI=.871-.992, SRMR=.035-.055), with significant individual differences apparent on the intercept, slope, and quadratic terms. For both mothers and fathers, child internalizing problems increased over time and then decreased slightly. However, mother and father models that assessed ages 9-17 (T3-T5) demonstrated no mean-level change over time (i.e., intercept-only models fit the data best, with significant intercept variance). Child-report models indicated that reports of internalizing problem behaviors modestly decreased across adolescence (N=993, χ^2=.659 (*ns*), RMSEA=.000, CFI=1.000, TLI=1.009, SRMR=.008), with significant variance on the intercept as well as slope terms. For internalizing reports from teachers, both T2-T5 and T3-T5, intercept-only models

once again fit the data best (N=917-952, χ^2=4.49-17.95 (p=.02-.05), RMSEA=.036-.039, CFI=.753-.760, TLI=.815-.820, SRMR=.050-.055), with significant variance on the intercept.

Distal Outcome Predictions

Do problem behaviors predict later timing of normative transitions?

Cohabitation age. Inconsistent with past findings (Bardone et al., 1996), no externalizing or internalizing problem behavior trajectory predicted the age at which young adults began cohabiting with a romantic partner.

Marital age.

Externalizing behaviors. The slope of mother-reported child externalizing problems from T1-T5 significantly predicted an earlier marital age (R^2=.129); children who decreased less in externalizing behaviors over the course of development had an earlier marital age.

With covariates included. The slope of mother-reported child externalizing problems (T1-T5) continued to predict children's marital age above and beyond covariates; no covariate contributed significantly to the prediction of marital age.

Internalizing behaviors. No internalizing problem behavior trajectory predicted subsequent marital age.

Age at first child.

Externalizing behaviors. Consistent with past findings that externalizing problems are related to earlier pregnancy (Bardone et al., 1996; Woodward & Fergusson, 1999), both teacher models (T2-T5 and T3-T5) and father models (T1-T5 and T3-T5) of child externalizing problems predicted an earlier age at first child (R^2 ranged from .078 for T2-T5 teacher-report to .203 for T1-T5 father-report). In addition, T3-T5 mother-reported externalizing behaviors (but not T1-T5) also significantly predicted an earlier age at first child (R^2=.041).

With covariates included. Once covariates were factored in, only the intercept of T3-T5 father-reported child externalizing behaviors continued to significantly predict an earlier age of first child later on. No covariate entered was uniquely predictive.

Internalizing behaviors. T3-T5 father-reported and T3-T5 teacher-reported internalizing behavior problems predicted an earlier age at first child (R^2=.058-.063), partially supporting past findings (Bardone et al., 1996) and suggesting that reports obtained later in childhood rather than earlier may carry greater implications in predicting later age at first child.

With covariates included. Neither model remained a significant predictor of later age at first child after covariates were included.

Do problem behaviors predict later romantic relationship quality?

Externalizing behaviors. In line with past work with earlier ages (Woodward et al., 2002), child-reported externalizing problems as well as teacher-reported externalizing problems significantly predicted lower relationship quality, R^2=.084 for the child-reported problems to R^2= for .135 for T3-T5 teacher-reported problems. For child-reported externalizing problems, the slope value significantly predicted later problems, while only intercept values were significantly predictive for teacher-based models.

With covariates included. Only teacher-reported externalizing problem behaviors remained significant predictors of relationship quality; child sex was also uniquely predictive. In the child-reported model, (higher) child FSIQ became a significant predictor of improved later relationship quality, and the slope of child externalizing problem behaviors was no longer a significant predictor of relationship quality.

Internalizing behaviors. In line with past work (Bardone et al., 1996), a higher intercept value for teacher-reported internalizing problem behaviors (both T2-T5 and T3-T5 models) also

predicted later lower relationship quality (R^2=.168-.236). In addition, a higher intercept for T3-T5 mother-reported internalizing problem behaviors (but not T1-T5) predicted lower relationship quality, R^2=.032.

With covariates included. Only teacher-reported internalizing problem behaviors remained significant predictors of lower relationship quality, although no covariate contributed additionally to the model. In the mother-reported model of internalizing problems, child FSIQ emerged as a significant predictor, but the intercept of internalizing problem behaviors was no longer a significant predictor of later romantic relationship quality.

Do problem behaviors predict later educational attainment?

Grade attainment.

Externalizing behaviors. Consistent with past work using younger participants (McLeod & Kaiser, 2004; Miech et al., 1999), every model of externalizing behaviors tested (both parents, T1-T5 as well as T3-T5; teacher T2-T5 as well as T3-T5; and child) indicated that a higher intercept value predicted fewer completed grades. T2-T5 teacher-reported externalizing behavior accounted for the largest amount of the variance in predicting grade completion (R^2=.220), while child-report contributed the least (R^2=.061). For both mother-reported and father-reported models from T1-T5, the slope of problem behaviors also significantly predicted grade attainment, in that children who decreased less over time in externalizing concerns achieved lower educational attainment.

With covariates included. All models continued to predict grade attainment above and beyond the covariates entered. Child FSIQ was also a significant predictor in all models. In multiple models, maternal education, paternal education, and maternal income also contributed significantly.

Internalizing behaviors. Consistent with limited prior work on earlier ages (Bardone et al., 1996; McLeod & Kaiser, 2004), greater intercepts of teacher-reported internalizing behaviors (both T2-T5 as well as T3-T5 models) as well as T3-T5 mother-reported internalizing behaviors predicted lower grade attainment (R^2=.036 for the mother-reported model; R^2=.224-.259 for the teacher-reported models). Neither child-reported nor father-reported models of internalizing behaviors predicted later grade completion.

With covariates included. As with externalizing behaviors, each model continued to predict grade attainment above and beyond the covariates. In each, child FSIQ and maternal education were unique predictors; for the mother-reported model, maternal income was additionally predictive.

Degree attainment.

Externalizing behaviors. As occurred with grade attainment and consistent with prior work (King et al., 2006), each model of externalizing behaviors (both parents, T1-T5 as well as T3-T5; teacher T2-T5 as well as T3-T5; and child) indicated that a higher intercept value predicted lower degree completion. T2-T5 teacher-reported externalizing behavior accounted for the largest amount of the variance (R^2=.137), while child-report contributed the least (R^2=.028). As with number of years of education, for both mother-reported and father-reported models from T1-T5, the slope of problem behaviors also significantly predicted degree attainment, in that children who decreased less over time in externalizing concerns achieved lower degree attainment.

With covariates included. All models except for child-reported externalizing problems continued to predict degree attainment above and beyond the covariates; the child-report model was rendered nonsignificant, with child FSIQ, maternal education, and paternal

education accounting for a notable portion of the variance. As with grade attainment, child FSIQ was a significant predictor in all models. In multiple models, maternal education, paternal education, and maternal income also contributed significantly.

Internalizing behaviors. As was the case with grade completion (e.g., McLeod & Kaiser, 2004), the intercepts of teacher-reported internalizing behaviors (T2-T5 and T3-T5 models) as well as T3-T5 mother-reported internalizing behaviors predicted lower degree attainment (R^2=.033 for the mother-reported model; R^2=.174-.197 for the teacher-reported models).

With covariates included. Each model continued to predict degree attainment above and beyond the covariates entered. In each, child FSIQ, maternal education, paternal education, and maternal education were common significant predictors.

Do problem behaviors predict later income?

Individual income.

Externalizing behaviors. In line with limited past evidence with younger participants (Bardone et al., 1996; Rohde et al., 2007), there was some connection between earlier externalizing concerns and later income. Both teacher-reported models (T2-T5 as well as T3-T5) as well as T3-T5 father-reported externalizing behaviors indicated that a higher intercept value predicted lower later individual income. Teacher-report (T2-T5) explained the greatest amount of variance in the outcome (R^2=.075), while father-report (T3-T5) explained the least (R^2=.015).

With covariates included. Only teacher-reported models of externalizing behaviors continued to predict later individual income above and beyond covariates in each model. In each model where covariates were entered, child FSIQ also emerged as uniquely predictive, and child sex was additionally predictive in teacher-reported models.

Internalizing behaviors. While prior work has not found a connection between internalizing concerns and later income (e.g., Bardone et al., 1996), the intercepts of teacher-reported internalizing behaviors (both T2-T5 as well as T3-T5 models) as well as child-reported internalizing behaviors predicted lower later individual income (R^2=.026 for the child-reported model; R^2=.108-.142 for the teacher-reported models). Neither mother-reported nor father-reported models of internalizing behaviors predicted later individual income.

With covariates included. Once covariates were included, only teacher-reported models continued to predict later individual income. Child FSIQ was a significant additional predictor in each model.

Family income.

Externalizing behaviors. Patterns of findings for later family income were fairly similar to individual income. Teacher-reported models (both T2-T5 and T3-T5) as well as T3-T5 father report predicted later family income (R^2=.048 for father-reported models, R^2=.079-.091 for teacher models), as did a higher intercept on T1-T5 mother-reported externalizing behavior (R^2=.085).

With covariates included. Only T1-T5 mother-reported externalizing behaviors and T2-T5 teacher-reported externalizing behaviors continued to predict lower later family income. In each, child FSIQ was also a significant predictor, and in several models, maternal education were additionally predictive.

Internalizing behaviors. As with individual income, the intercepts of teacher-reported internalizing behaviors (both T2-T5 as well as T3-T5 models) predicted lower later family income (R^2=.168-.220). Mother-reported, father-reported, and child-reported models of internalizing behaviors did not predict later family income.

With covariates included. Both teacher-reported models continued to predict later family income above and beyond all covariates. In both, greater maternal education was additionally predictive of greater subsequent family income.

Do problem behaviors predict later legal infractions?

Experience of arrest.

Externalizing behaviors. In line with considerable evidence connecting externalizing behaviors to later difficulty with law enforcement (e.g., Grella et al., 2005; Vander Stoep et al., 2002), each model of externalizing behaviors (both parents, T1-T5 as well as T3-T5; teacher T2-T5 as well as T3-T5; and child) indicated that a higher intercept value predicted a greater likelihood of later arrest. Teacher-report (T3-T5) explained the greatest amount of variance in the outcome (R^2=.192), while father-report (T1-T5) explained the least (R^2=.055).

With covariates included. All models of externalizing behaviors predicting arrest held above and beyond covariates in each model. In many of the models, (male) child sex also emerged as an additional significant predictor.

Internalizing behaviors. While past work has typically not found a connection between internalizing concerns and later arrest, the intercepts of teacher-reported internalizing behaviors (both T2-T5 as well as T3-T5 models) predicted a greater likelihood of subsequent arrest (R^2=.071-.079). In addition, mother-reported problems that spanned T3-T5 (not T1-T5) predicted subsequent arrest (R^2=.028).

With covariates included. Once covariates were included, arrest was not predicted by any informant report of internalizing behavior; (male) child sex instead emerged as the driving predictor in each model.

Legal infractions.

Externalizing behaviors. As with experiencing arrest and in line with past work (Darney et al., 2013; Hämäläinen & Pulkkinen, 1996), every model of externalizing behaviors tested across informants indicated that a higher intercept value predicted a greater likelihood of future legal infractions. Mother-report (T3-T5) explained the greatest amount of variance in the outcome (R^2=.115), while child-report (T1-T5) explained the least (R^2=.065).

With covariates included. All externalizing behavior models predicting legal infractions held above and beyond covariates in each model. In parent-report and child-report models, (male) child sex also emerged as an additional significant predictor. In teacher models, (less) maternal education was an additional significant predictor of later legal infractions.

Internalizing behaviors. As was the case with experience of arrest, the intercepts of teacher-reported internalizing behaviors (both T2-T5 as well as T3-T5 models) predicted a greater likelihood of subsequent arrest (R^2=.045-.060). In addition, mother-reported problems that spanned T3-T5 (not T1-T5) predicted subsequent arrest (R^2=.016).

With covariates included. Once covariates were included, the T3-T5 maternal and teacher reports were no longer significant; however, T2-T5 teacher internalizing reports continued to predict later legal infractions. Less maternal education as well as male child sex were significant predictors of increased legal infractions in both teacher models, while male child sex was a significant predictor in mother-reported models.

Do problem behaviors predict later alcohol use?

Drinks consumed within 24 hours.

Externalizing behaviors. In line with past work identifying a connection between externalizing behaviors with subsequent increased alcohol use and alcohol use disorder (Jester et

al., 2008; Zucker et al., 2007), the intercept values of most models of externalizing problem behaviors predicted respondents' later number of maximum alcoholic drinks consumed within a 24-hour period, with the exception of mother-reported T3-T5 model (mother-reported full T1-T5 externalizing problem model was predictive). Variance explained in the number of drinks consumed ranged from R^2=.021 (both T1-T5 and T3-T5 father-reported problems) to R^2=.067 (child-reported problems).

With covariates included. Three of the six models that predicted later drinks consumed remained significant after accounting for covariates, which included the two teacher-reported externalizing models (T2-T5, T3-T5) as well as the child-reported externalizing model; parent models were no longer significant. In all models, (male) child sex predicted a greater number of drinks consumed. In both father models (which were rendered non-significant), (lower) father income also emerged as a unique predictor.

Internalizing behaviors. No internalizing problem behavior trajectory predicted later number of drinks.

Alcohol-use problems.

Externalizing behaviors. While prior work has identified strong connections between earlier externalizing problems and later alcohol use problems (e.g., Chassin et al., 2002; Jester et al., 2008; Zucker et al., 2007), just two trajectories of externalizing problems predicted later alcohol-related problems, which included mother-reported problems from early childhood through late adolescence (T1-T5, but not T3-T5) as well as child-reported problems. A greater intercept for both models predicted more alcohol-related problems (R^2=.039-.053).

With covariates included. Both models predicting alcohol-related problems held above and beyond covariates in each model. In both models, (male) child sex and (lower)

maternal income emerged as additional unique predictors. In the child-reported predictive model, (lower) child FSIQ also predicted more alcohol-related problems.

Internalizing behaviors. No internalizing problem behavior trajectory predicted later alcohol-related problems, which is consistent with one prior study examining the unique aspects of internalizing problems connected to later alcohol use problems (Foster et al., 2018).

Are problem behavior elevations at specific time points especially important?

In investigating whether certain time periods were especially important in predicting distal outcomes, each T-period was tested in turn as a direct predictor of the outcome over and above the latent growth parameters for models that had at least one significantly predictive growth parameter. For example, because the intercept of teacher-reported internalizing behaviors (T2-T5) significantly predicted later lower relationship quality, each time period was entered in turn to determine if it was significant above and beyond the shape of change (i.e., T2 internalizing problems were included, then T3, then T4, then T5). Theoretically, if that time period significantly predicted the distal outcome, the implication would be that children's problem behavior at that developmental stage was especially predictive of the outcome.

Overall, the patterns of results from these analyses were largely contradictory with each other and did not fit prior theory, suggesting that these conclusions would be difficult to interpret with confidence. For example, a higher intercept of T2-T5 teacher-reported internalizing problems predicted fewer completed grades later on. As each T period was entered, T2 emerged as significant in a positive direction, and T4 emerged as a significant predictor with a virtually identical coefficient but in the negative direction. This pattern of opposite but equally strong coefficients was common.

A likely explanation for these findings may be that, for a majority of models, the best-fitting trajectory was an excellent fit, leaving little remaining systematic variance for which to account. Full results of these analyses have been detailed in the supplemental analyses.

DISCUSSION

This study addressed two primary questions: (1) Do the effects of externalizing and internalizing problem behaviors sustain or decay over time in their relation to later adulthood outcomes? (2) Which informants in children's lives (mothers, fathers, teachers, and the children themselves) provide the strongest predictive ability?

Hypotheses 1 and 2

The first hypothesis proposed that greater levels of childhood externalizing problems would predict an earlier age at transitions (age at cohabitation, marriage, and first child), lower romantic relationship quality, more legal infractions, greater likelihood of experiencing arrest, lower educational attainment, and greater alcohol use problems. The second hypothesis proposed that greater levels of internalizing concerns would predict earlier cohabitation age, lower romantic relationship quality, and lower educational attainment. Overall, these hypotheses were partially supported; an inherent assumption of these hypotheses, based on prior literature, was that childhood externalizing problem behaviors carry greater weight for later outcomes than internalizing concerns. This idea was supported in the sense that externalizing concerns were stronger predictors (i.e., accounted for a greater amount of the variance) of later outcomes than internalizing concerns, although not consistently across reporters for the specific outcomes that were predicted.

For child externalizing concerns, models explained up to 22% of the variance in a later outcome. Even after factoring in all covariates, most distal outcomes stemming from child externalizing problems were significantly predicted by at least two informants. More externalizing problems predicted lower relationship quality, lower educational attainment (in both years and degree), greater likelihood of experiencing arrest, greater number of legal

infractions committed, and greater alcohol use concerns (in both the maximum number of drinks consumed within 24 hours as well as alcohol use problems). Future individual income was predicted by teacher-reported externalizing behaviors only, but this prediction was significant across both time spans examined (T2-T5 as well as T3-T5). One outcome was universally *not* predicted, which was age at cohabitation. Indeed, age at transitions were the most weakly predicted outcomes from childhood externalizing concerns, which may in part be due to the lower covariance coverage in the timing of transitions compared to other outcomes. Greater childhood externalizing concerns predicted an earlier marital age (predicted by mother-report only across T1-T5) and age at first child (predicted by father-report only across T3-T5), but both only appeared significant by that particular informant over that specific span of time.

Prior research that has linked earlier externalizing problems to transition timing (e.g., Bardone et al., 1996) did not extend past age 21, suggesting that externalizing concerns may be a less robust predictor of transition age when outcomes are extended further in the future. However, other findings in the past literature were replicated, such as the connection between higher externalizing problems and increased difficulty in romantic relationship quality (Woodward et al., 2002). Similarly, the present study extended prior findings (e.g., Bardone et al., 1996; King et al., 2006; McLeod & Kaiser, 2004; Miech et al., 1999) that externalizing problems significantly reduced later educational attainment in both number of grades completed as well as degree obtainment.

A strong body of literature has identified that earlier externalizing problems predicted increased substance use in late adolescence (age 15-17; Jester et al., 2008) as well as young adult (age 18-23) diagnoses of alcohol use disorder (Zucker et al., 2007). The present study indicated that these effects of earlier externalizing problem behaviors persisted past emerging adulthood

and continued into early adulthood (age 24-32), past the ages of peak drinking seen in emerging adulthood (Masten, Faden, Zucker, & Spear, 2009) and into one in which there is typically a normative decline of alcohol related problems (Chassin, Sher, Hussong, & Curran, 2013). These effects were evident on measures of the maximum number of drinks consumed within a 24-hour period as well as alcohol-related problems. In a similar fashion, several studies have suggested that the presence of greater externalizing concerns in childhood as well as adolescence predict a greater likelihood of a criminal record (e.g., Darney et al., 2013; Hämäläinen & Pulkkinen, 1996; Bardone et al., 1996; Grella et al., 2005; Vander Stoep et al., 2002). Here, legal difficulties were conceptualized through rates of arrest as well as the more broadband measure of number of legal infractions; the significant links found, above and beyond the inclusion of covariates that included (among others) parental education and income as well as child sex, contributes further evidence to the strong connections between the presence of earlier rule-breaking, defiant, and aggressive behaviors to a greater likelihood of law-breaking behavior in adulthood.

For internalizing problem behaviors, it was predicted that higher levels would be linked to decreased relationship quality, given prior evidence that girls with depression at age 15 were at highly increased risk for intimate partner victimization by age 20 (Keenan-Miller et al., 2007). In the current study, a more comprehensive assessment of romantic relationship quality was utilized, and results indicated a partial confirmation of this hypothesis. That is, greater teacher-reported child internalizing concerns (but not parent- or child-reported concerns) predicted decreased relationship quality.

Prior literature has highlighted that internalizing problems may also be linked to reduced educational achievement. This hypothesis was confirmed in the current study, as decreased educational attainment was found even after including a comprehensive set of covariates that

contained, among others, child cognitive ability and parental education level. While teacher-reported child internalizing problems spanning child age 6-17 predicted this decreased educational attainment, when reports were examined for the pre-adolescent and adolescent period from ages 9-17, elevations on both mother-reported and teacher-reported child internalizing concerns predicted both decreased number of grades obtained as well as decreased degree attainment.

Unexpectedly, the presence of internalizing problem behaviors also predicted several other future outcomes, including a greater number of later legal infractions as well as lower income (both individual and family). Of note, the majority of these outcomes were predicted solely by teacher-reported internalizing problems, with occasional matching predictions by mother- or child-reported internalizing concerns. This finding regarding the strength of teacher report is in line with recent work that found that teachers' reports of personality were strongly predictive of a variety of later psychosocial aspects of children's functioning, often outperforming parents' reports (Clark, Durbin, Hicks, Iacono, & McGue, 2017).

Few studies have focused on connections between youth internalizing concerns and these later outcomes. Some past literature (Bardone et al., 1996) found that girls self-reporting depression at age 15 were more likely to be living with a partner and somewhat more likely to have given birth by age 21 than girls who were not depressed at age 15; however, as was also the case with externalizing problem behaviors, there was no evidence of this in the current sample, suggesting that this connection may not exist further in adulthood. In addition, there has been limited examination of a connection with income; the same study (Bardone et al., 1996) found that girls self-reporting depression at age 15 were not more likely to receive federal financial assistance at age 21 than girls without depression at age 15. However, in the current study,

greater internalizing concerns as reported by teachers predicted lower income on an individual level as well as a family level.

Typically, the presence of internalizing behaviors has not been related to later law-breaking behaviors, which is commonly defined as the presence or absence of a criminal record (e.g., Bardone et al., 1996), but this binary measure may capture more severe behaviors than dimensional assessments; here, when examining a broader picture of legal infractions and considered from childhood, greater teacher-reported internalizing problem behaviors significantly predicted a greater number of legal infractions, above and beyond covariates, while prior arrest (likely a more severe measure) was not significant after covariates were included. Finally, alcohol use concerns are typically considered to be related to earlier externalizing but not internalizing concerns; for example, Foster et al. (2018) found that the shared variance of internalizing and externalizing problems (e.g., negative emotionality, distress) predicted increase alcohol use from ages 9-11 through ages 21-23, while the unique aspects of internalizing problems predicted decreased alcohol use problems. The current study also served to extend these past findings past emerging adulthood, indicating that early internalizing problems are not, on average, related to later drinking problems.

Overall, as hypothesized, externalizing problem behaviors were more strongly related to later outcomes than internalizing problem behaviors. Several outcome measures relied on metrics that indicated greater "acting out" behaviors (e.g., legal infractions), while fewer metrics targeted processes that might be more explicitly affected by the continuity of internalizing concerns (e.g., job satisfaction). Alternatively, it is possible that the presence of internalizing problem behaviors across childhood result in less experiential canalization than externalizing problem behaviors, or have a decreased likelihood of "getting in the way" of positive outcomes.

In addition, the shared overlap between internalizing and externalizing problems (e.g., negative emotionality, general distress) may explain the prediction of later shared outcomes (e.g., experience of arrest).

Hypothesis 3

Given that most prior work has either not included multiple informants or has simply averaged reports, Hypothesis 3 sought to examine how conclusions regarding whether certain domains of problem behaviors predicted certain future behaviors would differ based on which informant provided information. The hypothesis that conclusions would differ between informants was largely supported, and the amount of the variance each informant explained differed widely.

When examining the mean proportion of the variance explained in outcomes, mothers' reports of child externalizing behavior explained about 10% of the variance in significant outcomes, while mothers' reports of child internalizing behavior explained about 3% across several significant outcomes. Fathers' reports explained slightly less of the variance in outcomes when considering externalizing problems (mean of 8%); for internalizing problems, fathers' reports significantly predicted age at first child through T1-T5 reports (30%) and T3-T5 (6%, nonsignificant after covariates), for an average of 18% that may be presumed to be a less stable or generalizable estimate. Children's reports of their externalizing problems yielded the lowest amount of later variance explained in significant outcomes, ranging from a mean of 2% from internalizing problems to 6% from externalizing problems. Teachers, however, consistently explained the greatest amount of the variance on average as well as predicted the greatest number of outcomes, ranging from an average of 11% from their reports of child externalizing problems to 15% of the outcome variance from their reports of child internalizing problems.

In addition, of the 11 outcomes predicted by at least one informant's report of child problem behaviors (collapsed across time point spans as well as internalizing vs. externalizing concerns, maintained above and beyond the inclusion of control variables), teachers' reports explained the greatest amount of the variance in six of these outcomes, including children's later experience of arrest (19%), highest educational grade (26%), highest educational degree (20%), individual income (14%), family income (22%), and romantic relationship quality (24%); interestingly, aside from experience of arrest, these highest effect sizes were found for teachers' reports of child internalizing behaviors rather than externalizing behaviors. Mothers' reports explained the greatest variance in 3 outcomes, including children's later legal infractions (12%), alcohol-use problems (5%), and marital age (13%), each from maternal reports of child externalizing problems. Fathers' reports as well as children's reports explained the greatest amount of the variance in one outcome each, both from reports of child externalizing problems: Fathers explained the greatest amount of the variance in children's later age at first child (13%), while children's self-reports explained their later maximum number of drinks (7%). Furthermore, the percentage of the variance explained in a single outcome could vary widely between informants; for example, different informants who were both significant predictors of an outcome could explain from 3% up to 25% of that subsequent outcome (here, mother and teacher T3-T5 reports of child internalizing problems predicting later child highest education grade).

Very few outcomes were predicted by all four informants (mothers, fathers, teachers, and children). Of note, the outcomes that were "unanimous" in their connections to later outcomes were all predicted by informants' reports of externalizing behavior, with higher levels of externalizing behaviors predicting a later greater likelihood of arrest and legal infractions, as well as lower educational attainment. In addition, some outcomes were "unanimous" in that they

were *not* systematically related to earlier functioning; children's internalizing behavior was unrelated to later cohabitation age, marital age, maximum number of alcoholic drinks, and alcohol-use problems. Similarly, for externalizing problems, none of the informants' report of children's externalizing concerns predicted later cohabitation age.

It could be argued that we can be most confident that a problem behavior predicts a later outcome if that finding is replicated across multiple informants within the same sample, and with similar effect sizes. It seems quite likely that replication across informants can help us to have increased assurance in the association between an earlier concern and a later aspect of functioning. However, simply because an outcome is predicted by one informant but not others does not imply that finding is spurious, even if associations are not always as hypothesized; for example, we typically conceptualize teachers to be more sensitive to externalizing problem behaviors given that rule-breaking and disruptive behaviors stand out in a classroom, and being typically less sensitive to internalizing problem behaviors, given that worries or a depressed mood may not require a teacher's attention (e.g., Berg-Nielsen et al., 2012; De Los Reyes & Prinstein, 2004; Rescorla et al., 2013; Zahner & Daskalakis, 1998). Instead, teachers' strength may lie in a *decreased rate* of "false positive" reporting of meaningful internalizing concerns, which might actually explain why teachers' reports of internalizing problems far outperformed parents' and children's reports of internalizing concerns in the strength with which they tell us information about those children's functioning more than a decade later. Descriptively, compared to parents and children themselves, teachers typically reported a slightly lower average rating of children's internalizing problems at each time point, which is in line with past work (Rescorla et al., 2013, 2014). Notably, there was no outcome predicted by earlier internalizing concerns that parents or children predicted more strongly than teachers.

Tentatively, I hypothesized that teachers' reports of child externalizing problem behaviors would be a stronger predictor than parents' reports, while teachers' reports of child internalizing problems would fail to tell us information about functioning later on. I also hypothesized that adolescents' reports of internalizing concerns would carry greater weight for predicting later functioning than parents' reports, given children's privileged access to their internal state. However, these hypotheses were largely not supported. Overall, the reports provided by parents and teachers of children's externalizing behavior predicted outcomes with similar effect sizes, with teachers' reports often (although not always) predicting slightly more of the variance in aspects of later functioning than parents' reports. This effect was even more pronounced for reports of child internalizing problems. While child-reported problem behaviors predicted several aspects of their own later functioning, other reporters tended to contribute information that carried greater predictive ability.

Hypotheses 4 and 5

Hypothesis 4 posited that parents' depressive symptoms and/or neuroticism may impact their assessment of children's problem behaviors, which would in turn affect their prediction of distal outcomes. Overall, this hypothesis was not supported.

When examining bivariate correlations between average parent reports of problem behaviors, maternal depressive symptoms as well as neuroticism were weakly to moderately related to mother-reported child problems, and paternal depressive symptoms and neuroticism were similarly modestly related to father-reported child problem behaviors. Of note, teacher-reported child externalizing concerns as well as child self-reports of problem behaviors demonstrated quite weak connections to maternal and paternal depressive symptoms and neuroticism, while examination of cross-parent correlations yielded correlations between these

ranges. In *r*-to-z score comparisons, the correlations between maternal depressive symptoms and neuroticism with mother-reported child problem behaviors (both internalizing and externalizing) were significantly higher than maternal depressive symptoms and neuroticism with any other informants' reports of children's problem behaviors, including the other parent; the same was true for fathers. These cross-informant correlations as well as the statistical comparisons suggest that children of parents with greater depressive symptoms and neuroticism have greater problem behaviors, regardless of informant source; however, parents with greater depressive symptoms or neuroticism likely inflate their reports of their children's problem behaviors relative to the reports of children, teachers, and the other parent.

However, when parents' reports were used to predict later distal outcomes, controlling for parents' depressive symptoms and neuroticism did not affect most models' predictions of later functioning. When all covariates were entered together, neither parent depressive symptomology nor parent neuroticism emerged as uniquely significant predictors in any model of distal outcomes. Notably, parents' alcohol use disorder symptoms and antisocial personality disorder symptoms, which were also tested in parent models to determine if parent externalizing-spectrum behaviors were important to explain outcomes, were similarly and largely non-significant predictors of distal outcome measures. This surprising outcome may be due in part to the high-risk nature of the current sample.

Other person-specific factors, however, emerged repeatedly as important in predicting distal outcomes. Hypothesis 5 suggested that child sex, child cognitive ability, parental education level, and parental income would emerge as notable and significant predictors of later outcomes, and in some cases carry greater weight towards the distal outcome than problem behaviors. This hypothesis was largely confirmed.

Consistent with past work, both child cognitive ability and child sex frequently emerged as significant contributors, with child cognitive ability significant in 36 models (41% of models) and child sex in 29 models (39%). Typically, child sex emerged as significant across reporters in predicting experience of arrest, legal infractions, maximum number of alcoholic drinks, and alcohol-use problems, which was hypothesized given the higher male base rates of concern (e.g., Heimer et al., 2009; Nolen-Hoeksema, 2004; Steffensmeier et al., 2005). As predicted, children's cognitive abilities were significantly predictive of greater later educational attainment and income in most models (e.g., Dubow et al., 2009). In addition, both maternal and paternal education explained a significant portion of the variance in children's later educational attainment; in line with this, both maternal and paternal educational attainment yielded significantly higher zero-order correlations to child educational attainment than maternal and paternal average ratings of problem behaviors. Maternal education also contributed significantly to children's later income in some models.

Multiple covariates often emerged as significant; for example, in predicting later number of grades completed from child-reported externalizing problems, in addition to externalizing problems themselves, significant covariates also included maternal education, paternal education, and child cognitive ability; the percentage of the variance explained increased from 6.1% with externalizing behaviors alone to 45.8% with externalizing behaviors and covariates, demonstrating the critically important role of these additional family- and child-based factors. However, from other reporters, the percentage of the outcome variance explained by different contributing aspects changed dramatically. When looking at later number of grades completed from *teacher*-reported externalizing problems, significant covariates again included maternal education, paternal education and child FSIQ; however, here, externalizing problem behaviors

explained 20.9% of the variance in the outcome, while 46.2% of the outcome was explained once covariates were factored in. That is, the role of externalizing problems in predicting later outcomes rose from a small percentage to nearly half of the explained variance, even as the total explained variance remained virtually the same.

Thus, consistently, previously-established or previously-theorized covariates pertaining to socioeconomic status (i.e., parent education and income) as well as child factors (cognitive ability and sex) were important in understanding later aspects of children's functioning, while other aspects such as parent internalizing (depressive symptoms, neuroticism) and externalizing (alcohol use disorder symptoms, antisocial personality disorder symptoms) concerns appeared less important. However, even after all of these factors were accounted for, the majority of predictions held, indicating that child internalizing and externalizing problem behaviors have sustaining and critical effects on these children's later adulthood outcomes.

Developmental Timing

Because children could report on their own problem behaviors from pre-adolescence through late adolescence (T3-T5), both parent and teacher reports were examined with this restricted age range as well as their full available models. This method allowed for the examination of similarities and differences between models that targeted the adolescent time span compared to models that assessed concerns from earlier in childhood.

Overall, comparison of these timespan differences indicated that examining the full span of development was optimal for understanding later outcomes stemming from increased externalizing behaviors. However, for internalizing behaviors, focusing on the span of adolescence yielded predictions with greater effect sizes as well as a greater number of significant predictions. That is, for internalizing problem behaviors spanning from early

childhood through late adolescence, mother- and father-reported internalizing concerns did not predict a single later outcome (while teacher-reported internalizing concerns predicted seven out of the 12 tested outcomes, or six after covariates were included, with effect sizes ranging from R^2=8-22%, mean R^2=14%). On the other hand, mother-reported internalizing concerns that spanned only from pre-adolescence through late adolescence *did* carry implications for later functioning; mother-reported internalizing problems spanning this limited range predicted five out of 12 outcomes with effect sizes ranging from R^2=2-4%, mean R^2=3% (two remained significant after covariates were included; for comparison, teacher in this limited range predicted eight outcomes, R^2=5-26%, mean R^2=15%; five remained significant after covariates were included). As this example indicates, the difference in timespan between both externalizing and internalizing problem behaviors was *not* apparent for teacher-reported problem behaviors, whose reports of children's problem behaviors yielded almost identical results when examined from middle childhood through late adolescence as when examined from pre-adolescence through late adolescence.

A secondary question of timing centered on whether a particular time point was especially important in predicting a later outcome, above and beyond the average change over time of problem behaviors (e.g., do problem behaviors in late adolescence carry greater implications for subsequent educational attainment, given normative education-specific choices made in later high school?). However, results indicated marked difficulty in pinpointing particular developmental periods as important. No period meaningfully stood out as especially notable, but rather the overall "gestalt" of problem behaviors appeared key. This was also indicated by the strong tendency of intercept factors to emerge as significant rather than slope factors, indicating that the overall level of a particular problem behavior type was of greater

importance than its change over time; this may in part be due to the greater reliability and variance in intercept estimates over slope estimates. However, there were some exceptions to this, particularly for educational attainment. The slope component for both mother- and father-reported externalizing problems across development significantly predicted later educational attainment (both number of years of education achieved as well as degree obtained), including after all covariates were included. This finding suggested that children who began with more externalizing problem behaviors *as well as* who decreased less in those externalizing problems over time achieved less academically— that is, the children who began with greater externalizing problems and maintained this higher level had lower attainment.

Implications for Researchers and Clinicians

A seemingly unending question for researchers centers on how we can best capture the constructs in which we are most interested when there are limitations on our time and resources. In assessing children's problem behaviors, we can never have a fully clear picture of what is or is not "true" about a child's functioning— we are limited by the lens we use, be it naturalistic observation, laboratory tasks, psychophysiology data, reports by informants, or the multitude of other creative methods. Questionnaire-based reporting tends to have the advantage of requiring less participant time to complete, fewer associated research costs, straightforward resulting data to analyze, and the advantage of capturing low base-rate behaviors from the perspective of a person more familiar with a target child than we as researchers can be. Given these reasons, how do we decide from whom we gather this information?

One way of conceptualizing the importance of reports is what the information provided by an informant can convey, on average, about a child's later functioning. In the current study, mothers and fathers had little systematic advantage over one another in their assessments of

youth problem behaviors towards distal outcomes; however, mothers' reports were slightly less likely than fathers' reports to be rendered nonsignificant after covariates were included. Children themselves were relatively poor predictors of later outcomes; thus, their reports may tell us more about their current, in-the-moment perceptions and internal processes without those perceptions weighted towards later adulthood functioning. Finally, teachers were the break-away stars of informants. Researchers should keep in mind, however, that this does not anoint teachers as the "most accurate" reporters, and indeed they may systematically under-report internalizing concerns relative to parents and children (Berg-Nielsen et al., 2012; De Los Reyes & Prinstein, 2004; Rescorla et al., 2013; Zahner & Daskalakis, 1998). However, when externalizing and internalizing concerns are detected by teachers— that is, they are appearing in a context where many demands are placed on children outside of the home environment— the reporting of these behaviors becomes especially important to predicting future outcomes for these kids a decade or more later. The setting of school and the challenges faced within a classroom may be especially salient in understanding children's ability to be successful later on, and problems here might be understood as early precursors to later difficulties. Despite the time and resources it requires for researchers to obtain teacher reports, the current results indicate that when longitudinal outcomes are of interest, the effort may be well worth it.

For clinicians, best practice dictates that receiving information from multiple perspectives is key to understanding children's functioning, but information on how to effectively understand non-converging information is lacking. There have been some recommendations offered on how clinicians can prioritize based on the age of the child, the problem type, and the setting in which reports are collected (e.g., Smith, 2007), but these perspectives often assert that one informant is more accurate or reliable than another informant. Instead, the current study supports the

perspective that there are likely context-specific variations in behavior that are captured by different reporters, and these reporters may or may not be assessing concerns in such a way that informs efforts to understand functioning a decade later. For example, children themselves may be accurately reporting that they frequently feel worried about social judgments from others. However, decreased knowledge about what is normative as well as the depth of that privileged access may actually decrease later predictive ability— that is, if a *teacher* has noticed that a child is frequently worried about social judgments from others, that elevation may signal a greater impact in that child's life than the child his/herself endorsing that worry. While clinicians are appropriately more concerned with proximal rather than distal outcomes, a working knowledge of distal outcomes may be useful.

Thus, for a clinician, it might be helpful to not only think through the questions of, "What does it mean if *this behavior* is observed in *that context*, but not this other context? Why might that be happening?" but also consider the questions of, "What does it mean that this child's teacher noticed this behavior, but not the child's mother? How does that behavior display itself?" or "What does it mean that this child rates herself very highly on measures of internalizing problems, but both parents and teachers do not report these problems? How is this child coping with these concerns in such a way that they are not on the radar of her parents and teacher?" Prior studies of child psychotherapy have found that greater discrepancies between perspectives of children's problem behaviors predicts worse therapy treatment outcomes and process (e.g., more canceled or missed sessions, earlier drop-out; Brookman-Frazee, Haine, Gabayan, & Garland, 2008; Ferdinand, van der Ende, & Verhulst, 2006; Jensen-Doss & Weisz, 2008), suggesting that an important responsibility of clinicians is to examine these questions and

discuss discrepancies with a curious and empathic stance in order to facilitate shared understanding of a particular child.

Strengths and Limitations

Strengths. There were several notable advantages to the current study. First, the high-risk and unique characteristics of the study population provided greater variability in subsequent aspects of later functioning than would be expected from a higher-functioning community sample. This sample was then followed through multiple decades of assessment, providing remarkably detailed and thorough information into participants' lives over time from multiple perspectives. To my knowledge, no other study has examined such a variety of outcomes past emerging adulthood from early childhood problem behaviors, and certainly none has examined these outcomes from the lens of multiple informants.

In addition, this study utilized a thorough set of covariates in an effort to determine if later significant predictions were simply the result of important other factors in participants' lives. These covariates— including parental education level, parental income, parental depressive symptoms, parental neuroticism, parental antisocial personality disorder symptoms, parental alcohol use disorder symptoms, parental relationship quality, child sex, and child cognitive ability— represent a far more comprehensive set of explanatory variables than is conventional. In addition, the data utilized for this study were collected across children's lifespan, rather than relying on retrospective reports, enhancing confidence in data quality.

The findings carry implications for understanding informant reports of children's problem behaviors for researchers as well as clinicians. In addition, methodology used in this study allowed for the comparison of adolescence-focused models to models from early childhood in order to determine the relative strength of the developmental periods covered in predicting

later outcomes. Importantly, the amount of variance explained by the majority of these models tended to be relatively large, which speaks to the importance of the findings here: That childhood problem behaviors carry notable, sustained associations with later adulthood functioning.

Limitations. However, despite the strengths of the sample, methodology, and importance of its conclusions, there are also several limitations that warrant attention.

Two of these issues pertain to the assessment of problem behaviors. First, by design, informants responded to different questions. That is, the wording of particular items that comprise each scale differs in order to capitalize on context-specific observations (e.g., the teacher-report form does not ask about children's difficulty sleeping). It is unclear how or if the different questions used to tap into the same underlying construct changed the meaning of that internalizing or externalizing latent construct. While this question has received slight attention (e.g., Konold, Walthall, & Pianta, 2004), it has gone largely unexplored. Second, it is possible there was non-invariance in problem behaviors across time. Typically, measurements gathered across multiple occasions should first be tested for invariance in order to be sure that any growth or change that is observed is due to actual change, rather than differences in how respondents are reacting to questions. However, the nature of developmental psychopathology actually predicts that there is some expected developmental change over time in their manifestations; for example, anxiety may manifest more in younger children as somatic concerns, whereas older children may express more cognitively-based worries. Currently, there are no effective recommendations for researchers on how best to integrate these conflicting ideas and little research on how the Achenbach and Rescorla (2001) measures fare, although work done on the dysregulation profile of the CBCL found it to have partial invariance across time (Deutz, Geeraerts, van Baar, Deković, & Prinzie, 2016).

Another limitation was that parental neuroticism, depressive symptoms, antisocial behaviors, and alcohol use disorder symptoms were averaged across youth development. It was beyond the scope of the current study to model these as time-varying covariates. However, it is possible that the timing of these parent factors is important. For example, one avenue through which parental depressive symptoms may affect youths' development is through parenting behaviors, as meta-analytic examinations have demonstrated connections between maternal as well as paternal depression and parenting behaviors, particularly negative behaviors (Lovejoy, Graczyk, O'Hare, & Neuman, 2000; Wilson & Durbin, 2010). This impact on parenting behaviors may have a different effect on children in, for example, early childhood relative to late adolescence.

Next, the present study did not account for the frequent co-occurrence of externalizing and internalizing concerns (e.g., Gilliom & Shaw, 2004; Keiley et al., 2003; Youngstrom, Findling, & Calabrese, 2003). It is fully possible that some children would have elevated externalizing problems but not internalizing problems, other children would have elevated internalizing problems but not externalizing problems, still others may have elevations on both, and a final group might have no elevations, and these configurations may carry unique connections to later adulthood outcomes that were not captured here. Instead, the analyses conducted here operationalized children's problems as a spectrum of externalizing concerns and a separate spectrum of internalizing concerns, but not both, despite only moderate correlations between them (averaged internalizing-externalizing correlations ranged from r=.41 for children's self-report to r=.67 for father-reports). It is feasible that examining the dimension of overall problems may have simply yielded an average of the current findings (e.g., the conclusions

identified by externalizing problem behaviors but made weaker through the inclusion of internalizing problems), but without specific examination we cannot be sure.

While a major strength of this study was the inclusion of multiple perspectives, we have little concrete understanding of *why* the reports of certain informants were consistently more strongly related to later outcomes than other informants. Child self-report in particular was unexpectedly weak in its connection to later functioning. Prior research has found that children with higher levels of externalizing problems may underestimate their own concerns, while children with higher internalizing concerns may overestimate them (e.g., Hoza, Pelham, Dobbs, Owens, & Pillow, 2002; Hoza et al., 2004; De Los Reyes & Prinstein, 2004; Owens et al., 2007); this potential distortion at higher levels of each type of problem behavior is a possibility for why children's reports did not typically predict their own later outcomes. However, future studies are warranted to better understand this phenomenon.

As with any non-experimental design, we must also use caution in separating the concept that these earlier problem behaviors "cause" these later outcomes, and rather focus on the idea that these earlier problem behaviors can *tell us something* about later outcomes through conceptualizing earlier issues as risk factors or indicators. A third (combined with a fourth, fifth, and likely many more) variable may well be relevant in our understanding of connections between earlier and later functioning, and why the associations between childhood problems and adulthood outcomes are sustained. While it is certainly possible that childhood problem behaviors make it more difficult or "get in the way" of youth optimally functioning later on, the present analyses do not allow for causal inferences.

Summary

Overall, the present study furthered understanding of long-term outcomes in two important ways. The first question explored whether the effects of children's internalizing and externalizing problem behaviors sustain or decay in early adulthood. Findings here suggest that children's internalizing problem behaviors, overall, carry less weight towards long-term functioning than the presence of externalizing problem behaviors, while externalizing problem behaviors carry implications even above relevant familial and child-based covariates for later marital age, age at first child, romantic relationship quality, educational attainment, income, and legal infractions. The second question assessed the critical question of context and informant sensitivity in providing information that conveys these longer-term associations, and this body of work suggests that parents possess reasonably good predictive validity, especially mothers, but that teachers in particular possess an often-underappreciated wealth of valuable insight. Taken together, results support the critical importance of studying developmental psychopathology, not only for its concurrent prominence in children's functioning, but also for its implications in children's future.

APPENDICES

APPENDIX A

Tables

Table 1: *Means (and Standard Deviations) for CBCL, TRF, and YSR Problem Behavior Scales*

	Time Point				
	T1	T2	T3	T4	T5
Number of Cases					
Mother	513	575	577	643	594
Father	491	525	485	528	463
Teacher	-	410	597	732	701
Child	-	-	428	878	892
Externalizing M (SD)					
Mother	11.22 (6.94)	9.72 (7.23)	8.83 (7.18)	8.43 (7.63)	7.44 (8.18)
Father	10.95 (6.94)	8.72 (6.67)	7.68 (6.65)	7.35 (6.86)	6.65 (7.00)
Teacher	-	6.48 (9.67)	6.09 (9.28)	5.47 (9.07)	5.12 (9.03)
Child	-	-	8.74 (6.29)	10.04 (7.50)	11.03 (7.55)
Internalizing M (SD)					
Mother	4.50 (4.35)	5.65 (5.21)	5.33 (5.47)	5.85 (5.43)	5.44 (5.45)
Father	4.43 (4.22)	5.10 (4.91)	4.32 (4.54)	4.88 (4.87)	4.48 (5.02)
Teacher	-	5.07 (5.72)	5.36 (6.75)	4.51 (6.01)	4.22 (6.28)
Child	-	-	9.06 (7.77)	8.67 (7.63)	8.55 (7.49)

Table 2: *Within-Informant Correlations between CBCL, TRF, and YSR Scales for Externalizing Problems*

	Time Point			
	T2	T3	T4	T5
Mother				
T1	.62***	.53***	.48***	.41***
T2	-	.70***	.57***	.50***
T3		-	.65***	.52***
T4			-	.64***
Father				
T1	.63***	.57***	.51***	.43***
T2	-	.73***	.63***	.48***
T3		-	.68***	.52***
T4			-	.66***
Teacher				
T2	-	.45***	.43***	.27***
T3		-	.39***	.36***
T4			-	.34***
Child				
T3	-	-	.51***	.32***
T4			-	.48***

Note. ***p<.001. **p<.01. *p<.05.

Table 3: *Within-Informant Correlations between CBCL, TRF, and YSR Scales for Internalizing Problems*

	Time Point			
	T2	T3	T4	T5
Mother				
T1	.41***	.42***	.30***	.22***
T2	-	.60***	.54***	.42***
T3		-	.65***	.46***
T4			-	.57***
Father				
T1	.47***	.41***	.26***	.30***
T2	-	.63***	.52***	.48***
T3		-	.59***	.40***
T4			-	.56***
Teacher				
T2	-	.13*	.17**	.08
T3		-	.15**	.25**
T4			-	.21***
Child				
T3		-	.44***	.30***
T4			-	.37***

Note. ***p<.001. **p<.01. *p<.05.

Table 4: *Correlations of Cross-Parent Reports of Externalizing and Internalizing Problem Behaviors*

	Mother-Report of Externalizing Problems					Father-Report of Internalizing Problems				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
Father-Report of Externalizing										
T1	**.43*****	.37***	.30***	.32***	.34***	**.67*****	.41***	.42***	.38***	.35***
T2	.31***	**.55*****	.43***	.37***	.31***	.41***	**.64*****	.49***	.48***	.39***
T3	.29***	.44***	**.58*****	.46***	.39***	.39***	.51***	**.61*****	.48***	.34***
T4	.31***	.40***	.48***	**.57*****	.45***	.29***	.39***	.39***	**.61*****	.42***
T5	.24***	.36***	.36***	.46***	**.59*****	.32***	.36***	.33***	.42***	**.62*****
Mother-Report of Internalizing										
T1	**.59*****	.33***	.31***	.28***	.20***	**.26*****	.23***	.10†	.02	.06
T2	.39***	**.59*****	.42***	.43***	.29***	.15**	**.36*****	.24***	.19**	.07
T3	.35***	.41***	**.61*****	.47***	.31***	.12*	.26***	**.31*****	.21***	.09
T4	.29***	.33***	.42***	**.60*****	.37***	.12*	.15**	.23***	**.28*****	.16**
T5	.20***	.29***	.31***	.37***	**.58*****	-.01	.12*	.17**	.15**	**.24*****

Note. ****p*<.001. ***p*<.01. **p*<.05. †*p*<.10. Values in bold signify externalizing-internalizing correlations at the same point in time.

Table 5: *Correlations of Parent-Reported Externalizing Problems with Informant Reports of Problem Behaviors*

	Mother-Report of Externalizing Problems					Father-Report of Externalizing Problems				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
Teacher-Report of Externalizing										
T2	.27***	**.41***	.31***	.28***	.41***	.15*	**.26***	.18**	.21**	.32***
T3	.26***	.32***	**.45***	.30***	.30***	.18**	.25***	**.38***	.37***	.25***
T4	.20***	.30***	.25***	**.36***	.39***	.19***	.25***	.29***	**.35***	.25***
T5	.26***	.26***	.26***	.29***	**.26***	.16**	.23***	.22***	.32***	**.39***
Child-Report of Externalizing										
T3	.15*	.21***	**.28***	.27***	.13*	.19**	.18**	**.22***	.25***	.20**
T4	.24***	.17***	.22***	**.40***	.27***	.16**	.16**	.22***	**.35***	.30***
T5	.23***	.20***	.20***	.26***	**.36***	.25***	.13**	.16***	.27***	**.37***
Teacher-Report of Internalizing										
T2	.06	**.10**	-.02	.01	.16*	-.01	**.08**	.01	-.09	.18*
T3	.06	-.01	**.12***	.11*	.09†	.05	.07	**.15****	.08	-.05
T4	.01	.03	-.01	**.15**	.15**	.04	.01	.07	**.15****	.05
T5	.23***	.16**	.15**	.23***	**.17***	.06	.11*	.13*	.18***	**.23***
Child-Report of Internalizing										
T3	-.05	.03	**.09**	.19**	.01	-.07	.01	**.04**	.14*	.01
T4	.13**	.07	.10*	**.22***	.16***	.08	.09†	.08	**.19***	.15**
T5	.07	.02	.03	.02	**.09***	.09†	‘.05	.00	.00	**.09**†

Note. ****p*<.001. ***p*<.01. **p*<.05. †*p*<.10. Values in bold signify correlations at the same point in time.

Table 6: *Correlations of Parent-Reported Internalizing Problems with Informant Reports of Problem Behaviors*

	Mother-Report of Internalizing Problems					Father-Report of Internalizing Problems				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
Teacher-Report of Internalizing										
T2	-.02	**.13***	.11†	.10	.13*	.07	**.14****	.09	-.01	.15*
T3	.02	.05	**.22*****	.20***	.17**	.04	.05	**.14****	.11*	.03
T4	.02	.02	.08†	**.15****	.20***	-.01	.08	.11*	**.21*****	.11*
T5	.07	.01	.12*	.11*	**.11***	.00	.00	.13*	.17**	**.23*****
			.							
Child-Report of Internalizing										
T3	.04	.04	**.19*****	.21***	.13†	.01	.00	**.13***	.20**	.12
T4	.10*	.11*	.16***	**.25*****	.15**	.04	.14**	.11*	**.23*****	.27***
T5	.05	.08	.10*	.12**	**.23*****	.07	.00	.07	.03	**.18*****
Teacher-Report of Externalizing										
T2	.09	.09	.17**	.12*	.21**	.04	.09	-.01	.10	.22**
T3	.01	.00	.17***	.09*	.11*	.01	.07	.15**	.18**	.06
T4	.04	.08	.06	.07	.17***	.05	.12*	.13**	.15**	.10†
T5	.06	.01	.05	.04	.01	.03	.06	.04	.16**	.18***
Child-Report of Externalizing										
T3	.11†	.08	.17**	.11†	.02	.10	.10	.13*	.19**	.17*
T4	.16***	.08†	.09*	.16***	.10*	.27***	.05	.09	.22***	.20***
T5	.16**	.17***	.10*	.11**	.23***	.17***	.07	.08	.14**	.18***

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. Values in bold signify correlations at the same point in time.

Table 7: *Correlations of Teacher- and Child-Reported Problems*

	Teacher-Report of Externalizing Problems				Child-Report of Externalizing Problems		
	T2	T3	T4	T5	T3	T4	T5
Teacher-Report of Internalizing							
T2	**.29*** **	.02	.13*	-.07	.01	-.03	.01
T3	.05	**.39*** **	.02	.08	**.02**	.05	.00
T4	.15**	.13**	**.39*** **	.04	.08	**.07**†	.00
T5	.22***	.22***	.20***	**.48*** **	.06	.18***	**.13*** **
Child-Report of Internalizing							
T3	.01	**.06**	.04	.06	**.61*** **	.32***	.20***
T4	.09†	.05	**.03**	.08*	.33***	**.61*** **	.31***
T5	.05	.00	-.04	**.01**	.20***	.21***	**.58*** **

	Child-Report of Internalizing Problems		
Child-Report of Externalizing	*Teacher-Report of Internalizing*	T3	T4	T5
T3 .15*	**.21*** ** .18** .20**	T2	-.06	.06	.05
T4 .10†	.22*** **.26*** ** .26***	T3	**.12***	.11*	.09*
T5 .14*	.15*** .18*** **.25*** **	T4	.09†	**.07***	.04
		T5	.06	.13**	**.10***

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. Values in bold signify correlations at the same point in time.

Table 8: *Correlations between Average Problem Behaviors and Distal Outcomes*

	Mother Ext.	Father Ext.	Teacher Ext.	Child Ext.	Mother Int.	Father Int.	Teacher Int.	Child Int.
Cohabitation age	-.12†	-.03	-.02	-.07	-.07	-.04	.01	-.16*
Marital age	-.07	.01	-.07	-.07	.00	.06	-.07	-.08
Age at first child	-.17*	-.23**	-.24**	-.07	-.06	-.05	-.15†	.02
Relationship quality	-.13*	-.12*	-.23***	-.14*	-.13*	-.09	-.24**	-0.06
Highest grade completed	-.29***	-.25***	-.34***	-.21***	-.13**	-.10*	-.28***	-.08†
Highest degree obtained	-.24***	-.22***	-.28***	-.14**	-.10*	-.07	-.26***	-.06
Individual income	-.09†	-.14**	-.19***	-.11*	-.05	-.05	-.20***	-.12**
Family income	-.15**	-.12*	-.22***	-.10†	-.08	.00	-.27***	-.10†
Prior arrest	.24***	.20***	.26***	.22***	.09*	.06	.13**	.05
Legal infractions	.24***	.17***	.17***	.24***	.11*	.02	.13**	.16***
Alcohol consumed	.13**	.11*	.16***	.18***	-.01	.05	.01	-.04
Alcohol-use problems	.12**	.04	.09*	.16***	.02	.02	-.02	.04

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. Informant reports of problem behaviors were averaged across wave within each child.

Table 9: *Correlations between Average Problem Behaviors and Childhood Variables*

	Mother Ext.	Father Ext.	Teacher Ext.	Child Ext.	Mother Int.	Father Int.	Teacher Int.	Child Int.
Child FSIQ	-.19***	-.19***	-.28***	-.04	-.06	-.05	-.20***	-.05
Fathers:								
Highest grade completed	-.21***	-.23***	-.23***	-.13***	-.06†	-.09*	-.19***	-.05
Highest degree obtained	-.15***	-.20***	-.22***	-.12**	-.04	-.08*	-.19***	-.03
Income	-.24***	-.24***	-.30***	-.17***	-.06†	-.12**	-.16***	-.80*
Average depressive symptoms	.19***	.25***	.07†	.17***	.14***	.26***	.09*	.15***
Average neuroticism	.14***	.23***	.11**	.14***	.06	.24***	.09**	.16***
Relationship quality	-.14***	-.17***	-.06†	-.13***	-.21**	-.16***	-.02	-.11**
Mothers:								
Highest grade completed	-.18***	-.18***	-.19***	-.15***	-.05†	-.07*	-.09**	-.06†
Highest degree obtained	-.12***	-.12**	-.13***	-.10**	-.05	-.06†	-.10**	-.03
Income	-.22***	-.25***	-.30***	-.15***	-.03	-.08*	-.16***	-.08*
Average depressive symptoms	.28***	.18***	.16***	.17***	.31***	.17***	.08*	.18***
Average neuroticism	.27***	.15***	.12***	.15***	.31***	.14***	.12***	.18***
Relationship quality	-.14***	-.17***	-.06†	-.13***	-.12**	-.16***	-.02	-.11**

Note. ***p<.001. **p<.01. *p<.05. †p<.10. Informant reports of problem behaviors were averaged across wave within each child.

Table 10: *Correlations among Distal Outcomes*

	1	2	3	4	5	6	7	8	9	10	11
1. Cohabitation age	-										
2. Marital age	.58***	-									
3. Age at first child	.22*	.46***	-								
4. Relationship quality	.14†	.05	.09	-							
5. Highest grade completed	.30***	.29***	.44***	.29***	-						
6. Highest degree obtained	.26***	.26***	.44***	.23***	.84***	-					
7. Individual income	.11	.17*	.35***	.10†	.33***	.38***	-				
8. Family income	.15*	.24**	.42***	.21**	.41***	.41***	.67***	-			
9. Prior arrest	.02	-.06	-.07	-.29***	-.28***	-.22***	-.13**	-.21***	-		
10. Legal infractions	-.08	-.10	-.12	-.21***	-.22***	-.16***	-.10*	-.31***	.55***	-	
11. Alcohol consumed	.01	.21**	.10	.00	-.11*	.09†	.11*	.04	.20***	.19***	-
12. Alcohol-use problems	.15*	.24**	.09	.01	-.02	.04	.07†	.00	.31***	.22***	.61***

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10.

Table 11: *Best-Fitting Model Trajectory Estimates: Externalizing Behavior Problems*

	Mother T1-T5	Father T1-T5	Teacher T2-T5	Child T3-T5	Mother T3-T5	Father T3-T5	Teacher T3-T5
N	967	917	953	994	857	742	918
df	6	6	8	1	1	1	4
χ^2	4.31	6.15	13.38	.01	.00	.25	4.39
RMSEA	.000	.005	.027	.000	.000	.000	.010
RMSEA CI	.000-.034	.000-.043	.000-.051	.000-.021	.000-.000	.000-.077	.000-.052
CFI	1.000	1.000	.958	1.000	1.000	1.000	.994
TLI	1.004	1.000	.969	1.014	1.014	1.009	.996
SRMR	.022	.017	.051	.001	.000	.004	.031
Intercept M (SE)	9.10 (.31)***	8.06 (.30)***	5.90 (.30)***	10.12 (.22)***	8.42 (.31)***	7.33 (.30)***	5.74 (.31)***
Intercept Var (SE)	43.01 (4.87)***	36.05 (4.30)***	35.52 (4.55)***	26.81 (2.51)***	43.12 (5.23)***	34.02 (4.26)***	33.68 (4.97)***
Slope M (SE)	-.87 (.10)***	-.97 (.09)***		.86 (.19)***	-.67 (.17)***	-.64 (.16)***	
Slope Var (SE)	2.78 (.53)***	2.13 (.48)***		9.46 (2.18)***	4.61 (2.51)†	6.82 (2.24)**	
Slope with Intercept (SE)	1.66 (1.07)	-.30 (.92)		.83 (1.42)	3.86 (2.34)†	1.77 (1.81)	
Quad M (SE)	.09 (.06)	.17 (.06)**					
Quad Var (SE)	.69 (.19)***	.59 (.21)**					
Quad with Intercept (SE)	-1.86 (.72)**	-1.63 (.64)*					
Quad with Slope (SE)	.25 (.30)	.45 (.27)†					

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. Adult models were estimated with full data as well as abbreviated time point data (T3-T5) to match the child self-report model. Only best-fitting model estimations are detailed here; model building steps can be found in the supplemental material.

Table 12: *Best-Fitting Model Trajectory Estimates: Internalizing Behavior Problems*

	Mother T1-T5	Father T1-T5	Teacher T2-T5	Child T3-T5	Mother T3-T5	Father T3-T5	Teacher T3-T5
N	967	917	952	993	857	742	917
df	6	6	8	1	4	4	4
χ^2	14.39*	32.35***	17.95*	.66	7.04	10.85*	9.49†
RMSEA	.038	.069	.036	.000	.030	.048	.039
RMSEA CI	.012-.064	.047-.093	.013-.059	.000-.078	.000-.065	.014-.083	.000-.071
CFI	.982	.923	.753	1.000	.990	.946	.760
TLI	.969	.871	.815	1.009	.992	.959	.820
SRMR	.035	.048	.055	.008	.037	.055	.050
Intercept M (SE)	5.69 (.23)***	4.85 (.21)***	4.75 (.17)***	8.83 (2.44)***	5.59 (.21)***	4.57 (.20)***	4.66 (.18)***
Intercept Var (SE)	19.95 (2.71)***	15.46 (2.26)***	7.21 (1.19)***	24.22 (2.53)***	17.81 (2.35)***	12.37 (1.47)***	8.29 (1.42)***
Slope M (SE)	.19 (.07)**	-.03 (.07)		-.39 (.21)†			
Slope Var (SE)	1.09 (.35)**	.05 (.36)**		6.87 (3.15)*			
Slope with Intercept (SE)	1.38 (.54)**	.76 (.38)*		-2.31 (2.09)			
Quad M (SE)	-.17 (.05)**	-.10 (.05)*					
Quad Var (SE)	.32 (.15)*	.41 (.14)**					
Quad with Intercept (SE)	-1.93 (.53)***	-1.47 (.42)**					
Quad with Slope (SE)	-.10 (.17)	-.03 (.20)					

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. Adult models were estimated with full data as well as abbreviated time point data (T3-T5) to match the child self-report model. Only best-fitting model estimations are detailed here; model building steps can be found in the supplemental material.

Table 13: *Outcomes from Problem Behavior Trajectories: Mother-Reported Externalizing Problems, T1-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	Outcome on Quad (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Marital age	12.9%	-6.70-32.50%	23.75 (.37)***	-.02 (.03)	-.43 (.18)*	-.48 (.53)	35.9%	Yes
Grade completion	17.1%	8.08-26.12%	15.00 (.30)***	-.10 (.02)***	-.35 (.12)**	.01 (.32)	47.6%	Yes
Degree obtained	12.2%	5.34-19.06%	1.24 (.14)***	-.04 (.01)***	-.15 (.06)**	-.07 (.14)	37.6%	Yes
Family income	8.5%	-0.71-17.71%	4.74 (.24)***	-.04 (.02)*	-.18 (.10)†	-.30 (.24)	29.1%	Yes
Arrest	10.3%	-0.09-20.69%	*Threshold:* 1.27 (.18)***	.31 (.08)***	.15 (.12)	-.02 (.15)	24.9%	Yes
Legal infractions	10.2%	2.56-17.84%	.51 (.14)***	.04 (.01)***	.14 (.06)*	.10 (.16)	15.0%	Yes
Alcohol consumed	2.7%	-1.22-6.62%	10.35 (1.01)***	.22 (.08)**	-.28 (.40)	.85 (.93)	13.9%	No
Alcohol-use problems	5.3%	-1.95-12.55%	2.36 (.60)***	.13 (.05)**	-.36 (.26)	-.23 (.55)	12.9%	Yes

Note. ***p<.001. **p<.01. *p<.05. †p<.10. As a reminder, T1-T5 mother-reported externalizing problems best fit a quadratic shape. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 14: *Outcomes from Problem Behavior Trajectories: Father-Reported Externalizing Problems, T1-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	Outcome on Quad (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	20.3%	-2.63-43.23%	22.70 (1.01)***	-.11 (.07)	-.90 (.42)*	.81 (1.18)	33.7%	Yes
Grade completion	12.6%	1.62-23.58%	14.66 (.38)***	-.08 (.03)**	-.45 (.19)*	-.02 (.46)	49.2%	Yes
Degree obtained	9.2%	0.77-17.63%	1.12 (.19)***	-.04 (.01)**	-.19 (.09)*	03 (.21)	37.2%	Yes
Arrest	5.5%	-0.18-11.18%	*Threshold:* 2.32 (.44)***	.09 (.03)**	.13 (.21)	.07 (.46)	20.4%	Yes
Legal infractions	7.9%	-1.51-17.31%	.71 (.23)**	.03 (.01)*	.19 (.11)†	-.04 (.26)	14.3%	Yes
Alcohol consumed	2.1%	-0.84-5.04%	11.31 (1.34)***	.19 (.09)*	.22 (.60)	.05 (1.38)	14.3%	No

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T1-T5 father-reported externalizing problems best fit a quadratic shape. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 15: *Outcomes from Problem Behavior Trajectories: Teacher-Reported Externalizing Problems, T2-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	7.8%	-2.78-18.38%	23.60 (.44)***	-.15 (.06)**	23.5%	No
Relationship quality	12.7%	1.53-23.87%	116.84 (1.68)***	-1.10 (.26)***	20.0%	Yes
Grade completion	20.9%	11.69-30.11%	15.41 (.17)***	-.17 (.02)***	46.5%	Yes
Degree obtained	13.7%	7.23-20.17%	1.40 (.08)***	-.07 (.01)***	36.5%	Yes
Individual income	7.5%	1.42-13.58%	7.45 (.16)***	-.11 (.02)***	15.1%	Yes
Family income	9.1%	1.26-16.94%	4.97 (.14)***	-.08 (.02)***	20.8%	Yes
Arrest	18.2%	7.62-28.78%	*Threshold:* 2.66 (.26)***	.14 (.03)***	25.6%	Yes
Legal infractions	7.7%	1.04-14.36%	.47 (.09)***	.06 (.02)***	10.4%	Yes
Alcohol consumed	4.2%	-1.29-9.69%	1.41 (.08)***	.21 (.07)**	13.4%	Yes

Note. ***p<.001. **p<.01. *p<.05. †p<.10. As a reminder, T2-T5 teacher-reported externalizing problems best fit an intercept-only model. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 16: *Outcomes from Problem Behavior Trajectories: Child-Reported Externalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Relationship quality	8.4%	-3.75-20.55%	115.71 (3.05)***	-.30 (.30)	-1.59 (.76)*	16.4%	No
Grade completion	6.1%	-0.17-12.37%	15.38 (.30)***	-.09 (.03)**	-.11 (.09)	45.8%	Yes
Degree obtained	2.8%	-1.12-6.72%	1.33 (.14)***	-.03 (.01)*	-.03 (.04)	34.7%	No
Arrest	10.2%	2.16-18.24%	*Threshold:* 3.00 (.36)***	.11 (.03)***	.06 (.08)	23.4%	Yes
Legal infractions	6.5%	0.42-12.58%	.18 (.15)	.06 (.02)***	.03 (.04)	10.9%	Yes
Alcohol consumed	6.7%	-0.55-13.95%	9.10 (.95)***	.37 (.09)***	-.28 (.22)	15.8%	Yes
Alcohol-use problems	3.9%	-1.00-8.80%	2.14 (.60)***	.16 (.06)**	.11 (.15)	10.7%	Yes

Note. ***p<.001. **p<.01. *p<.05. †p<.10. As a reminder, T3-T5 child-reported externalizing problems best fit a linear shape. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 17: *Outcomes from Problem Behavior Trajectories: Mother-Reported Externalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	4.1%	-1.39-9.59%	23.48 (.59)***	-.09 (.04)*	-.08 (.26)	34.2%	Yes
Grade completion	15.7%	5.90-25.50%	15.27 (.29)***	-.12 (.02)***	-.15 (.16)	48.2%	Yes
Degree obtained	13.0%	4.18-21.82%	1.29 (.14)***	-.04 (.01)***	-.10 (.07)	32.8%	Yes
Arrest	12.2%	3.97-20.43%	*Threshold:* 2.55 (.28)***	.10 (.02)***	.04 (.12)	26.6%	Yes
Legal infractions	11.5%	-0.06-23.06%	.52 (.13)***	.04 (.01)**	.11 (.08)	16.9%	Yes

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T3-T5 mother-reported externalizing problems best fit a linear shape. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 18: *Outcomes from Problem Behavior Trajectories: Father-Reported Externalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	12.7%	2.12-23.28%	24.32 (.52)***	-.20 (.05)***	-.01 (.16)	27.3%	Yes
Grade completion	12.1%	1.12-23.08%	15.01 (.24)***	-.09 (.03)**	-.19 (.11)†	48.1%	Yes
Degree obtained	8.0%	0.75-15.25%	1.28 (.12)***	-.04 (.01)**	-.07 (.05)	36.6%	Yes
Individual income	1.5%	-1.24-4.24%	7.23 (.21)***	-.05 (.02)*	.03 (.09)	14.7%	No
Family income	4.8%	-2.26-11.86%	4.76 (.17)***	-.04 (.02)*	-.10 (.07)	19.5%	No
Arrest	6.9%	0.24-13.56%	*Threshold:* 2.18 (.26)***	.07 (.02)**	.10 (.09)	22.6%	Yes
Legal infractions	6.6%	-0.85-14.05%	.53 (.11)***	.04 (.01)**	.07 (.05)	12.0%	Yes
Alcohol consumed	2.1%	-0.84-5.04%	11.18 (.73)***	.20 (.07)**	.02 (.28)	13.9%	No

Note. ***p<.001. **p<.01. *p<.05. †p<.10. As a reminder, T3-T5 father-reported externalizing problems best fit a linear shape. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 19: *Outcomes from Problem Behavior Trajectories: Teacher-Reported Externalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	9.3%	-2.07-20.67%	23.69 (.44)***	-.17 (.06)**	21.0%	No
Relationship quality	13.5%	1.54-25.46%	117.03 (1.72)***	-1.16 (.28)***	21.9%	Yes
Grade completion	22.0%	11.61-32.39%	15.44 (.19)***	-.18 (.03)***	46.5%	Yes
Degree obtained	14.2%	7.34-21.06%	1.41 (.09)***	-.07 (.01)***	36.4%	Yes
Individual income	7.4%	1.32-13.48%	7.45 (.17)***	-.11 (.02)***	14.9%	Yes
Family income	7.9%	0.26-15.54%	4.94 (.15)***	-.08 (.02)***	20.2%	No
Arrest	19.2%	7.64-30.76%	1.34 (.10)***	.44 (.07)***	26.3%	Yes
Legal infractions	6.9%	0.43-13.37%	.48 (.09)***	.05 (.02)***	9.3%	Yes
Alcohol consumed	3.9%	-1.78-9.58%	11.18 (.63)***	.27 (.11)*	13.2%	Yes

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T3-T5 teacher-reported externalizing problems best fit an intercept-only model. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 20: *Outcomes from Problem Behavior Trajectories: Father-Reported Internalizing Problems, T1-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	Outcome on Quad (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	30.0%	-2.54-62.54%	22.50 (.76)***	.02 (.17)	-1.89 (.72)**	.51 (1.43)	39.6%	Yes

Note. ***p<.001. **p<.01. *p<.05. †p<.10. As a reminder, T1-T5 father-reported internalizing problems best fit a quadratic shape. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 21: *Outcomes from Problem Behavior Trajectories: Teacher-Reported Internalizing Problems, T2-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Relationship quality	16.8%	1.51-32.09%	123.88 (3.23)***	-2.77 (.66)***	20.6%	Yes
Grade completion	22.4%	9.07-35.73%	16.26 (.39)***	-.39 (.08)***	44.7%	Yes
Degree obtained	17.4%	6.42-28.38%	1.81 (.18)***	-.17 (.04)***	36.6%	Yes
Individual income	10.8%	0.41-21.19%	8.15 (.38)***	-.28 (.08)***	16.2%	Yes
Family income	16.8%	4.20-33.20%	5.72 (.29)***	-.26 (.06)***	28.3%	Yes
Arrest	7.9%	-2.88-18.68%	*Threshold:* 2.64 (.45)***	.20 (.08)*	18.2%	No
Legal infractions	6.0%	-0.66-12.66%	.28 (.16)†	.11 (.04)**	9.6%	Yes

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T2-T5 teacher-reported internalizing problems best fit an intercept-only model. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 22: *Outcomes from Problem Behavior Trajectories: Child-Reported Internalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	Outcome on Slope (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Individual income	2.6%	-1.52-6.72%	7.50 (.29)***	-.08 (.03)*	-0.5 (.13)	12.7%	No

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T3-T5 child-reported internalizing problems best fit a linear shape. Only the outcome that was significantly predicted by a growth factor has been detailed here. Model coefficient values are provided without covariates included.

Table 23: *Outcomes from Problem Behavior Trajectories: Mother-Reported Internalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Relationship quality	3.2%	-3.86-10.26%	115.69 (2.40)***	-.84 (.40)*	13.4%	No
Grade completion	3.6%	-0.71-7.91%	14.97 (.21)***	-.10 (.03)**	41.3%	Yes
Degree obtained	3.3%	-1.40-8.00%	1.24 (.10)***	-.04 (.01)**	32.7%	Yes
Arrest	2.8%	-2.10-7.70%	*Threshold:* 2.06 (.25)***	.07 (.03)*	20.2%	No
Legal infractions	1.6%	-1.54-4.74%	.58 (.12)***	.04 (.02)*	8.6%	No

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T3-T5 mother-reported internalizing problems best fit an intercept-only model. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 24: *Outcomes from Problem Behavior Trajectories: Father-Reported Internalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	5.8%	-4.98-16.58%	23.68 (.56)***	-.23 (.11)*	23.7%	No

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T3-T5 father-reported internalizing problems best fit an intercept-only model. Only the outcome that was significantly predicted by a growth factor has been detailed here. Model coefficient values are provided without covariates included.

Table 25: *Outcomes from Problem Behavior Trajectories: Teacher-Reported Internalizing Problems, T3-T5*

Distal Outcome	% Variance Explained	95% CI for Variance Explained	Outcome Intercept (SE)	Outcome on Intercept (SE)	% Variance Explained with all Covariates	Prediction Maintained with Covariates?
Age at first child	6.3%	-5.26-17.86%	23.95 (.71)***	-.29 (.13)*	21.1%	No
Relationship quality	23.6%	6.16-41.04%	124.96 (3.05)***	-3.07 (.63)***	25.6%	Yes
Grade completion	25.9%	10.61-41.19%	16.24 (.40)***	-.40 (.08)***	45.4%	Yes
Degree obtained	19.7%	7.55-31.85%	1.79 (.19)***	-.17 (.04)***	37.0%	Yes
Individual income	14.2%	1.85-26.55%	8.22 (.37)***	-.30 (.08)***	17.8%	Yes
Family income	22.0%	5.93-38.07%	5.70 (.29)***	-.26 (.06)***	30.7%	Yes
Arrest	7.1%	-3.09-17.29%	*Threshold:* 2.51 (.41)***	.17 (.07)*	17.7%	No
Legal infractions	4.5%	-1.58-10.58%	.38 (.16)*	.09 (.04)*	8.0%	No

Note. ***p<.001. **p<.01. *p<.05. $^{\dagger}p$<.10. As a reminder, T3-T5 mother-reported internalizing problems best fit an intercept-only model. Only outcomes that were significantly predicted by a growth factor are detailed here. Model coefficient values are provided without covariates included.

Table 26: *Summary Table of Outcomes Predicted by Each Parent Model of Child Problem Behaviors*

	Mean Outcome Variance Explained in Significant Models	Mean Outcome Variance in Models Maintaining Significance after Including Covariates	Range of Variance Explained in Sig. Models (& Only Models Maintaining Sig. After Covar. Incl., if Varied)	Number of Outcomes Significantly Predicted (& After Including Covariates)	List of Outcomes Significantly Predicted (Non-Significant Outcome after Covariates Included)
Mothers:					
Mother T1-T5 Externalizing	9.90%	10.93%	2.70-17.10% (5.30-17.10%)	8 (7)	Marital age, grade completion, degree obtained, family income, prior arrest, legal infractions, (alcohol consumed), alcohol-use problems
Mother T3-T5 Externalizing	11.30%	11.30%	4.10-15.70%	5 (5)	Age at first child, grade completion, degree obtained, prior arrest, legal infractions
Mother T1-T5 Internalizing	No outcomes predicted.	--	--	--	--
Mother T3-T5 Internalizing	2.90%	3.45%	1.60-3.60% (3.30-3.60%)	5 (2)	(Relationship quality), grade completion, degree obtained, (prior arrest), (legal infractions)
Fathers:					
Father T1-T5 Externalizing	9.60%	11.10%	2.10-20.30% (5.50-20.30%)	6 (5)	Age at first child, grade completion, degree obtained, prior arrest, legal infractions, (alcohol consumed)
Father T3-T5 Externalizing	6.84%	9.26%	1.50-12.70% (6.60-12.70%)	8 (5)	Age at first child, grade completion, degree obtained, (individual income), (family income), prior arrest, legal infractions, (alcohol consumed)

Table 26 (cont'd)

	Mean Outcome Variance Explained in Significant Models	Mean Outcome Variance in Models Maintaining Significance after Including Covariates	Range of Variance Explained in Sig. Models (& Only Models Maintaining Sig. After Covar. Incl., if Varied)	Number of Outcomes Significantly Predicted (& After Including Covariates)	List of Outcomes Significantly Predicted (Non-Significant Outcome after Covariates Included)
Father T1-T5 Internalizing	30.00%	30.00%	--	1 (1)	Age at first child
Father T3-T5 Internalizing	5.80%	--	--	1 (0)	(Age at first child)

Note. This summary table reflects the significant outcomes predicted by each growth model stemming from parents' reports.

Table 27: *Summary Table of Outcomes Predicted by Each Teacher and Child Model of Child Problem Behaviors*

	Mean Outcome Variance Explained in Significant Models	Mean Outcome Variance in Models Maintaining Significance after Including Covariates	Range of Variance Explained in Sig. Models (& Only Models Maintaining Sig. After Covar. Incl., if Varied)	Number of Outcomes Significantly Predicted (& After Including Covariates)	List of Outcomes Significantly Predicted (Non-Significant Outcome after Covariates Included)
Teachers:					
Teacher T2-T5 Externalizing	11.31%	11.75%	4.20-20.90%	9 (8)	(Age at first child), relationship quality, grade completion, degree obtained, individual income, family income, prior arrest, legal infractions, alcohol consumed
Teacher T3-T5 Externalizing	11.59%	12.44%	3.90-22.00%	9 (7)	(Age at first child), relationship quality, grade completion, degree obtained, individual income, (family income), prior arrest, legal infractions, alcohol consumed
Teacher T2-T5 Internalizing	14.29%	15.35%	6.00-22.40%	7 (6)	Relationship quality, grade completion, degree obtained, individual income, family income, (prior arrest), legal infractions
Teacher T3-T5 Internalizing	15.41%	20.08%	4.50-25.90% (14.20%-25.90%)	8 (5)	(Age at first child), relationship quality, grade completion, degree obtained, individual income, family income, (prior arrest), (legal infractions)

Table 27 (cont'd)

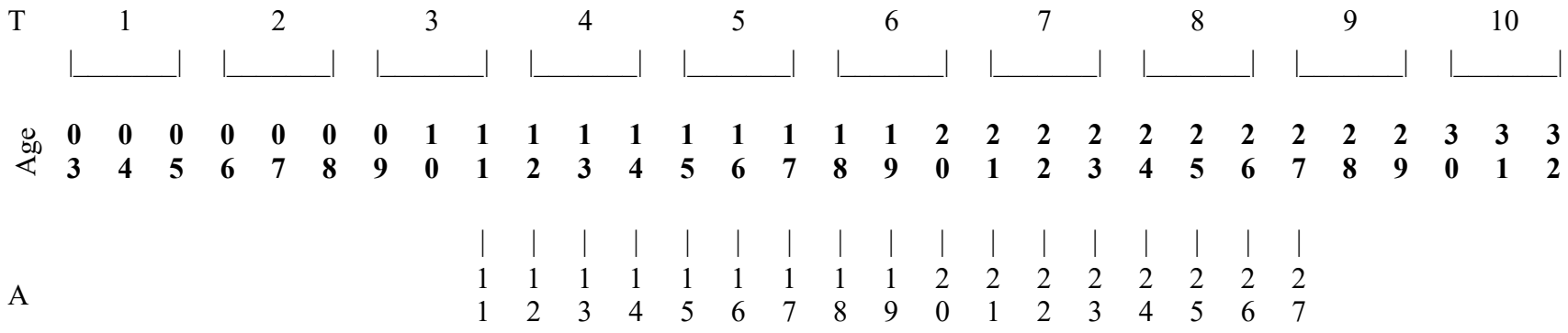
	Mean Outcome Variance Explained in Significant Models	Mean Outcome Variance in Models Maintaining Significance after Including Covariates	Range of Variance Explained in Sig. Models (& Only Models Maintaining Sig. After Covar. Incl., if Varied)	Number of Outcomes Significantly Predicted (& After Including Covariates)	List of Outcomes Significantly Predicted (Non-Significant Outcome after Covariates Included)
Children:					
Child T3-T5 Externalizing	6.37%	6.68%	2.80-10.20% (3.90-10.20%)	7 (5)	(Relationship quality), grade completion, (degree obtained), prior arrest, legal infractions, alcohol consumed, alcohol-use problems
Child T3-T5 Internalizing	2.60%	--	--	1 (0)	(Individual income)

Note. This summary table reflects the significant outcomes predicted by each growth model stemming from teacher-report and children's self-report.

APPENDIX B

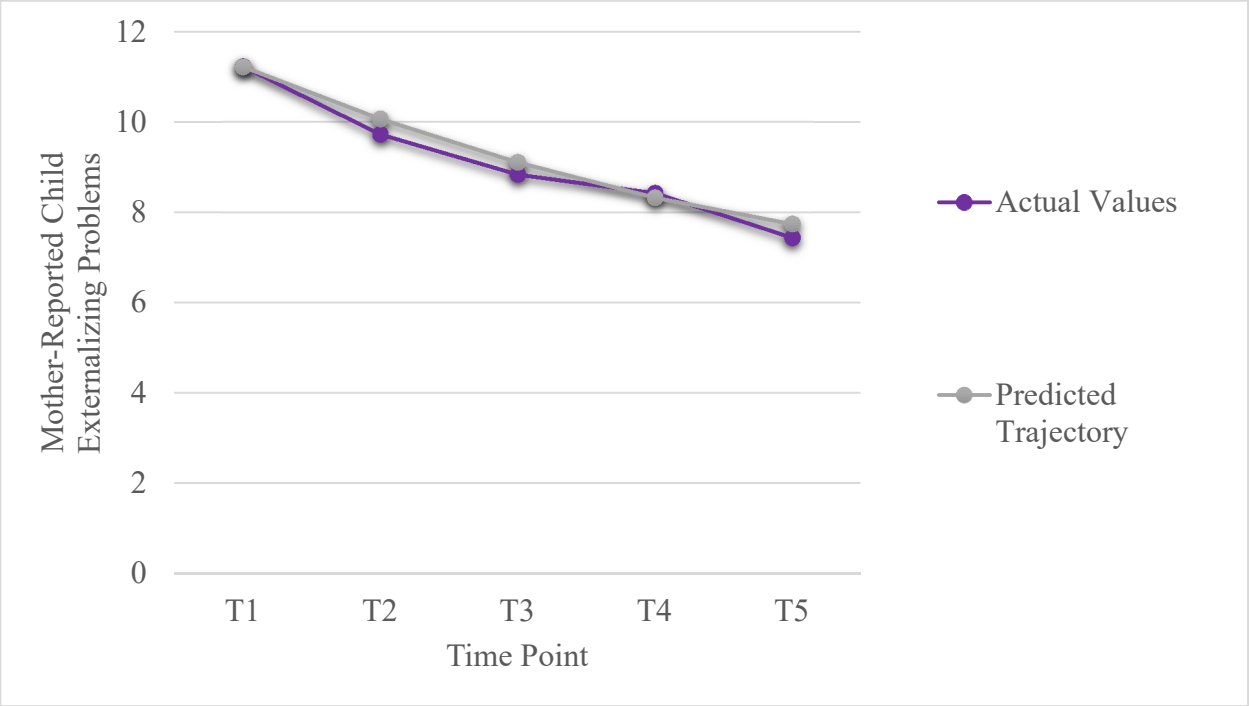
Figures

Figure 1: *Schematic of Assessments*



Note. Age=child age at assessment; T=time period; A=annual mailing. Annual mailings were not conducted during the year the T-period assessment was conducted; annuals will be used to supplement missing T-period data.

Figure 2: *Mother-Reported Child Externalizing and Internalizing Actual and Predicted Trajectories*



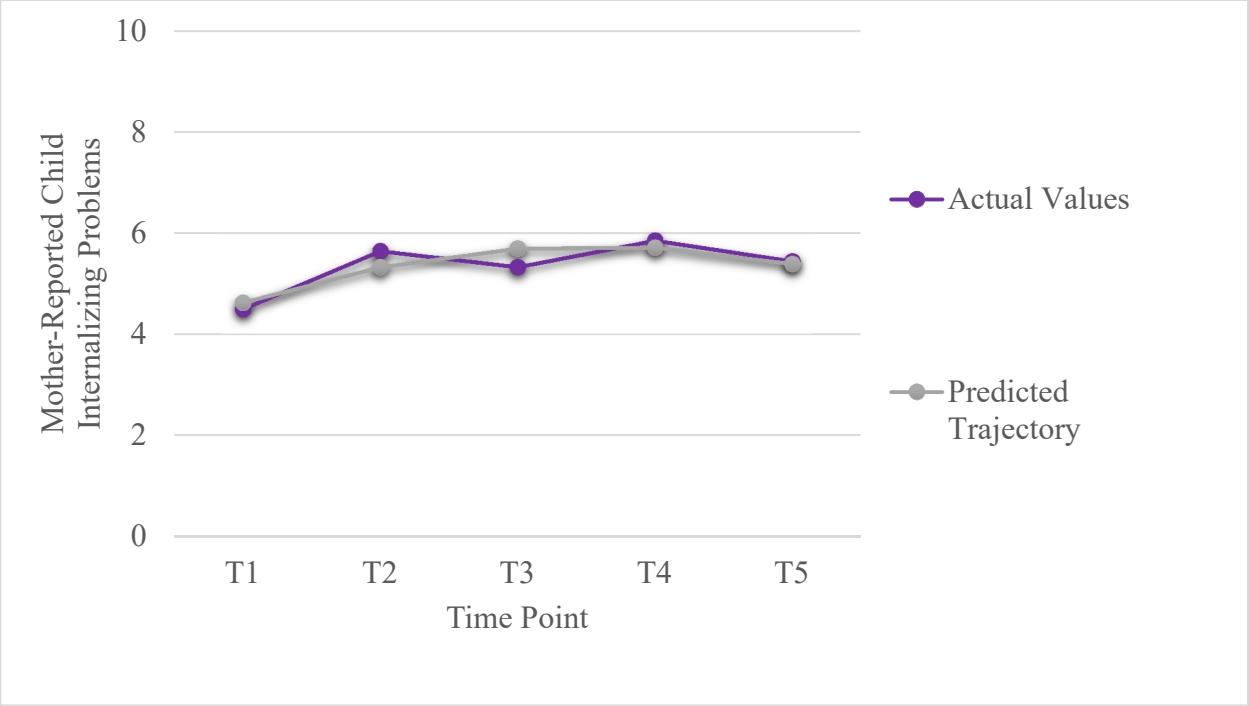
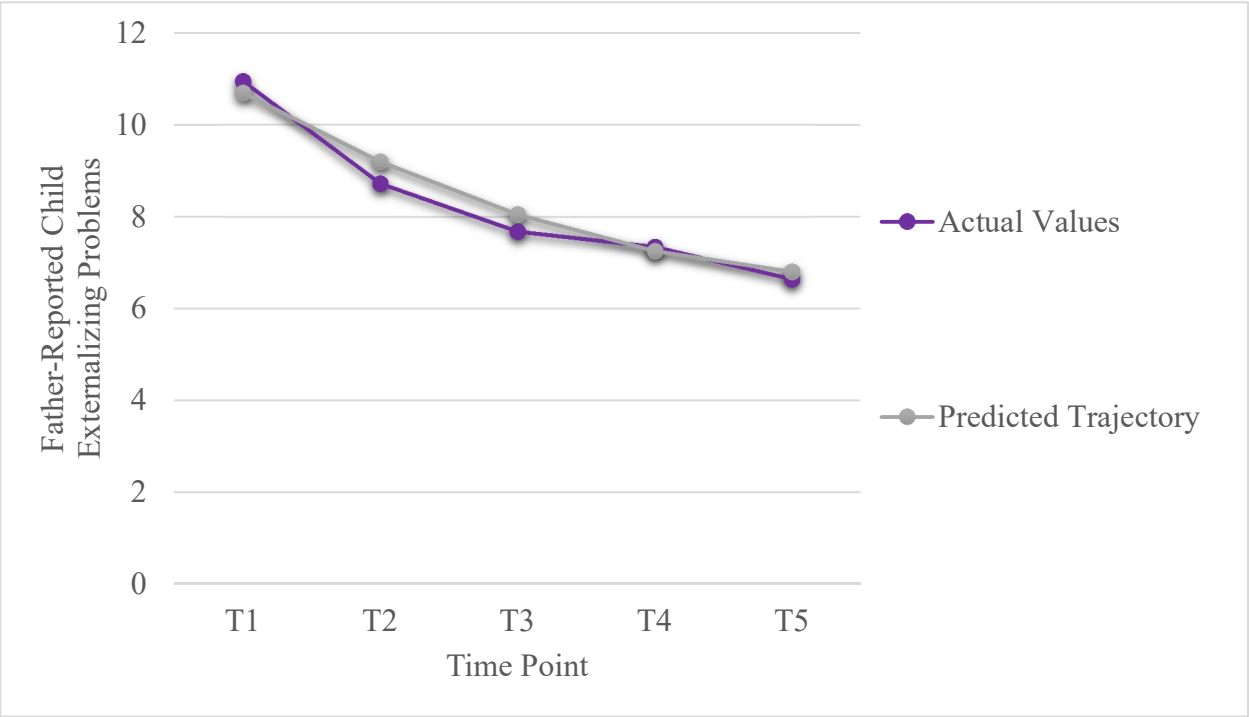


Figure 3: *Father-Reported Child Externalizing and Internalizing Actual and Predicted Trajectories*



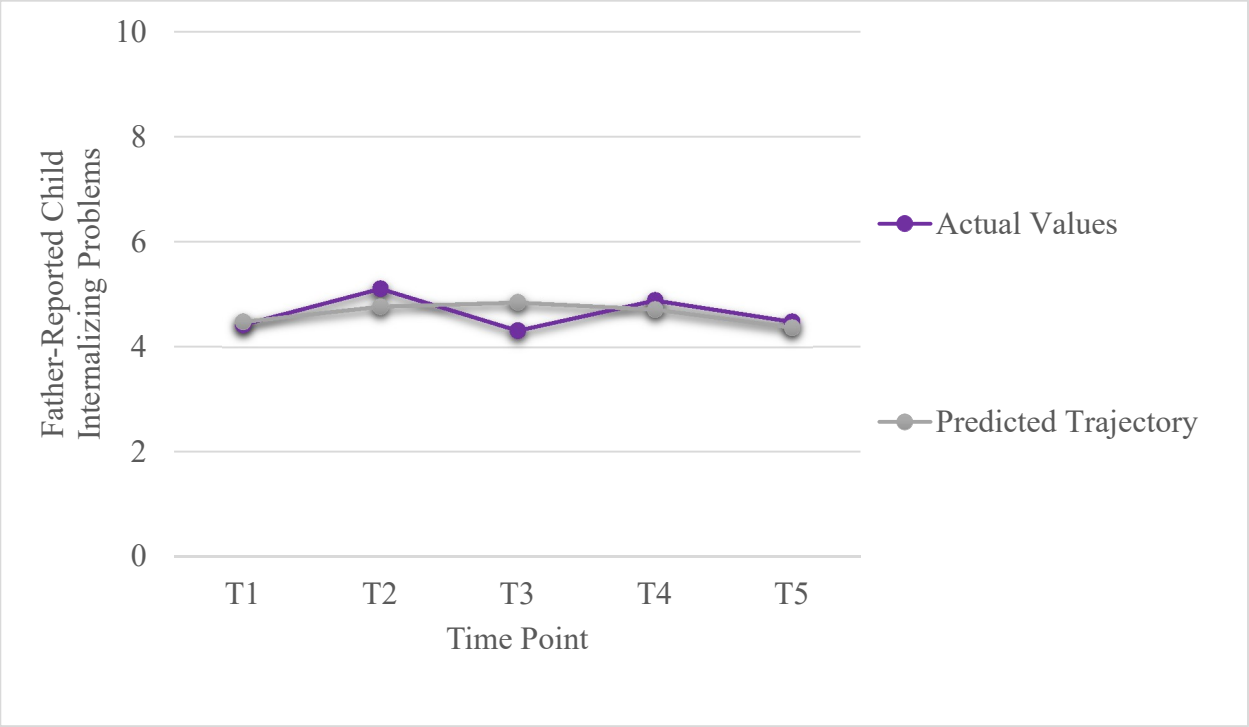
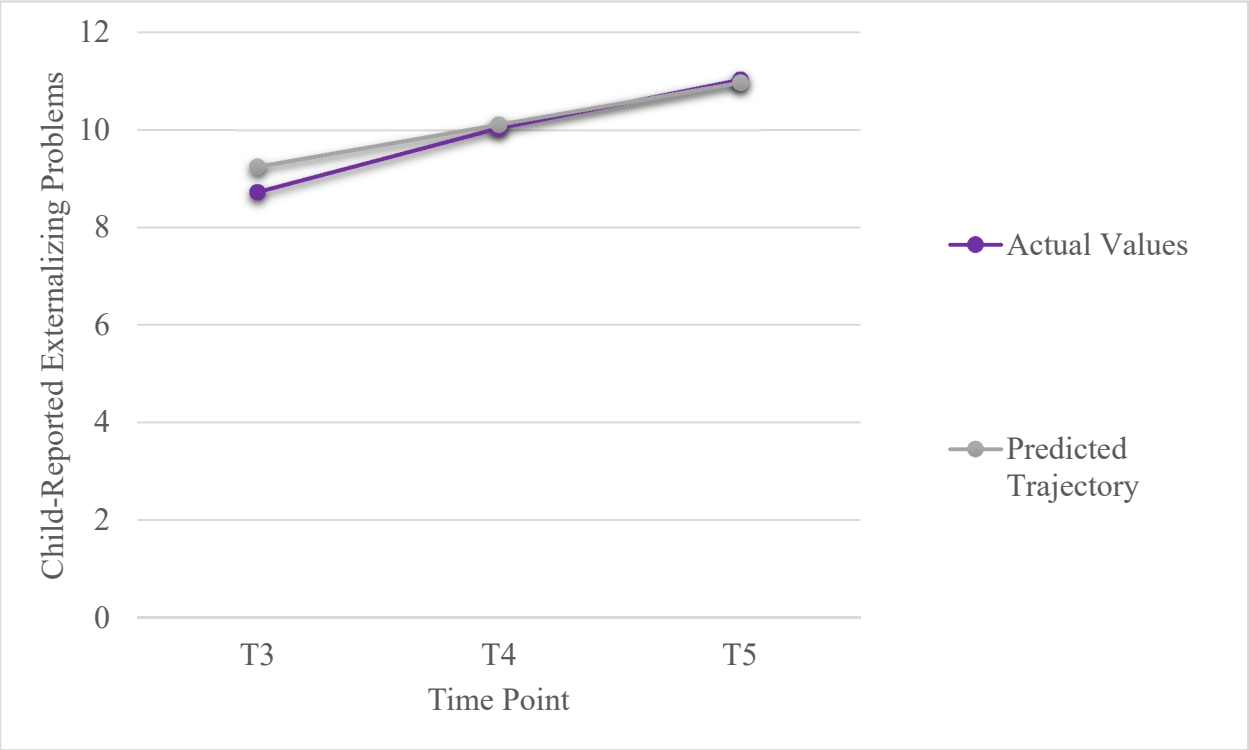


Figure 4: *Child Self-Reported Externalizing and Internalizing Actual and Predicted Trajectories*



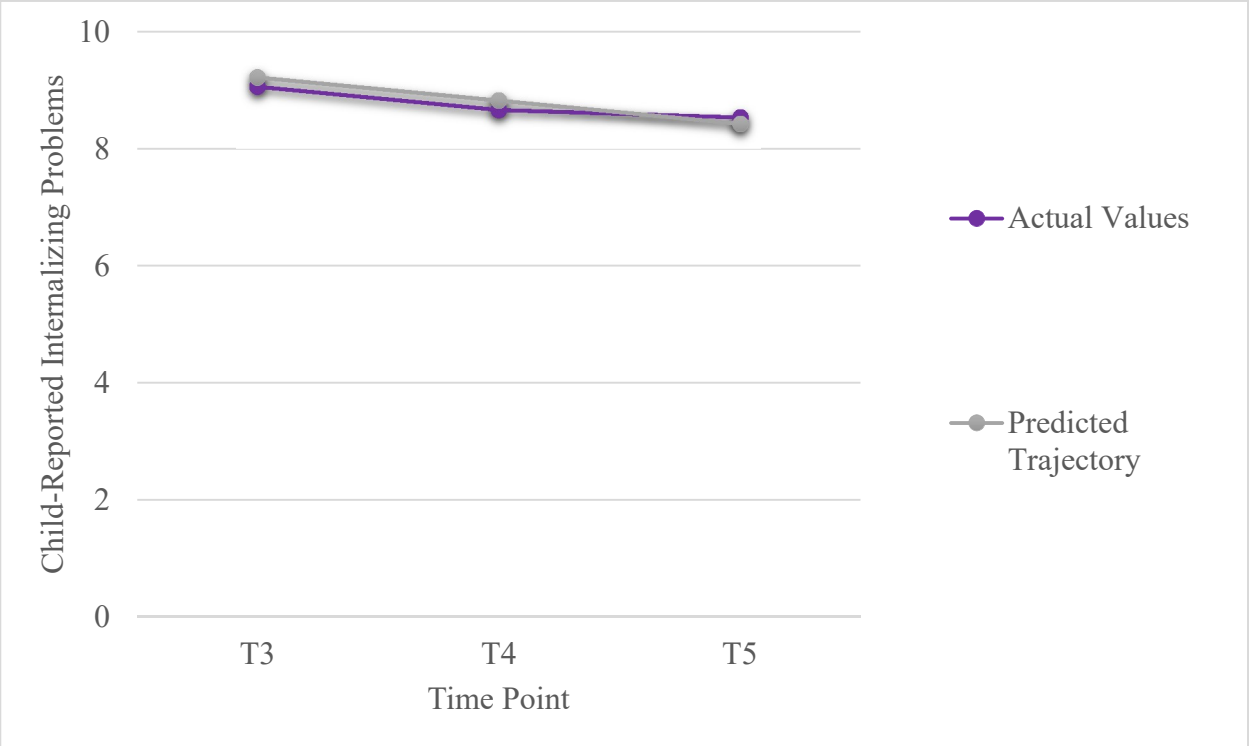
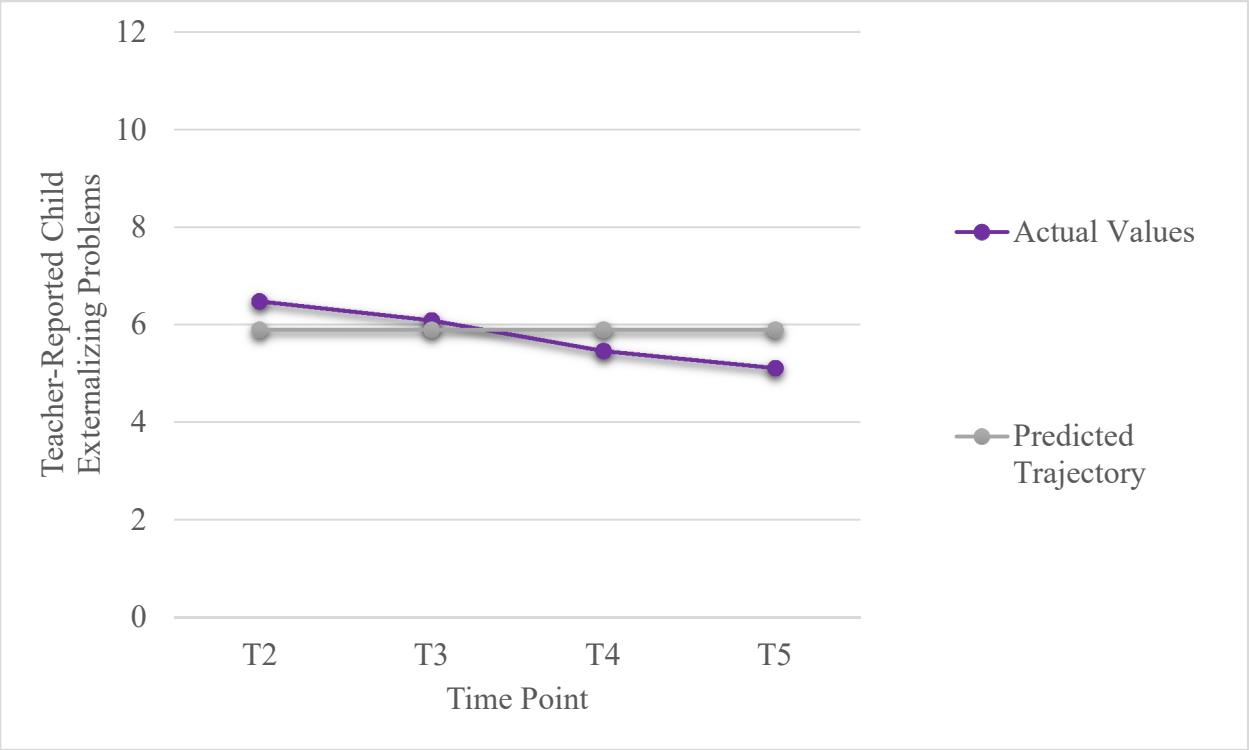


Figure 5: *Teacher-Reported Child Externalizing and Internalizing Actual and Predicted Trajectories*



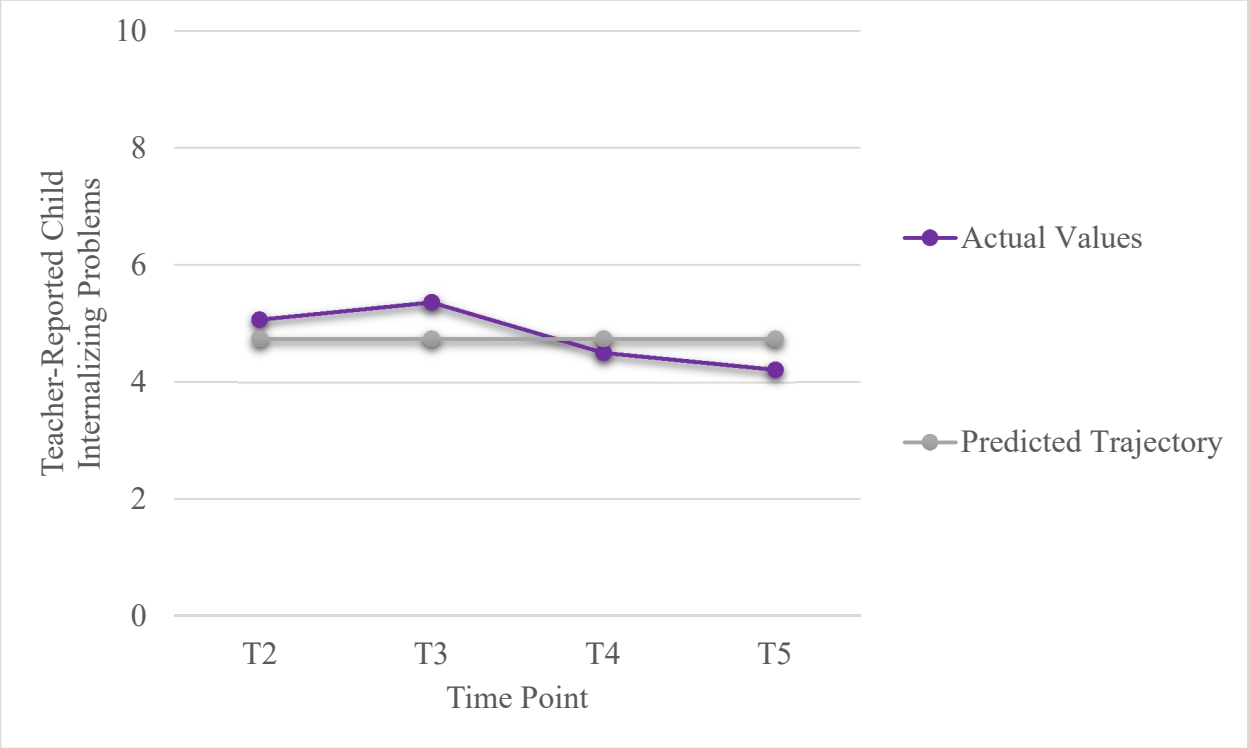
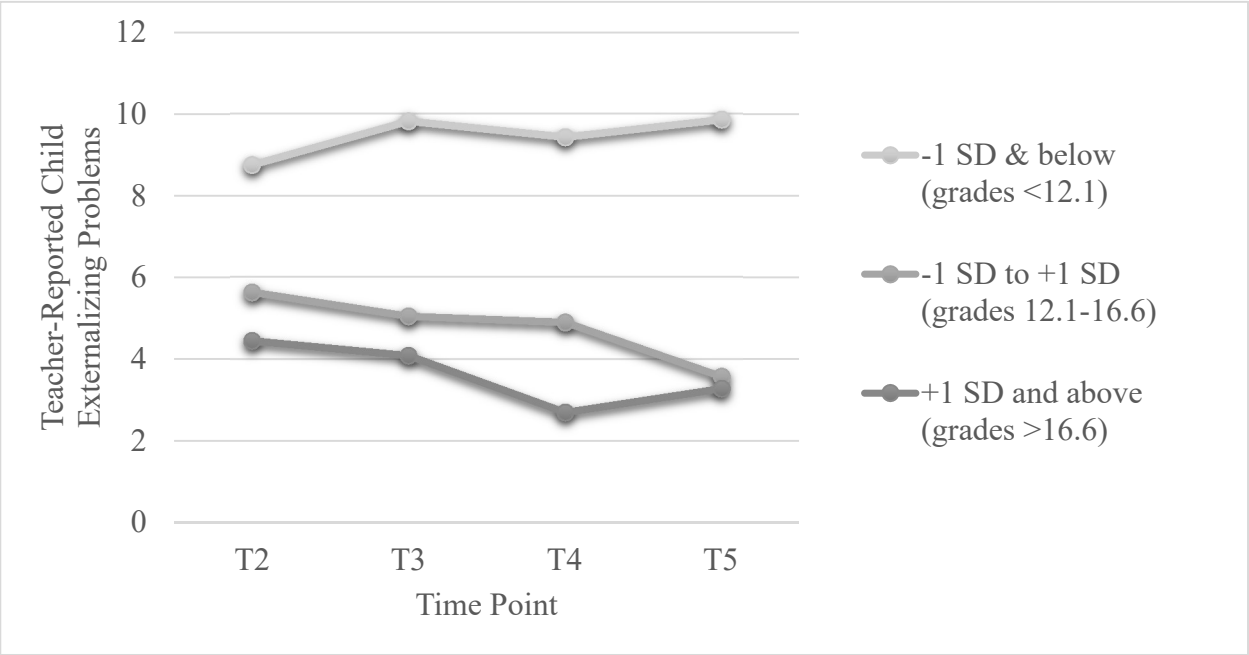
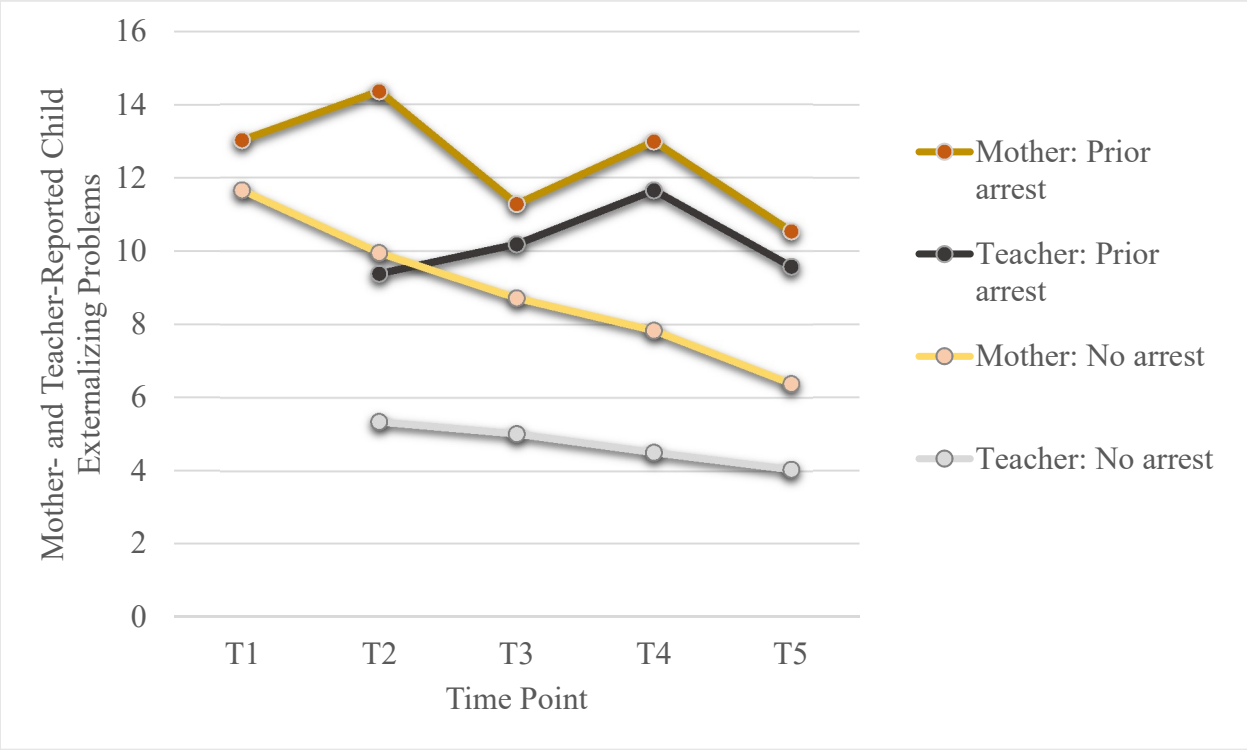


Figure 6: *Plotted Actual Values for Teacher-Reported Child Externalizing Behavior Separated by Later Educational Grade Attainment*



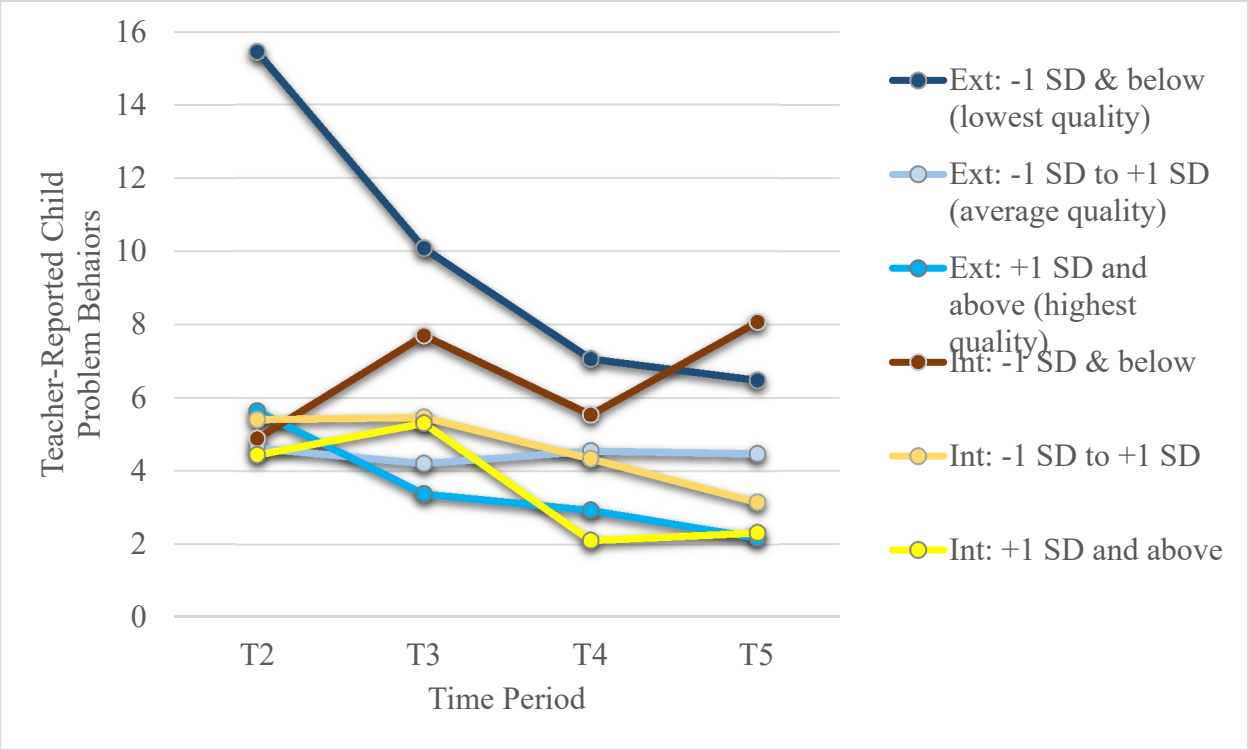
Note. The intercept of teacher-reported child externalizing problem behaviors (T2-T5) significantly predicted later child grade attainment (R^2=.209). As a reminder, T2-T5 teacher-reported externalizing problems best fit an intercept-only model. Here, the actual average values of teacher-reported child externalizing behaviors at each time point are plotted, separated by children's later educational grade attainment (*M*=14.39, *SD*=2.26).

Figure 7: *Plotted Actual Values for Mother- and Teacher-Reported Child Externalizing Behavior Separated by Later Arrest*



Note. The intercepts of both mother-reported (T1-T5) and teacher-reported child externalizing problem behaviors (T2-T5) significantly predicted later arrest (R^2=.103 for mothers, .182 for teachers). As a reminder, T1-T5 mother-reported externalizing problems best fit a quadratic model and T2-T5 teacher-reported externalizing problems best fit an intercept-only model. Here, the actual average values of mother- and teacher-reported child externalizing behaviors at each time point are plotted, separated by whether children later ever did or did not experience arrest.

Figure 8: *Plotted Actual Values for Teacher-Reported Child Externalizing and Internalizing Behavior Separated by Later Relationship Quality*



Note. The intercepts of both teacher-reported externalizing and internalizing problem behaviors significantly predicted later romantic relationship quality (R^2=.127 for externalizing, .168 for internalizing). As a reminder, T2-T5 teacher-reported externalizing as well as internalizing problems best fit intercept-only models. Here, the actual average values of teacher-reported child externalizing behaviors at each time point are plotted, separated by later romantic relationship quality. Higher scores in romantic relationship quality indicate better relationship functioning, whereas scores below one SD indicate worse later relationship functioning.

REFERENCES

REFERENCES

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Department of Psychiatry, University of Vermont.

Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science*, *15*(2), 94-98.

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment. University of Vermont. *Research Center for Children, Youth, & Families*.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213.

Amato, P. R., Landale, N. S., Havasevich-Brooks, T. C., Booth, A., Eggebeen, D. J., Schoen, R., & McHale, S. M. (2008). Precursors of young women's family formation pathways. *Journal of Marriage and Family*, *70*(5), 1271-1286.

Baltes, P. B., & Nesselroade, J. R. (Eds.). (1979). *Longitudinal research in the study of behavior and development*. Academic Press.

Bardone, A. M., Moffitt, T. E., Caspi, A., Dickson, N., & Silva, P. A. (1996). Adult mental health and social outcomes of adolescent girls with depression and conduct disorder. *Development and Psychopathology*, *8*(4), 811-829.

Barlow, D. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford Press.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77-100.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561-571.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238.

Berg-Nielsen, T. S., Solheim, E., Belsky, J., & Wichstrom, L. (2012). Preschoolers' psychosocial problems: In the eyes of the beholder? Adding teacher characteristics as determinants of discrepant parent–teacher reports. *Child Psychiatry & Human Development*, *43*(3), 393-413.

Bingham, C. R., Loukas, A., Fitzgerald, H. E., & Zucker, R. A. (2003). Parental ratings of son's behavior problems in high-risk families: Convergent validity, internal structure, and interparent agreement. *Journal of Personality Assessment*, *80*(3), 237-251.

Blair, C., & Raver, C. C. (2012). Child development in the context of adversity: experiential canalization of brain and behavior. *American Psychologist*, *67*(4), 309.

Bongers, I. L., Koot, H. M., Van Der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, *75*(5), 1523-1537.

Bongers, I. L., Koot, H. M., Van der Ende, J., & Verhulst, F. C. (2003). The normative development of child and adolescent problem behavior. *Journal of Abnormal Psychology*, *112*(2), 179.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., ... & Lynam, D. R. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, *39*(2), 222.

Brookman-Frazee, L., Haine, R. A., Gabayan, E. N., & Garland, A. F. (2008). Predicting frequency of treatment visits in community-based youth psychotherapy. *Psychological Services*, *5*(2), 126.

Carroll, J., Willoughby, B., Badger, S., Nelson, L., McNamara, C. B., & Madsen, S. (2007). So close, yet so far away: The impact of varying marital horizons on emerging adulthood. *Journal of Adolescent Research, 22(3),* 219–247.

Caspi, A., Bem, D. J., & Elder, G. H. (1989). Continuities and consequences of interactional styles across the life course. *Journal of Personality*, *57*(2), 375-406.

Chan, S. W., Goodwin, G. M., & Harmer, C. J. (2007). Highly neurotic never-depressed students have negative biases in information processing. *Psychological Medicine*, *37*(9), 1281-1291.

Chassin, L., Pitts, S. C., & Prost, J. (2002). Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: predictors and substance abuse outcomes. *Journal of Consulting and Clinical Psychology*, *70*(1), 67.

Chi, T. C., & Hinshaw, S. P. (2002). Mother–child relationships of children with ADHD: The role of maternal depressive symptoms and depression-related distortions. *Journal of Abnormal Child Psychology*, *30*(4), 387-400.

Child Trends (2006). Charting parenthood: A statistical portrait of fathers and mothers in America. Retrieved from http://fatherhood.hhs.gov/charting02/index.htm. Accessed 1 Jun 2018.

Cicchetti, D., & Toth, S. L. (1995). A developmental psychopathology perspective on child abuse and neglect. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*(5), 541-565.

Clark, D. A., Durbin, C. E., Hicks, B. M., Iacono, W. G., & McGue, M. (2017). Personality in the age of industry: Structure, heritability, and correlates of personality in middle childhood from the perspective of parents, teachers, and children. *Journal of Research in Personality*, *67*, 132-143.

Colder, C. R., Mott, J. A., & Berman, A. S. (2002). The interactive effects of infant activity level and fear on growth trajectories of early childhood behavior problems. *Development and Psychopathology*, *14*(1), 1-23.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory*. Odessa, FL: Psychology Assessment Resources.

Coyne, C. A., & D'Onofrio, B. M. (2012). Some (but not much) progress toward understanding teenage childbearing: A review of research from the past decade. In *Advances in child development and behavior* (Vol. 42, pp. 113-152). JAI.

Darney, D., Reinke, W. M., Herman, K. C., Stormont, M., & Ialongo, N. S. (2013). Children with co-occurring academic and behavior problems in first grade: Distal outcomes in twelfth grade. *Journal of School Psychology*, *51*(1), 117-128.

Day, J. C., & Newburger, E. C. (2002). The Big Payoff: Educational Attainment and Synthetic Estimates of Work-Life Earnings. Special Studies. Current Population Reports.

De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *40*(1), 1-9.

De Los Reyes, A., & Prinstein, M. J. (2004). Applying depression-distortion hypotheses to the assessment of peer victimization in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *33*(2), 325-335.

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858.

De Los Reyes, A., Goodman, K. L., Kliewer, W., & Reid-Quinones, K. (2008). Whose depression relates to discrepancies? Testing relations between informant characteristics and informant discrepancies from both informants' perspectives. *Psychological Assessment*, *20*(2), 139.

De Los Reyes, A., Goodman, K. L., Kliewer, W., & Reid-Quinones, K. (2010). The longitudinal consistency of mother–child reporting discrepancies of parental monitoring and their ability to predict child delinquent behaviors two years later. *Journal of Youth and Adolescence*, *39*(12), 1417-1430.

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*, *37*(5), 637-652.

Ellis, D. A., Pallas, D. M., Puttler, L. I., Refior, S. K., Fitzgerald, H. E., & Zucker, R. A. (2002). The clinical and social ecology of childhood for children of alcoholics: Description of a study and implications for a differentiated social policy. In *Children of Addiction* (pp. 125-158). Routledge.

Essau, C.A. (2003). Conduct and oppositional defiant disorder: Epidemiology, risk factors and treatment. Mahwah, NJ: Erlbaum.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, *26*(1), 57-63.

Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2006). Prognostic value of parent–adolescent disagreement in a referred sample. *European Child & Adolescent Psychiatry*, *15*(3), 156-162.

Ferdinand, R. F., Verhulst, F. C., & Wiznitzer, M. (1995). Continuity and change of self-reported problem behaviors from adolescence into young adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*(5), 680-690.

Foster, K. T., Hicks, B. M., & Zucker, R. A. (2018). Positive and negative effects of internalizing on alcohol use problems from childhood to young adulthood: The mediating and suppressing role of externalizing. *Journal of Abnormal Psychology, 127(4)*, 394.

Gaughan, M. (2002). The substitution hypothesis: The impact of premarital liaisons and human capital on marital timing. *Journal of Marriage and Family, 64(2),* 1741–3737.

Gazelle, H., & Ladd, G. W. (2003). Anxious solitude and peer exclusion: A diathesis–stress model of internalizing trajectories in childhood. *Child Development*, *74*(1), 257-278.

Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and Psychopathology*, *16*(2), 313-333.

Gottlieb, G. (1991). Experiential canalization of behavioral development: Theory. *Developmental Psychology*, *27*(1), 4.

Gottlieb, G. (1997). Synthesizing nature-nurture: Prenatal roots of instinctive behavior. Mahwah, NJ: Lawrence Erlbaum.

Grella, C. E., Stein, J. A., & Greenwell, L. (2005). Associations among childhood trauma, adolescent problem behaviors, and adverse adult outcomes in substance-abusing women offenders. *Psychology of Addictive Behaviors*, *19*(1), 43.

Grimm, K. J., & Ram, N. (2012). Growth curve modeling from a structural equation modeling perspective. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 411-431). New York, NY, US: Guilford Press.

Grimm, K. J., & Widaman, K. F. (2010). Residual structures in latent growth curve modeling. *Structural Equation Modeling*, *17*(3), 424-442.

Hämäläinen, M., & Pulkkinen, L. (1996). Problem behavior as a precursor of male criminality. *Development and Psychopathology*, *8*(2), 443-455.

Hicks, B. M., Blonigen, D. M., Kramer, M. D., Krueger, R. F., Patrick, C. J., Iacono, W. G., & McGue, M. (2007). Gender differences and developmental change in externalizing disorders from late adolescence to early adulthood: A longitudinal twin study. *Journal of Abnormal Psychology*, *116*(3), 433.

Hinde, R. A. (1992). Developmental psychology in the context of other behavioral sciences. *Developmental Psychology*, *28*(6), 1018.

Hoza, B., Gerdes, A. C., Hinshaw, S. P., Arnold, L. E., Pelham Jr, W. E., Molina, B. S., ... & Odbert, C. (2004). Self-perceptions of competence in children with ADHD and comparison children. *Journal of Consulting and Clinical Psychology*, *72*(3), 382.

Hoza, B., Pelham Jr, W. E., Dobbs, J., Owens, J. S., & Pillow, D. R. (2002). Do boys with attention-deficit/hyperactivity disorder have positive illusory self-concepts? *Journal of Abnormal Psychology*, *111*(2), 268.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Isen, A., & Stevenson, B. (2010). *Women's education and family behavior: Trends in marriage, divorce and fertility* (No. w15725). National Bureau of Economic Research.

Jaffee, S. R., Moffitt, T. E., Caspi, A., Fombonne, E., Poulton, R., & Martin, J. (2002). Differences in early childhood risk factors for juvenile-onset and adult-onset depression. *Archives of General Psychiatry*, *59*(3), 215-222.

Jensen-Doss, A., & Weisz, J. R. (2008). Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics. *Journal of Consulting and Clinical Psychology*, *76*(5), 711.

Jester, J. M., Buu, A., & Zucker, R. A. (2016). Longitudinal phenotypes for alcoholism: Heterogeneity of course, early identifiers, and life course correlates. *Development and Psychopathology*, *28*, 1531-1546.

Jester, J. M., Nigg, J. T., Adams, K., Fitzgerald, H. E., Puttler, L. I., Wong, M. M., & Zucker, R. A. (2005). Inattention/hyperactivity and aggression from early childhood to adolescence: Heterogeneity of trajectories and differential influence of family environment characteristics. *Development and Psychopathology*, *17*(1), 99-125.

Jester, J. M., Nigg, J. T., Buu, A., Puttler, L. I., Glass, J. M., Heitzeg, M. M., ... & Zucker, R. A. (2008). Trajectories of childhood aggression and inattention/hyperactivity: differential effects on substance abuse in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(10), 1158-1165.

Keenan-Miller, D., Hammen, C., & Brennan, P. (2007). Adolescent psychosocial risk factors for severe intimate partner violence in young adulthood. *Journal of Consulting and Clinical Psychology*, *75*(3), 456.

Keiley, M. K., Lofthouse, N., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2003). 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*, *31*(3), 267-283.

Kerr, D. C., Lunkenheimer, E. S., & Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and Psychiatry*, *48*(10), 967-975.

Khan, A. A., Jacobson, K. C., Gardner, C. O., Prescott, C. A., & Kendler, K. S. (2005). Personality and comorbidity of common psychiatric disorders. *The British Journal of Psychiatry*, *186*(3), 190-196.

King, K. M., Meehan, B. T., Trim, R. S., & Chassin, L. (2006). Marker or mediator? The effects of adolescent substance use on young adult educational attainment. *Addiction*, *101*(12), 1730-1740.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd edition). Guilford publications.

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin*, *136*(5), 768.

Kovacs, M. (1985). The children's depression inventory (CDI). *Psychopharmacology Bulletin*, *21*, 995-998.

Kovacs, M., & Devlin, B. (1998). Internalizing disorders in childhood. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(1), 47-63.

Kroes, G., Veerman, J. W., & De Bruyn, E. E. (2003). Bias in parental reports? Maternal psychopathology and the reporting of problem behavior in clinic-referred children. *European Journal of Psychological Assessment*, *19*(3), 195.

Leve, L. D., & Fagot, B. I. (1997). Gender-role socialization and discipline processes in one-and two-parent families. *Sex Roles*, *36*(1-2), 1-21.

Leve, L. D., Kim, H. K., & Pears, K. C. (2005). Childhood temperament and family environment as predictors of internalizing and externalizing trajectories from ages 5 to 17. *Journal of Abnormal Child Psychology*, *33*(5), 505-520.

Lindsey, E. W., Mize, J., & Pettit, G. S. (1997). Differential play patterns of mothers and fathers of sons and daughters: Implications for children's gender role development. *Sex Roles*, *37*(9-10), 643-661.

Loukas, A., Zucker, R. A., Fitzgerald, H. E., & Krull, J. L. (2003). Developmental trajectories of disruptive behavior problems among sons of alcoholics: Effects of parent psychopathology, family conflict, and child undercontrol. *Journal of Abnormal Psychology*, *112*(1), 119.

Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*, *116*(1), 155.

Martel, M. M., Pierce, L., Nigg, J. T., Jester, J. M., Adams, K., Puttler, L. I., ... & Zucker, R. A. (2009). Temperament pathways to childhood disruptive behavior and adolescent substance abuse: Testing a cascade model. *Journal of Abnormal Child Psychology*, *37*(3), 363.

McLeod, J. D., & Kaiser, K. (2004). Childhood emotional and behavioral problems and educational attainment. *American Sociological Review*, *69*(5), 636-658.

Mesman, J., & Koot, H. M. (2000). Child-reported depression and anxiety in preadolescence: I. Associations with parent-and teacher-reported problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(11), 1371-1378.

Miech, R. A., Caspi, A., Moffitt, T. E., Wright, B. R. E., & Silva, P. A. (1999). Low socioeconomic status and mental disorders: a longitudinal study of selection and causation during young adulthood. *American Journal of Sociology*, *104*(4), 1096-1131.

Moffitt, T. E. (1993). The neuropsychology of conduct disorder. *Development and Psychopathology*, *5*(1-2), 135-151.

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology*, *14*(1), 179-207.

Muthén, L.K. and Muthén, B.O. (1998-2018). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Ohrt, T., Sjödin, I., & Thorell, L. H. (1999). Cognitive distortions in panic disorder and major depression: Specificity for depressed mood. *Nordic Journal of Psychiatry*, *53*(6), 459-464.

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical Child and Family Psychology Review, 10(4),* 335-351.

Raley, R. K., Crissey, S. R., & Muller, C. (2007). Of sex and romance: Later adolescent relationships and young adult union formation. *Journal of Marriage and Family, 69(5),* 1210–1226.

Reitz, E., Deković, M., & Meijer, A. M. (2005). The structure and stability of externalizing and internalizing problem behavior during early adolescence. *Journal of Youth and Adolescence, 34(6),* 577-588.

Rescorla, L. A., Bochicchio, L., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., ... & Fombonne, E. (2014). Parent–teacher agreement on children's problems in 21 societies. *Journal of Clinical Child & Adolescent Psychology*, *43*(4), 627-642.

Rescorla, L. A., Ginzburg, S., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., ... & Döpfner, M. (2013). Cross-informant agreement between parent-reported and adolescent self-reported problems in 25 societies. *Journal of Clinical Child & Adolescent Psychology*, *42*(2), 262-273.

Rhule-Louie, D. M., & McMahon, R. J. (2007). Problem behavior and romantic relationships: Assortative mating, behavior contagion, and desistance. *Clinical Child and Family Psychology Review*, *10*(1), 53-100.

Richters, J. E. (1992). Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychological Bulletin*, *112*(3), 485.

Richters, J., & Pellegrini, D. (1989). Depressed mothers' judgments about their children: An examination of the depression-distortion hypothesis. *Child Development*, 1068-1075.

Rohde, P., Lewinsohn, P. M., Seeley, J. R., Klein, D. N., Andrews, J. A., & Small, J. W. (2007). Psychosocial functioning of adults who experienced substance use disorders as adolescents. *Psychology of Addictive Behaviors*, *21*(2), 155.

Rusting, C. L. (1999). Interactive effects of personality and mood on emotion-congruent memory and judgment. *Journal of Personality and Social Psychology*, *77*(5), 1073.

Sameroff, A. J. (1983). Developmental systems: Contexts and evolution. In P. Mussen (Ed.), *Handbook of child psychology* (Vol. 1, pp. 237–294). New York, NY: Wiley.

Saulsman, L. M., & Page, A. C. (2004). The five-factor model and personality disorder empirical literature: A meta-analytic review. *Clinical Psychology Review*, *23*(8), 1055-1085.

Singer, J. D., & Willet, J. B. (2003). Applied Longitudinal Data Analysis. New York, NY.

Sivo, S., Fan, X., & Witta, L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation Modeling*, *12*(2), 215-231.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, 15-28.

Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent–child relationships? *Journal of Abnormal Child Psychology*, *31*(1), 13-27.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort difference on the children's depression inventory: A meta-analysis. *Journal of Abnormal Psychology, 111(4),* 578.

U.S. Census Bureau. (2011). Table MS-2. Estimated median age at first marriage, by sex: 1890 to the Present.

Uecker, J., & Stokes, C. (2008). Early marriage in the United States. *Journal of Marriage and Family, 70(4),* 835–846.

van Os, J., Jones, P., Lewis, G., Wadsworth, M., & Murray, R. (1997). Developmental precursors of affective illness in a general population birth cohort. *Archives of General Psychiatry*, *54*(7), 625-631.

Vander Stoep, A., Weiss, N. S., McKnight, B., Beresford, S. A., & Cohen, P. (2002). Which measure of adolescent psychiatric disorder—diagnosis, number of symptoms, or adaptive functioning—best predicts adverse young adult outcomes? *Journal of Epidemiology & Community Health*, *56*(1), 56-65.

Verhulst, F. C., & van der Ende, J. (1991). Assessment of child psychopathology: relationships between different methods, different informants and clinical judgment of severity. *Acta Psychiatrica Scandinavica*, *84*(2), 155-159.

Verhulst, F. C., & van der Ende, J. (1992). Six-year developmental course of internalizing and externalizing problem behaviors. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(5), 924-931.

Watson, D., & Clark, L. A. (1984). Negative affectivity: the disposition to experience aversive emotional states. *Psychological Bulletin*, *96*(3), 465.

Watson, D., & Clark, L. A. (1992). On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of Personality*, *60*(2), 441-476.

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children: Manual*. Psychological Corporation.

Woodward, L. J., & Fergusson, D. M. (1999). Early conduct problems and later risk of teenage pregnancy in girls. *Development and Psychopathology*, *11*(1), 127-141.

Woodward, L. J., Fergusson, D. M., & Horwood, L. J. (2002). Romantic relationships of young people with childhood and adolescent onset antisocial behavior problems. *Journal of Abnormal Child Psychology*, *30*(3), 231-243.

Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2003). Who are the comorbid adolescents? Agreement between psychiatric diagnosis, youth, parent, and teacher report. *Journal of Abnormal Child Psychology*, *31*(3), 231-245.

Youngstrom, E., Wolpaw, J. M., Kogos, J. L., Schoff, K., Ackerman, B., & Izard, C. (2000). Interpersonal problem solving in preschool and first grade: Developmental change and ecological validity. *Journal of Clinical Child Psychology*, *29*(4), 589-602.

Zahner, G. E., & Daskalakis, C. (1998). Modelling sources of informant variance in parent and teacher ratings of child psychopathology. *International Journal of Methods in Psychiatric Research*, *7*(1), 3-16.

Zucker, R. A., & Noll, R. B. (1980). The antisocial behavior checklist. *Unpublished manuscript, Michigan State University*.

Zucker, R. A., Ellis, D. A., Bingham, C. R., & Fitzgerald, H. (1996). The development of alcoholic subtypes: Risk variation among alcoholic families during the early childhood years. *Alcohol, Health, and Research World, 20*, 46-54.

Zucker, R. A., Fitzgerald, H. E., & Noll, R. B. (1990). Drinking and drug history. *Unpublished manuscript, Michigan State University*.

Zucker, R. A., Jester, J.M., Puttler, L. I., Buu, A., & Shedden, K. (2007). Very early externalizing behavior and elevated substance abuse in early adulthood: What are the risk pathways? *Alcoholism: Clinical & Experimental Research, 31(6), 284.*