Effects of the Family Check-Up on Reducing Growth in Conduct Problems from Toddlerhood through School Age: An Analysis of Moderated Mediation

Elizabeth C. Shelleby
Department of Psychology
Northern Illinois University
Psychology-Computer Science Building, Rm. 316
DeKalb, IL 60115
eshelleby@niu.edu

Daniel S. Shaw
Department of Psychology
University of Pittsburgh
210 S. Bouquet Street
4101 Sennott Square
Pittsburgh, PA 15260-0001

Thomas J. Dishion
Department of Psychology
Director, REACH Institute
Arizona State University
900 S. McAllister Rd.
Tempe, Arizona, 85287-1104

Melvin N. Wilson
Department of Psychology
University of Virginia
Gilmer Hall 306
Charlottesville, VA 22908

Frances Gardner
Department of Social Policy and Intervention
University of Oxford
Barnett House, 32 Wellington Square
Oxford OX1 2ER, UK

Abstract. Objective: The Family Check-Up (FCU) is a preventive intervention found to significantly reduce child conduct problems (CP). This study examined the extent to which parents reported that their child's CP was a problem for them at baseline (baseline CP) as a moderator of FCU effects into middle childhood and moderated mediation models to explore positive parent-child dyadic interaction and maternal depressive symptoms as mediators. Method: Participants included 731 mother-child dyads followed from child ages 2 to 9.5 (49% female; 28% African American, 50% European American, 13% biracial, and 9% other; 13% self-reported as Hispanic), half assigned to the FCU. Maternal depressive symptoms, observed parent and child behavior (positive dyadic interaction), and CP were assessed annually. Results: Support was found for baseline CP as a moderator of the FCU, with significant decreases in CP for children in the FCU demonstrating high baseline CP. The following associations did not differ between those with high versus low baseline CP. The FCU significantly increased positive dyadic interaction. Lower maternal depressive symptoms were associated with significantly lower CP. Conclusions: Findings add to existing evidence that preventive interventions are effective for high risk families, and the FCU is especially beneficial for children whose parents report high levels of CP in early childhood. Further, the FCU significantly improved positive dyadic interaction for families of children with both high and low baseline CP.

Public Health Significance: Results of this study suggest that the FCU intervention is effective in reducing child CP through the school age years, particularly for children whose parents report high baseline problems.

Key Words: Preventive intervention, moderation, conduct problems, parenting

Conduct problems (CP), which encompass behaviors such as disruptiveness, aggression, oppositionality, and rule-breaking behavior (American Psychiatric Association, 2013), are particularly common in young children, with estimates of between 5-13% of preschoolers showing moderate to severe CP (Lavigne et al., 1996). Longitudinal studies have shown that young children with early elevated CP are at risk for persisting in antisocial behavior across development (Hong, Tillman & Luby, 2015; Moffit & Caspi, 2001). Based on the rationale that behavior may be more malleable in early childhood and that early intervention may more effectively deescalate problems before they worsen (e.g., Dishion & Patterson, 1992), a wealth of research has focused on testing developmental models that would guide early prevention of CP.

Parenting is often targeted in interventions and has been posited to play a mediating role between various other risk factors (e.g., SES) and child CP (Dawson-McClure et al., 2015; Patterson, Reid, & Dishion, 1992). The Family Check-Up (FCU) is one preventive intervention that combines aspects of motivational interviewing with parent training for at-risk youth. The early childhood version of the FCU has shown significant effects in reducing growth in child CP, with effect sizes from .57 (teacher reports at age 7.5) to .85 (parent reports from ages 2 to 5) for families who attended annually repeated feedback follow-ups (Dishion, Brennan, McEachern, Shaw, Wilson, & Weaver, 2014). At a broader level, meta-analyses of parenting interventions for child CP have shown small to moderate effects (e.g., Piquero et al., 2016); however, studies have found that a sizeable group (e.g., 1/5 to 1/3 of children) fail to improve (e.g., Webster-Stratton, 1990). Prevention studies have shown similar differential effects. For example, *ParentCorps*, a family-centered, universal preventive intervention, found greater reduction in CP for boys with high baseline behavioral dysregulation (23% of the sample) (Dawson-McClure et al., 2015).

Moderator analyses offer an opportunity to identify subgroups to elucidate if

developmental or intervention processes differ for different groups or risks (Hinshaw, 2002; Kraemer et al., 2002). Several studies and a limited number of reviews have investigated whether sociodemographic, child, and family factors operate as moderators of family-based intervention effects (Shelleby & Shaw, 2014). Results are mixed, with many studies finding that individuals facing stressors such as lower income and higher life stress, benefit as much from interventions as those at lower risk (e.g., Lavigne et al., 2008; McGilloway et al., 2012; McTaggart & Sanders, 2007). Others find poor intervention outcomes for those at high risk in domains such as maternal depressive symptoms (Baydar, Reid, & Webster Stratton, 2003). Still others show greater effects for those at higher risk, such as lower mother-child relationship quality (Tein et al., 2004). It is important to explore differential effects by risks because children who demonstrate CP are heterogeneous (Stadler, Poustka, & Sterzer, 2010). Families facing multiple risks such as extreme aggression, impoverished neighborhoods, less optimal parenting, and parent depression (e.g., Hong et al., 2015) are more likely to have high CP persist (Moffitt & Caspi, 2001).

The Family Check-Up (FCU) Parenting Intervention

The Family Check-Up (FCU) is a brief, ecological, individually-tailored preventive intervention based on motivational interviewing (MI) techniques that has been implemented with families from the toddler period through adolescence (Dishion et al., 2008). The core of the intervention addresses family management issues, including parenting skills within three broad domains: positive behavior support, monitoring and limit setting, and relationship quality. The Everyday Parenting curriculum is used (Dishion, et al., 2011), based on a well-established social learning parent management training intervention (Forgatch & Patterson, 2010).

Initially, in a growth mixture analysis that identified high and low problem behavior groups, Dishion et al. (2008) found the effect of FCU on growth in problem behavior to age 4

was strongest in the high problem group. Some studies using the same sample have explored individual factors as moderators of outcome. After exploring several demographic and family adjustment factors, Gardner et al. (2009) found that children of partnered parents and children of parents with lower education demonstrated greater benefit than those of single parents and whose parents had greater education, respectively. Other variables explored –relationship quality, daily hassles, maternal depressive symptoms, substance use, teen parent - did not moderate effects. A recent study on neighborhood deprivation found level of deprivation moderated the impact of the FCU on teacher ratings of CP at age 9.5, with the FCU having greater effects on teacher reported child CP at age 9.5 in neighborhoods with moderately versus extremely high deprivation; however, indirect effects of the FCU were present even for children living in extreme deprivation when positive parenting was improved between ages 2-3 (Shaw et al., 2016). Finally, Pelham et al. (2017) explored effects across different latent classes and found that the effect of the FCU was significantly greater in families characterized by child neglect, legal problems, and parental mental health issues. The current study seeks to extend these findings by examining if the child's initial level of problem behavior continues to moderate child outcomes of the FCU and explores maternal depressive symptoms and positive dyadic interaction as mediators.

Initial Level of Child Conduct Problems

It is essential to know if interventions are effective for children who may be in most need of services and who are at risk for future CP and more serious antisocial behavior in middle childhood and adolescence. Although only a limited number of studies on effectiveness of parenting interventions have employed formal tests of moderation, a recent review of existing studies (Shelleby & Shaw, 2014) found support for moderation by baseline or initial level of child CP. In samples in which not all children were clinically elevated on initial levels of CP, the

more problematic a child's behavior was rated at baseline, the greater change in problem behavior occurred (e.g., Chamberlain et al., 2008; Tein et al., 2004). In the current study, we focus on parental reports of how problematic the child's CP is as our baseline CP measure (Robinson, Eyberg, & Ross, 1980). Although most studies exploring moderation have examined intensity of CP, some have investigated similar constructs, such as clinician-rated level of impairment as a moderator (Lavigne et al., 2008). Others have used the degree to which the behavior represents a problem for the parent as the CP outcome, the same construct explored in the current study (e.g., Gardner et al., 2010; Dishion et al., 2008).

In addition to understanding moderators of direct effects, it is also important to examine how moderating variables may influence mediational pathways. Moderated mediation analysis assesses whether mediating mechanisms are differentially influenced by risk variables and can elucidate if a mediated effect is larger or only apparent within a subgroup (Hayes, 2013; Preacher, Rucker, & Hayes, 2007). A limited number of studies on parenting interventions have demonstrated significant moderation of mediational mechanisms. For example, Chamberlain et al. (2008) found that the indirect effect of an intervention on child CP through increases in parent positive reinforcement was stronger for children with higher baseline CP. Similarly, Zhou et al. (2008) found that for higher risk children (risk including higher CP, parent conflict, negative life events, maternal distress, reduced father contact, and financial hardship), the mediational effects of a multicomponent intervention on child externalizing, internalizing, and mental disorder symptoms were mediated through changes in parent-child relationship quality. For the low risk group, neither the direct intervention nor mediation effects were significant.

The two mediators explored in the current study are positive dyadic interaction and maternal depressive symptoms. Some of the most prominent parenting interventions have their

origins in behaviorism and are designed to shift social contingencies so that parents positively reinforce child prosocial behavior and ignore/punish aversive and disruptive behavior (Serketich & Dumas, 1996). The Family Check-Up is a parent-focused intervention designed to strengthen parents' use of positive behavior support strategies and improve parent-child interactions (Dishion et al., 2008). As such, for a parenting-focused intervention such as the FCU that works directly with parents, a change in child behavior is expected to be linked to intervention induced changes in parenting and improvements in interactions between parents and children.

Targeting parent-child interaction through the FCU has considerable face validity and empirical support (Shaw et al., 2009), but research has elucidated other important mechanisms (e.g., maternal depressive symptoms) that may also explain how preventive interventions reduce child CP. Mothers with higher depressive symptoms have been found to demonstrate higher negativity, coercion, punitive discipline, and inconsistency, and to be less responsive and affectionate toward children than non-depressed mothers (e.g., Lovejoy et al., 2000). With regard to the FCU, the initial assessment covers a broad array of factors that may contribute to child CP, such as economic disadvantage, neighborhood risk, parental conflict, and parental depressive symptoms (Shaw et al., 2009). As noted by Shaw et al. (2009), it seems reasonable to expect that providing broad assistance and support to parents facing multiple challenges might also lead to reductions in maternal depressive symptoms, which, in turn, might be linked to lower child CP.

Previous research on the FCU within the current sample has demonstrated its effects in increasing positive parenting and positive engagement (Dishion et al., 2008), decreasing coercive engagement in early childhood (Sitnick et al., 2015), increasing positive parent-child dyadic interaction (Dishion et al., 2017; Chang et al., 2016), and reducing both the emergence of early problem behavior (Dishion et al., 2008; Shaw et al., 2009) and the growth in conduct problems

through the school-age years (Dishion et al., 2014). As noted above, the two family-based mediating pathways through which the FCU has been found to reduce child CP in early childhood are by improving parental interactions with children (positive parenting) (Dishion et al., 2008) and improving maternal depressive symptoms (Shaw et al., 2009). Dishion et al. (2008) demonstrated that the FCU was effective in increasing primary caregivers' level of positive behavior support, which is a cluster of positive parenting techniques found to decrease the likelihood of coercive escalations, or a child's attempt to avoid parental requests by escalating their negative behavior (Patterson, 1982). More recently, Dishion et al. (2017) demonstrated that the FCU improved parent-child dynamics and that children from families with less adaptive parent-child dynamics demonstrated more chronic behavior problems. The current study seeks to investigate whether the indirect effect through which the FCU is linked to lower CP through improved parent-child interaction could be stronger for children with higher baseline CP. Shaw et al. (2009) showed that decreases in maternal depressive symptoms brought about by the FCU mediated the link between FCU and reduced CP, after accounting for changes in parenting. As research has shown bidirectional effects between maternal depressive symptoms and CP during early childhood (Gross et al., 2008), parents with elevated depression who have children with higher baseline CP may see engagement in the FCU as a way to learn management strategies to improve functioning. Thus, the indirect effect through which the FCU is linked to lower CP through reduced maternal depressive symptoms could be stronger for children with higher baseline CP. To our knowledge, no studies have examined moderated mediation with maternal depressive symptoms as the mechanism of change.

The current study seeks to advance our understanding of how baseline ratings of child CP (parent ratings of how problematic a child's behavior is for them) may influence the efficacy of

the FCU. Analyses build upon prior results, which demonstrated significant effects of the FCU on parent reports through age 7.5 (see Dishion et al., 2014) and teacher reports through age 9.5, by examining additional waves of parent-report data through child age 9.5, and exploring if mediators found to influence CP at earlier ages remain important mechanisms in decreasing CP, while considering baseline CP as a moderator. First, we explored whether the influence of the FCU on reducing CP was maintained to age 9.5 based on parent reports, a replication and extension of previous findings within this trial (Shaw et al., 2016). Second, we examined if baseline CP moderated the effect of the FCU on CP, such that children with higher baseline CP would show greater reductions in CP from the FCU across ages 2 to 9.5. Third, we investigated if baseline CP moderated the mediational path through which the FCU is linked with higher positive dyadic interaction. It was hypothesized that higher baseline CP would be associated with stronger links between FCU and positive dyadic interaction from child ages 3 to 5. Fourth, we investigated whether baseline CP would moderate the mediational path through which the FCU is linked with reductions in maternal depressive symptoms, such that high baseline CP would be associated with stronger links between FCU and reduced depression from ages 3 to 5.

Method

Participants in this study are part of the Early Steps Multisite Study (ESM) designed to examine the effectiveness of a tailored, family-based intervention for children identified as at risk for CP on the basis of child, family, and sociodemographic factors (Dishion et al., 2008). Participants included 731 primary caregiver-child dyads recruited from Women, Infants, and Children (WIC) Nutritional Supplement Centers between 2002 and 2003. Because the vast majority of primary caregivers were mothers (e.g., > 90% across all ages), the terms “mothers” and “maternal” will be used throughout when referring to primary caregivers. Families with

children age 2 years to 2 years 11 months were asked to participate and screened to ensure that they met risk criteria defined as one standard deviation (*SD*) or more above normative averages on at least two of the following three domains: (a) child behavior (CP, high conflict relationships with adults), (b) family problems (maternal depressive symptoms, parenting challenges, substance use, teen parents), and (c) socio-demographic risk (low educational achievement and low family income using WIC criteria). Participants were from Pittsburgh, PA (37%), Charlottesville, VA (26%), and Eugene, OR (37%). After the initial age 2 assessment, participants were randomly assigned to the intervention (n = 367) or control group. (n = 364). Reports of primary caregivers' race/ethnicity was 28% African American, 50% European American, 13% biracial, and 9% other, and 13% self-reported as Hispanic. Over two-thirds of families had annual income of less than $20,000 at recruitment (in 2002-3). Observational and parent report data were collected at home visits when the target child was 2, 3, 4, 5, 7.5, 8.5, and 9.5 years old. Parents who agreed to participate in the study and went through informed consent procedures were scheduled for annual home visits. All procedures performed were in accordance with established ethical guidelines, and appropriate approval was obtained from institutional review boards all sites. Assessment visits were identical for control and intervention participants and involved approximately 2.5-3 hours of structured and unstructured play activities for the target child with mothers and alternative caregivers and siblings. Research staff were blind to intervention assignment. After an initial free play (15 min), each dyad participated in the following tasks: cleanup (5 min), delay of gratification (5 min), four teaching tasks (3 min each), second free play (4 min), second cleanup (4 min), presentation of two inhibition inducing toys (2 min each), and meal preparation (20 min). Protocols comparable in content, structure, and length, occurred at all ages with slight variations to make tasks developmentally appropriate. Mothers

were compensated for each assessment, and children began to receive compensation at age 7.5.

Of the 731 families who initially participated, 659 (90%) participated at age 3, 629 (86%) at age 4, 621 (85%) at age 5, 566 (77%) at age 7.5, 565 (77%) at age 8.5, and 587 (80%) at age 9.5. At ages 3, 4, 5, and 7.5, analyses to explore whether likelihood of participant drop-out systematically varied revealed no significant differences related to intervention group; project site; children's race/ethnicity, or gender; maternal depressive symptoms; or parent-reported child behavior. However, at age 8.5, families in the FCU ($p < .05$), families with lower age 2 maternal depressive symptoms ($p < .05$), and families with higher age 2 ratings of externalizing problem behavior were less likely to participate ($p < .05$), and there were significant differences based on site, with families from Eugene more likely to be retained than families from Charlottesville ($p < .05$) and Pittsburgh (at a trend level; $p < .10$). There continued to be no significant differences at age 8.5 related to race/ethnicity, or gender. At age 9.5, only the following selective attrition effect remained, reflecting the 'come and go' behavior of participants: those with higher age 2 internalizing and externalizing problems were less likely to participate ($p < .05$).

Intervention Protocol: The Family Check-Up (FCU)

The FCU is a brief intervention based on MI techniques and modeled after the Drinker's Check-Up (Miller & Rollnick, 2002). Following developmentally-based assessments, FCU families were scheduled to meet with a parent consultant for an initial contact meeting and a feedback session, and follow-up treatment sessions typically focused on parenting. Age 2 assessments were completed before random assignment results were known to either the research staff or the family. Intervention families received a $25 gift certificate for completing the FCU and feedback. After the home assessment, the second visit called the "get to know you" meeting consisted of the parent consultant meeting with the caregiver(s) and discussing their concerns

with a focus on current family issues that were most critical to their child's and family's functioning. For the third meeting, the feedback, parent consultants utilized MI to summarize the assessment results and highlight strengths and areas in need of attention. Feedback objectives involve assessing the caregiver's willingness to change problematic parenting practices, identifying ways to support strengths, and providing services appropriate to the family's needs. Parents have the option to participate in follow-up sessions focused on parenting and contextual issues (see Dishion et al., 2014 for additional information). Fidelity of implementation has been shown to be acceptable (see Smith et al., 2013). Of the 367 FCU families, 343 (93.5 %) engaged in the FCU at least once between ages 2 and 9.5 (Chang et al., 2016). The following are the proportion who had annual assessments and (a) feedback, (b) participated in follow-ups, and (c) in parentheses, average number of follow-ups: age 2: 76 %, 72% (3.4); age 3: 69 %, 70% (3.1); age 4: 70 %, 74 % (3.5); age 5: 66 %, 68% (3.6); age 7.5: 64%, 65% (2.4); age 8.5: 71 %, 72 % (2.6); An intention-to-treat design was used. See Figure 1 for a participant flow chart and see Appendix for additional information about the data.

Measures

Demographics questionnaire. Several variables will be used as covariates, including age 2 family income (M = \$19,938, SD=\$12,242), and dummy coded race/ethnicity, gender, and site.

Childhood behavior problems. The Child Behavior Checklist (CBCL) for ages 1.5 - 5 (Achenbach & Rescorla, 2000) and ages 6 -18 (Achenbach & Rescorla, 2001) is a questionnaire that assesses behavior problems in children and was administered to mothers at each home assessment. Data from the age 2, 3, 4, 5, 7.5, 8.5, and 9.5 assessments will be used to explore growth in CP as an outcome. To generate a factor of CP behavior that was both developmentally meaningful and clinically relevant, individual items from the CBCL were chosen that mapped

onto *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013) criteria for Oppositional Defiant Disorder and aggressive items from Conduct Disorder. Only the 8 items that were continuously present on both versions of the CBCL across ages were included. This oppositional/aggressive (Opp/Agg) factor will serve as the primary outcome measure of CP. Composite variables were computed by summing values for these items at each assessment, and alpha values were between .71 and .84 at all ages. The following items are included in the composite: cruel to animals, destroys own things, destroys others' things, gets in many fights, physically attacks people, is defiant, is disobedient, and has temper tantrums. Means and SDs (in parentheses) at each age were as follows: age 2 M = 5.21 (2.68); age 3 M = 4.42 (2.78); age 4 M = 3.94 (2.82); age 5 M = 3.61 (2.86); age 7.5 M = 3.83 (3.13); age 8.5 M = 3.11 (2.99); age 9.5 M = 2.99 (2.85). The 36-item Eyberg Child Behavior Inventory (ECBI, Robinson et al., 1980) was also administered to mothers. Data from the ECBI Problem factor at age 2 served as the measure of initial level of child CP (α = .85). The Problem factor, which asks parents whether each behavior is a problem for them (i.e., "yes" or "no"), was selected rather than the Intensity factor due to a specific interest in understanding how baseline parental ratings of how problematic a child's behavior is might moderate intervention effectiveness. In addition, because the Intensity factor is similar in content and structure to the CBCL scales, the Problem factor was selected such that the measures of child behavior problems utilized as a moderator and as the main outcome variable were significantly related yet distinct constructs.

Positive dyadic interaction. At child ages 2, 3, 4, and 5, child and caregiver behaviors during interactive tasks at the home assessment were coded using the Relationship Affect Coding System (RACS; Peterson et al. 2008), a micro-social coding system of relational behavior and affect within dyadic interaction (see also Chang et al., 2016; Dishion et al., 2017). At any given

moment during an interaction, the child and caregiver were given one individual code from each of the three dimensions of behavior: a) verbal: general conversation (positive, negative, or neutral) and verbal attempts at behavior change of the partner (directives, negative directives, and positive structure); b) physical: physical interaction (positive physical contact, negative physical contact, and neutral physical contact); c) affect: general emotional display (or lack of) in an interaction (positive affect, validation, anger/disgust, distress, ignore). Six behavior clusters were created to summarize the three data streams for each person in the interaction: positive, neutral, directive, negative, no talk, and ignore. For example, the positive cluster included behaviors such as positive verbal statement, positive physical contact, positive structure, or positive affect, as well as validation. The negative cluster included anger and disgust, and negative verbal and physical interaction (also see Sitnick et al. 2015). When the three codes did not converge, priority was given to codes that captured behavior or affect more salient to the relationship. Thus, the trumping was: ignore, negative, positive, directive, no talk, neutral.

Once behavior clusters for each dyadic member were coded, we derived data on dyadic states by combining the behavior cluster of the child and that of the caregiver at each time point (see Dishion et al., 2017). The total duration that each dyad was observed in a particular dyadic state was calculated as a proportion score in relation to the overall session time. The dyadic states that were formulated were: positive engagement, neutral engagement, coercive engagement, and non-interactive. Of those, we were primarily interested in dyadic states that involved both participants mutually engaged in particular sets of interactive behavior. Specifically, a summary score of dyadic positive engagement reflected the duration of time that the child or caregiver engaged in positive or neutral behavior while the other member of the dyad also displayed positive or neutral engagement. Likewise, a summary score of dyadic coercive

engagement encompassed interactions in which either participant was negatively engaged or directive, and the partner responding by negative engagement, directive, ignoring, or not talking. The coding was recorded using Noldus Observer XT, Version 11.0, which enables continuous coding of an interaction of caregivers and children simultaneously. Inter-rater reliability was based on 20 % of the sample: overall $\kappa = .93$ at each age; agreement =93 %, 94 %, 93 %, and 94% at ages 2, 3, 4, and 5, respectively. Coders were trained undergraduates, graduate students, and project staff. Throughout the duration of the project, inter-rater reliability checks were conducted. Between 15-25% of the entire sample was coded twice to ensure inter-rater reliability. Extensive training was conducted to ensure that coding was culturally sensitive.

Scores of dyadic positive and coercive engagement were used to construct a latent factor of positive dyadic interaction at the ages of 2, 3, 4 and 5 years. Means and SDs (in parentheses) at each age were as follows. Positive engagement: age 2 $M = 0.33$ (0.14); age 3 $M = 0.35$ (0.14); age 4 $M = 0.27$ (0.13); age 5 $M = 0.37$ (0.14). Coercive engagement: age 2 $M = 0.11$ (0.08); age 3 $M = 0.09$ (0.06); age 4 $M = 0.09$ (0.05); age 5 $M = 0.06$ (0.05). This construct has been used in previous studies with the same sample (see Chang et al., 2016; Dishion et al., 2017). As reported in Chang et al. (2016), dyadic positive engagement and dyadic coercive engagement scores were significantly correlated within each time (rs ranged from −.61 to −0.47, all ps < 0.01) and also across the four waves (−.61 to −0.22, all ps < 0.01). A second order latent factor of positive dyadic interaction was created based on latent factors of positive dyadic interaction at 3, 4, and 5 years. The latent factor of age-2 positive dyadic interaction was kept separate, as it captured variability in relational qualities that existed prior to random assignment to the FCU.

Maternal depressive symptoms. The CES-D (Radloff, 1977) is a 20-item measure of depressive symptoms that was administered to mothers at each home assessment. Data from the

ages 2, 3, 4, and 5 assessments will be used in analyses. Mothers report how frequently they experienced a list of symptoms during the past week on a scale from 0 (less than a day) to 3 (5–7 days). Sample items include, “I felt depressed” and “I was bothered by things that usually don’t bother me.” Items are summed to create a depressive symptoms score. Alphas across ages 2 to 5 ranged from .74 to .77. Means and SDs (in parentheses) at each age were as follows: age 2 *M* = 16.75 (10.66); age 3 *M* = 15.39 (10.95); age 4 *M* = 14.99 (10.85); age 5 *M* = 14.81 (11.32). Scores of 16 or greater indicate clinically significant symptoms; the following percentages reflect individuals within the clinical range at each age: 45.4% at age 2, 41.6% at age 3, 39.2% at age 4, and 39.1% at age 5. A latent factor of depressive symptoms was created based on scores at child age 3, 4, and 5 years. The age 2 score was kept separate; it captured variability prior to the FCU.

Data Analytic Plan

Structural equation modeling (SEM) was utilized to examine latent factors of positive dyadic interaction and maternal depressive symptoms. Latent growth curve modeling (LGM) was employed to examine growth in child CP as well as the effects of FCU on the growth trajectories. LGM can model both inter- and intra-individual change (Selig & Preacher, 2009). Mplus 7.31 (Muthén & Muthén, 1998-2012), using maximum likelihood estimation with robust standard errors (MLR), was applied for hypothesis testing. MLR is robust to nonnormality and adjusts for missing data by estimating parameters of all available data for the estimation of a specific parameter. Initial level of CP as a moderator was explored through a multi-group design. The sample was split into two risk groups: high and low baseline CP. Multi-group SEM was utilized rather than a single-group SEM with interaction terms because we were interested in examining the distinct direct and mediational processes within the high baseline CP group versus the low group. In addition, multi-group modeling allows more parameters to be compared than is

possible in analyses examining continuous interactions. For example, we can see whether the pathways from parenting and depression to growth in CP differ across high and low baseline CP groups whereas a continuous interaction term would only tell us whether the direct effect of intervention on CP and effect of intervention on parenting/depression are moderated by baseline CP. In addition, potential differences between groups on the ways covariates may operate can be explored. Because children in the sample had scores on baseline CP that were elevated compared to normative samples, the high CP group was defined as having a score one *SD* above the *sample* mean, equating to raw scores > 19 (equivalent *t*-scores above 65) on the ECBI problem factor (n = 158). The low group was defined as having a raw score ≤ 19 (n = 571). In multi-group models, first all parameters were constrained to equality, representing the most parsimonious model. Next, paths were individually freed. Satorra-Bentler scaled chi-square difference tests (Satorra, 2000) were conducted to determine if paths differed and should be freed. This test is appropriate for MLR, as it divides the chi-square by a scaling correction factor. Model fit was assessed using several indices. Unstandardized results are reported.

Results. Hypothesis 1: Intervention Effect across the Full Sample

As a first step, an unconditional growth curve model for Opp/Agg behavior across ages 2-9 was computed. The best fitting model included a quadratic function, which yields three growth factors: an intercept (i.e., initial status; in this case, Opp/Agg at age 2), a linear slope, and a quadratic term, indicating an acceleration or deceleration in growth. For each factor, the mean describes the average initial score (intercept) or change over time (slopes), and the variance specifies whether there is significant variability across individuals in the parameter. The unconditional model showed significant means and variances for the intercept (M = 5.046, p < .01; variance = 4.219, p < .01), linear slope (M = -0.557, p < .01; variance = 0.563, p < .01), and

quadratic slope (M = 0.040, p < .01; variance = 0.008, p < .01). The fit indices were as follows: χ^2 (df =19) = 92.066, p < 0.001; *CFI* = 0.957; *RMSEA* = .073; *SRMR* = .048. The linear and quadratic slopes negatively covaried, $B = -0.06$, p < 0.001, indicating that initially greater declines in CP were associated with slight "acceleration" later as children aged, potentially showing that those who changed more rapidly earlier on had less room for change at later points and even showed a slight increase. Next, the full model including FCU and covariates was computed, which provided good fit: χ^2 (df =51) = 137.849, p < 0.001; *CFI* = 0.958; *RMSEA* = .049; *SRMR* = .027. The FCU was significantly associated with the linear slope, such that those in the FCU showed greater decreases in CP (B = -0.216, p =.028; 95% CI = -0.408 to -0.054), and the quadratic slope (B = 0.025, p = .035; 95% CI = 0.002 to 0.045), demonstrating that the FCU was associated with a slight acceleration. Combining the effects of FCU on the linear and quadratic slopes, the results indicated that over time, the FCU had a convex rather than linear effect on Opp/Agg behaviors. Specifically, children who had high baseline CP benefited from the FCU, such that their Opp/Agg behaviors lowered over time, compared to the children in the control condition; however the magnitude of differences decreased over time.

Supplementary Figure 1 presents plotted means of Opp/Agg over time and means and standard deviations of Opp/Agg at each time by condition and within the high baseline CP group by condition (to be discussed in greater detail below). Although the quadratic term is positive and mean levels in the FCU group increase slightly from age 5 to 7.5 (when there was a 2.5 year gap between visits because of funding constraints), there was a subsequent decrease at ages 8.5 and 9.5. Effect sizes for the difference in Opp/Agg between FCU and control were d = .027 at age 8.5 and d = .078 at age 9.5. To further explore the association between the FCU and linear and quadratic slopes of Opp/Agg, additional analyses were conducted. The quadratic factor

describes the upturn or downturn over time beyond what is predicted by the linear factor. When a quadratic factor is present in a growth model, the linear term describes the rate of change at the intercept. By setting the intercept at different times in different models, the following can be seen of the mean of the linear slope across time: age *2* M = -0.557, $p < .01$; age 3 M = -0.478, $p < .01$; age 4 M = -0.398, $p < .01$; age 5 M = -0.318, $p < .01$; age 7.5 M = -0.199, $p < .01$; age 8.5 M = -0.039, $p = 0.293$; age 9.5 M = 0.041, $p = 0.405$. This demonstrates a significant linear decrease through age 7.5, and then leveling off, with nonsignificant linear terms at the two subsequent time-points (at which the mean linear slope is not significantly different from zero).

Hypothesis 2: Moderation of Direct Intervention Effect on Growth in CP

This model examined whether baseline CP moderated the direct effect of the intervention on growth in CP. Results from the multi-group model provided adequate fit: χ^2 (df =132) = 237.287, $p < 0.001$; *CFI* = 0.948; *RMSEA* = .047; *SRMR* = .041. Based on preliminary tests examining differences between residual variance of the indicators (i.e., measures of Opp/Agg behavior at each age) for the growth model, suggesting they differed, these parameters were allowed to vary freely across groups in this and all subsequent multi-group models. The significant Satorra-Bentler chi-square difference test indicated the association between the FCU and linear slope of Opp/Agg behavior significantly differed between the groups, such that the FCU was associated with decreasing Opp/Agg behavior for the high group only (B = -0.344, p = .001; 95% CI = -0.551 to -0.136), whereas this association was not significant in the low group (B = -0.163, p = .091). Again, the FCU was significantly associated with the quadratic slope (B = 0.025, p = .045; 95% CI = 0.001 to 0.046), demonstrating that the FCU was associated with a slight acceleration. This was not moderated by baseline CP. Within the subgroup with high baseline CP, effect sizes were d = .258 at age 8.5 and d = .462 at age 9.5.

Hypothesis 3: Moderated Mediation through Positive Dyadic Interaction

This model examined whether positive dyadic interaction across ages 3 to 5 mediated the effect of the FCU on Opp/Agg across ages 2 through 9.5 and whether such an effect might be stronger for families in which children demonstrated higher levels of baseline CP. In line with previous research with this sample (Chang et al., 2016), a second-order latent factor was created based on latent factors of positive dyadic interaction at 3, 4, and 5 years. The age-2 latent factor was kept separate, as it captured variability prior to the FCU. The initial single-group model had adequate fit: χ^2 (*df* =170) = 441.980, $p < 0.001$; *CFI* = 0.926; *RMSEA* = .047; *SRMR* = .043. In this model, the FCU was linked to significantly higher positive dyadic interaction (B = 0.022, p = .011; 95% CI = 0.005 to 0.036). There was a non-significant trend-level link between the FCU and both the linear (B = -0.193, p =.052) and quadratic (B = 0.023, p = .062) slopes. A non-significant trend-level link was also found between positive dyadic interaction and the linear slope, (B = -1.025, p = .10). The link between positive dyadic interaction and the quadratic slope was not significant (B = 0.117, p = .15). The indirect effect was not significant.

Next, a multigroup model was explored. Fit was adequate: χ^2 (*df* =406) = 785.321, $p <$ 0.001; *CFI* = 0.90; *RMSEA* = .051; *SRMR* = .062 (see Figure 2). The significant Satorra-Bentler chi-square difference test indicated that the link between the FCU and linear slope of Opp/Agg significantly differed between the groups, such that the FCU was associated with decreasing Opp/Agg for the high group (B = -0.354, p = .001; 95% CI = -0.566 to -0.114) but not the low group (B = -0.127, p = .195). The following associations did not differ between groups: the FCU was linked to significantly higher positive dyadic interaction (B = 0.022, p = .011; 95% CI = 0.005 to 0.039); a nonsignificant trend was found between the FCU and quadratic slope (B = 0.020, p = .087); a nonsignificant trend was found between positive dyadic interaction and linear

growth in CP (B = -1.050, p = .096); the link between positive dyadic interaction and the quadratic term was not significant (B = 0.118, p = 0.14). The indirect effect was not significant.

Hypothesis 4: Moderated Mediation through Maternal Depressive Symptoms

This model examined whether maternal depressive symptoms across ages 3 to 5 mediated the effect of the FCU on Opp/Agg behavior from ages 2 through 9.5 and whether such an effect might be stronger (or only significant) for families of children with higher baseline CP. A latent factor was created based on CES-D scores at child ages 3, 4, and 5 years. The measure of depression at age 2 was kept separate because it captured variability prior to the start of the FCU. The initial single-group model had adequate fit: χ^2 (df =96) = 235.371, p < 0.001; *CFI* = 0.950; *RMSEA* = .045; *SRMR* = .046. In this model, the FCU was associated with significantly lower maternal depressive symptoms (B = -1.231, p = .047; 95% CI = 0.326 to 0.464). A nonsignificant trend between the FCU and both the linear (B = -0.168, p =.094), and quadratic (B = 0.020, p = .10) slopes was found. Significant associations between maternal depressive symptoms and both the linear and quadratic terms of Opp/Agg were found. Lower maternal depressive symptoms were associated lower Opp/Agg; however, over time, this association became less steep (linear term: B = 0.053, p < 0.001; 95% CI = 0.037 to 0.067; quadratic term: B = -0.006, p < 0.001; 95% CI = -0.008 to -0.004). The indirect effect of the FCU on Opp/Agg (linear and quadratic terms) through maternal depressive symptoms was a non-significant trend (linear term: B = -0.065, p = .055; quadratic term: B = 0.007, p = .061).

Next, a multigroup mediation model was explored. Results provided adequate fit: χ^2 (df =244) = 399.361, p < 0.001; *CFI* = 0.94; *RMSEA* = .042; *SRMR* = .052 (see Figure 3). Satorra-Bentler chi-square difference tests indicated that that the association between the FCU and linear slope of Opp/Agg behavior significantly differed between the groups, such that the FCU was

associated with decreasing Opp/Agg behavior for the high group (B = -0.363, p = .001; 95% CI = -0.576 to -0.150), whereas this association was not significant in the low group (B = -0.130, p = .297). The following associations did not differ between the groups. There was a non-significant trend level link between the FCU and lower maternal depressive symptoms (B = -1.114, p = .067). Lower maternal depressive symptoms were associated lower Opp/Agg; however, over time, this association became less steep (linear term: B = 0.051, p < 0.001; 95% CI = 0.035 to 0.066); quadratic term: B = -0.005, p < 0.001; 95% CI = -0.007 to -0.003). The indirect effect of the FCU on Opp/Agg (linear and quadratic terms) through maternal depressive symptoms was a trend (linear term: B = -0.058, p = .073; quadratic term: B = 0.006, p = .079).

Discussion

The goals of the present study were to examine how baseline CP may moderate the effectiveness of the FCU in facilitating faster declines in CP over time (compared to declines in the control group) and to examine whether mediational pathways through which the FCU has been associated with faster declines in child CP were moderated by child CP at age 2. Consistent with prior work (e.g., Dishion et al., 2014), we found a group difference on the slope of Opp/Agg behavior. Initial analyses across the full sample showed that based on primary caregiver reports of child CP, the FCU was associated with reduced growth in parent-reported child CP from ages 2 through 9.5, extending previous findings (Dishion et al., 2008; Dishion et al., 2014; Shaw et al., 2009, 2016). The multi-group direct effect model demonstrated significant moderation by baseline CP, such that the FCU was effective in reducing growth in CP in the high baseline CP group whereas in the low group, this pattern was a non-significant trend.

As previously described, although the pathway from the FCU to the quadratic slope was also significant and mean levels in the FCU and control groups increased slightly from age 5 to

7.5 (when there was a 2.5 year gap between visits due to funding constraints), there was a subsequent decrease at ages 8.5 and 9.5. The association between the FCU and the quadratic term does not suggest that the intervention has an iatrogenic effect. Results of t-tests found that although CP between the intervention and control groups did not significantly differ across the whole sample at ages 8.5 or 9.5, restricting the sample to those with elevated baseline CP demonstrated a significantly lower mean on Opp/Agg for children in the FCU at age 9.5 (with a similar, yet non-significant difference seen at age 8.5, $p = .16$.). Although the quadratic term is significant and shows a slight increase, this does not negate the positive influence the FCU demonstrated. Despite this acceleration, the high baseline CP group in the intervention still demonstrated significantly lower Opp/Agg scores at age 9.5 than high baseline CP controls.

The only hypothesized difference between the high and low baseline CP groups that was supported in the mediational models was the main effect of the FCU on growth in Opp/Agg. However, the FCU was effective in increasing positive dyadic interaction in both high risk and low risk families. Interestingly, neither positive dyadic interaction nor maternal depressive symptoms were significant mediators, suggesting that future studies should investigate other mechanisms, or perhaps sequential mechanisms, that may help to explain the processes through which the FCU continues to reduce CP over time. The current findings add to our understanding of factors that moderate the effectiveness of the FCU. Similar to other previously published findings using the current sample, the effect of the FCU has been shown to differ across various risks, including stronger effects for children of parents with lower education and children of partnered parents (Gardner et al., 2010), for those living at moderate versus extreme levels of neighborhood deprivation (Shaw et al., 2016), and for families characterized by child neglect, legal problems, and parental mental health issues (Pelham et al., 2017). These current findings

demonstrate that parents' perceptions of how problematic their child's behavior is at baseline serves as a meaningful indicator of who may benefit most from the FCU. We hypothesized that the ECBI problem factor may be more closely related to parental *motivation* to than overall frequency of CP, since it taps into a parent's beliefs about whether a specific behavior is a problem for the parent. Future research should examine this question empirically.

The current sample is somewhat novel compared to participants from other parenting intervention studies in four respects, which are related: the FCU is preventive, the FCU occurs annually across several years, it is relatively brief in duration compared to even other brief interventions (e.g., Incredible Years, Parent Management Training Oregon model), and children have been followed longitudinally for more than 7 years. Because the FCU is preventive, many families do not need to reduce their relatively low levels of problem behavior, providing some context as to why those with higher baseline CP may benefit to a greater extent, similar to findings from other preventive interventions showing greater benefit for children with higher initial CP. Much of the other moderator research involves pre/post designs of more standardized or manual-based 12- to 16-week interventions. Other studies may have long term follow-up, but few involve a design similar to that of the current study. One exception is Fast Track, a multi-modal, intensive intervention administered for several years in which children were followed through adolescence. Although differing from the current trial in that main effects of the intervention on child CP were not significant, findings have shown lower symptoms and diagnoses of disorders across 3rd, 6th, 9th and 12th grades, and fewer severe arrests through age 19 for children from this high-risk sample who had the highest baseline problem behavior, (CPPRG, 2007, 2010). It will be interesting for FCU studies to examine similar long-term effects.

The mediational model involving dyadic interaction revealed that the FCU is beneficial in promoting more adaptive and less coercive interactions between parents and children through age 5. Interestingly, the indirect effect of the FCU reductions in Opp/Agg operating through dyadic interaction was not significant in the current study, differing from models examining similar constructs at earlier timepoints (e.g., Dishion et al., 2008). Based on direct effects of the FCU on both positive dyadic interaction and Opp/Agg, further research investigating other explanatory mechanisms that may play an important role in explaining these sequential processes appears warranted. Similarly, the mediational model involving maternal depressive symptoms demonstrated a non-significant trend level association between the FCU and lower maternal depressive symptoms and a non-significant trend level indirect effect of the FCU on Opp/Agg through maternal depressive symptoms. An interesting extension would be to examine differences among mothers with and without clinical depression to see if intervention effects on depression differ depending on level of initial risk. Gardner et al. (2009) found that maternal depressive symptoms did not moderate FCU effects on child CP in the current sample from child ages 2 to 4, but it remains possible that maternal depressive symptoms at child age 2 might be related to the FCU's influence on maternal depressive symptoms themselves or on parenting.

Several limitations should be noted when interpreting results. First, although focusing on a high risk sample is important and could be considered a strength, the sample is not representative of the population at large. In addition, observational measures were used when available, but mothers reported on depressive symptoms and child CP. RCTs do protect against bias to some extent, but additional measures such as observations and peer reports would also be of value. Although the coding manual used for the current study was developed using examples of culturally diverse coding categories, additional research is needed to provide psychometric

support for the validity of this and other methods of observational coding across different cultural groups. In addition, although the current study attempted to provide a meaningful cutoff, there are a multitude of ways to define high versus low child CP, which could influence results. Although analyses with continuous interactions are also illuminating, as noted, multi-group modeling is informative in considering whether there may be distinct mediational processes within different subgroups, a key aim in the current research. Other studies have created similar cut-offs to examine differences (see Chamberlain et al., 2008; Zhou et al., 2008). The high baseline CP group was defined by having a score equivalent to a *t*-score above 65, a more meaningful distinction than a simple median split. Further, multi-group modeling allows more parameters to be compared than is possible in analyses examining continuous interactions. Piecewise growth modeling may be an important future direction to allow for the most developmentally meaningful measures of behavior to be captured at younger versus older ages.

Despite these limitations, current findings suggest that the FCU was effective in reducing child CP across ages 2 to 9.5 for children with higher baseline CP and that the FCU was effective in improving positive dyadic interaction across ages 3-5 for families of children with both high and low baseline CP. In terms of clinical implications, from a cost/benefit analysis perspective, using limited resources to target children with the highest initial levels of CP may be an effective strategy for broader dissemination and implementation. However, targeting children with high scores is one criterion for identifying children at risk and for responding to the FCU; others may also be important. Recent research by Pelham et al. (2017), using the same data set, suggests that high rates of maternal depression, parent use of mental health services, parent criminality, and involvement with child welfare are reliable predictors of better response to the FCU. Additional studies investigating predictors of trajectories of behavior problems and treatment responsiveness

will serve multiple purposes of identifying those who are most likely to persist in problems, identifying those who respond well to existing interventions, and adapting interventions to serve the needs of those who may still be in need but who do not respond optimally.

References

Achenbach, T. M., & Rescorla, L. (2000). *Manual for the ASEBA preschool forms & profiles: An integrated system of multi-informant assessment*. Burlington: University of Vermont.

Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA School-Age Forms and Profiles.* Burlington: University of VT.

American Psychiatric Association. (2013). *DSM-V*. Washington, DC: Author.

Baydar, N., Reid, M. J., & Webster-Stratton, C. (2003). The role of mental health factors and program engagement in the effectiveness of a preventive parenting program for Head Start mothers. *Child Development, 74*(5), 1433-1453.

Chamberlain, P., Price, J., Leve, L. D., Laurent, H., Landsverk, J. A., & Reid, J. B. (2008). Prevention of behavior problems for children in foster care: Outcomes and mediation effects. *Prevention Science, 9*(1), 17-27. 10.1007/s11121-007-0080-7

Chang, H., Shaw, D.S., Shelleby, E.C., Dishion, T.J., & Wilson, M.N. (2016). The long-term effectiveness of the family check-up on peer preference: Parent-child interaction and child effortful control as sequential mediators. *Journal of Abnormal Child Psychology*. doi: 10.1007/s10802-016-0198-9

Conduct Problems Prevention Research Group. (2007). Fast Track randomized controlled trial to prevent externalizing psychiatric disorders: Findings from grades 3 to 9. *J Am Acad Child Adolesc Psychiatry, 46*(10), 1250-1262. doi: 10.1097/chi.0b013e31813e5d39

Conduct Problems Prevention Research Group. (2010). Fast Track Intervention Effects on Youth Arrests and Delinquency. *Journal of Experimental Criminology, 6*(2), 131-157. doi: 10.1007/s11292-010-9091-7

Dawson-McClure, S., Calzada, E., Huang, K., Kamboukos, D., Rhule, D., Kolawole, B….. Brotman, L.M. (2015). A population-level approach to promoting healthy child development and school success in low-income, urban neighborhoods: Impact on parenting and child conduct problems. *Prevention Science, 16*(2), 279-290. doi: 10.1007/s11121-014-0473-3

Dishion T. J., Brennan, L. M., McEachern, A., Shaw, D. S., Wilson, M. N., & Weaver, C. M. (2014). The prevention of early onset problem behavior through periodic family check-ups in early childhood: A health maintenance model with indigent, high risk families. *Journal of Abnormal Child Psychology, 42*, 343-354. doi: 10.1007/s10802-013-9768-2

Dishion, T.J., Mun, C.J., Tein, J.Y., Kim, H., Shaw, D.S., Gardner, F., Wilson, W.N., & Peterson, J. (2017). The validation of macro and micro observations of parent-child dynamics using the Relationship Affect Coding System in early childhood. *Prevention Science, 18*(3), 268-280. doi: 10.1007/s11121-016-0697-5

Dishion, T. J., & Patterson, G. R. (1992). Age effects in parent training outcome. *Behavior Therapy, 23*(4), 719-729. doi: 10.1016/s0005-7894(05)80231-x

Dishion, T. J., Shaw, D. S., Connell, A., Gardner, F., Weaver, C. M., & Wilson, M. N. (2008). The Family Check Up with high risk indigent families: Preventing problem behavior by increasing parents' positive behavior support in early childhood. *Child Development, 79*(5), 1395-1414. doi: 10.1111/j.1467-8624.2008.01195.x

Dishion, T. J., Stormshak, E. A., & Kavanagh, K. (2011). *Everyday parenting: A professional's guide to building family management skills.* Champaign, IL: Research Press.

Forgatch, M. S., & Patterson, G. R. (2010). *Parent Management Training—Oregon Model: An intervention for antisocial behavior in children and adolescents Evidence-based psychotherapies for children and adolescents* (Vol. 2, pp. 159-178). New York: Guilford.

Gardner, F., Connell, A., Trentacosta, C. J., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2009). Moderators of outcome in a brief family-centered intervention for preventing early problem behavior. *J of Consulting and Clinical Psychology, 77*(3), 543-553. doi: 10.1037/a0015622

Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *J Clin Child and Adolesc Psychol, 39*(4), 568–580. doi: 10.1080/15374416.2010.486315

Gross, H. E., Shaw, D. S., & Moilanen, K. L. (2008). Reciprocal associations between boys' externalizing problems and mothers' depressive symptoms. *Journal of Abnormal Child Psychology, 36*(5), 693-709. doi: 10.1007/s10802-008-9224-x

Hayes, A. F. (2013). Introduction to mediation, moderation and conditional process analysis: A regression based approach. New York, NY: Guilford Press.

Hinshaw, S. P. (2002). Prevention/intervention trials and developmental theory: Commentary on the Fast Track special section. *J Abnorm Child Psychol, 30*(1), 53-59. doi: 10.1023/A:1014279015195

Hong, J.S., Tillman, R., & Luby, J.L. (2015) Disruptive behavior in preschool children: Distinguishing normal misbehavior from markers of current and later childhood conduct disorder. *Journal of Pediatrics, 166*(3), 723-730. doi: 10.1016/j.jpeds.2014.11.041

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General*

Psychiatry, 59(10), 877-883. doi: 10.1001/archpsyc.59.10.877

Lavigne, J. V., Gibbons, R. D., Christoffel, K. K., Arend, R., Rosenbaum, D., Binns, H., . . . Isaacs, C. (1996). Prevalence rates and correlates of psychiatric disorders among preschool children. *J American Academy of Child & Adolescent Psychiatry, 35*, 204-214.

Lavigne, J. V., LeBailly, S. A., Gouze, K. R., Cicchetti, C., Jessup, B. W., Arend, R., . . . Binns, H. J. (2008). Predictor and moderator effects in the treatment of oppositional defiant disorder in pediatric primary care. *J Ped Psychol, 33*(5), 462-472. doi:10.1093/jpepsy/jsm075

Lovejoy, M. C., Graczyk, P. A., O'Hare, E., & Neuman, G. (2000). Maternal depression and parenting behavior: A meta-analytic review. *Clinical Psychology Review*, *20*(5), 561-592.

McGilloway, S., Mhaille, G. N., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., ... & Donnelly, M. (2012). A parenting intervention for childhood behavioral problems: a randomized controlled trial in disadvantaged community-based settings. *J Consult Clin Psychol*, *80*(1), 116-127.

McTaggart, P., & Sanders, M. (2007). Mediators and moderators of change in dysfunctional parenting in a school-based universal application of the Triple-P Positive Parenting Programme. *Journal of Children's Services*, *2*(1), 4-17.

Miller, W., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change*. New York: The Guilford Press

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology, 13*(2), 355-375. doi: 10.1017/s0954579401002097

Muthén, L. K., & Muthén, B. O. (1998-2012). *MPlus User's Guide. 7th Ed*. Los Angeles, CA.

Patterson, G. R. (1982). Coercive family process (Vol. 3): Castalia Publishing Company

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial Boys.* Eugene, OR: Castalia

Pelham, W.E., Dishion, T.J., Tein, J.Y., Shaw, D.S., & Wilson, M.N. (2017). What doesn't work for whom? Exploring heterogeneity in responsiveness to the Family Check-Up in early childhood using a mixture model approach. *Prev Sci.* doi: 10.1007/s11121-017-0805-1

Peterson, J., Winter, C., Jabson, J., & Dishion, T. J. (2008). Relationship affect coding system. *Unpublished coding manual, University of Oregon, Child and Family Center, Eugene*.

Piquero, A.R., Jennings, W.G., Diamond, B., Farrington, D.P., Tremblay, R.E., Welsch, B.C., & Reingle Gonzales, J.M. (2016). A meta-analysis update on the effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology, 12*, 229–248. doi: 10.1007/s11292-016-9256-0

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behav Res, 42*(1), 185-227.

Radloff, L. S. (1977). The CES-D scale. *Applied Psychological Measurement, 1*(3), 385-401.

Robinson, E., Eyberg, S., & Ross, A. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology, 9*, 22-28. doi: 10.1080/15374418009532938

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In Heijmans, R.D.H., Pollock, D.S.G. & Satorra, A. (eds.), *Innovations in multivariate statistical analysis* (pp.233-247). London: Kluwer Academic.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development, 6*(2-3), 144-164.

Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*(2), 171-186.

Shaw, D. S., Connell, A., Dishion, T. J., Wilson, M. N., & Gardner, F. (2009). Improvements in

maternal depression as a mediator of intervention effects on early childhood problem behavior. *Dev Psychopathol, 21*(2), 417-439. doi:10.1017/s0954579409000236

Shaw, D. S., Sitnick, S., Brennan, L. M., Choe, D. E., Dishion, T. J., Wilson, M. N., & Gardner, F. (2016). Moderation by neighborhood deprivation on long-term effectiveness of the Family Check-Up on School-Age Conduct Problems. *Dev Psychopathol, 1471-1486.*

Shelleby, E. C., & Shaw, D. S. (2014). Outcomes of parenting interventions for child conduct problems: A review of differential effectiveness. *Child Psychiatry and Human Development*, 45(5), 628-45. doi: 10.1007/s10578-013-0431-5

Sitnick, S. L., Shaw, D. S., Gill, A., Dishion, T., Winter, C., Waller, R., Gardner, F., Wilson, M. (2015). Parenting and the Family Check-Up: Developments and changes in observed parent-child interaction during early childhood. *J of Clinical Child and Adolescent Psychology, 44* (6), 970-984. doi:10.1080/15374416.2014.940623

Smith, J. D., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2013). Indirect Effects of Fidelity to the Family Check-Up on Changes in Parenting and Early Childhood Problem Behaviors. Journal of Consulting and Clinical Psychology, 81(6), 962-974. doi: 10.1037/a0033950

Stadler, C., Poustka, F., & Sterzer, P. (2010). The heterogeneity of disruptive behavior disorders – implications for neurobiological research and treatment. *Frontiers in Psychiatry*, 1-14.

Tein, J. Y., Sandler, I. N., MacKinnon, D. P., & Wolchik, S. A. (2004). How did it work? Who did it work for? Mediation in the context of a moderated prevention effect for children of divorce. *J Consult Clin Psychol, 72*(4), 617-624. doi: 10.1037/0022-006x.72.4.617

Webster-Stratton, C. (1990). Long-term follow-up of families with young conduct problem children: From preschool to grade school. *J of Clinical Child Psychology, 19*(2), 144-149. doi: 10.1207/s15374424jccp1902_6

Zhou, Q., Sandler, I. N., Millsap, R. E., Wolchik, S. A., & Dawson-McClure, S. R. (2008). Mother-child relationship quality and effective discipline as mediators of the 6-year effects of the New Beginnings Program for children from divorced families. *Journal of Consulting and Clinical Psychology*, *76*(4), 579-594. doi: 10.1037/0022-006x.76.4.579

Appendix 1: Data Transparency

The data reported in this manuscript were collected as part of a larger longitudinal project spanning nearly a decade. Findings have been reported in over 70 different manuscripts. Many previous manuscripts focused on outcomes measured at earlier timepoints and therefore, did not include the age 9.5 data that is presented in the current study. For more recent studies utilizing the same data collection timepoints, analyses in other manuscripts have not included moderated mediation models to examine differential effectiveness by baseline CP. Other moderators that have been examined, as articulated within the current submission, include single parent status, maternal education, relationship quality, daily hassles, maternal depressive symptoms, substance use problem, teen parent status, neighborhood deprivation, and experiences of child neglect, legal problems, and parental mental health problems.

Figure 1: Participant flow chart

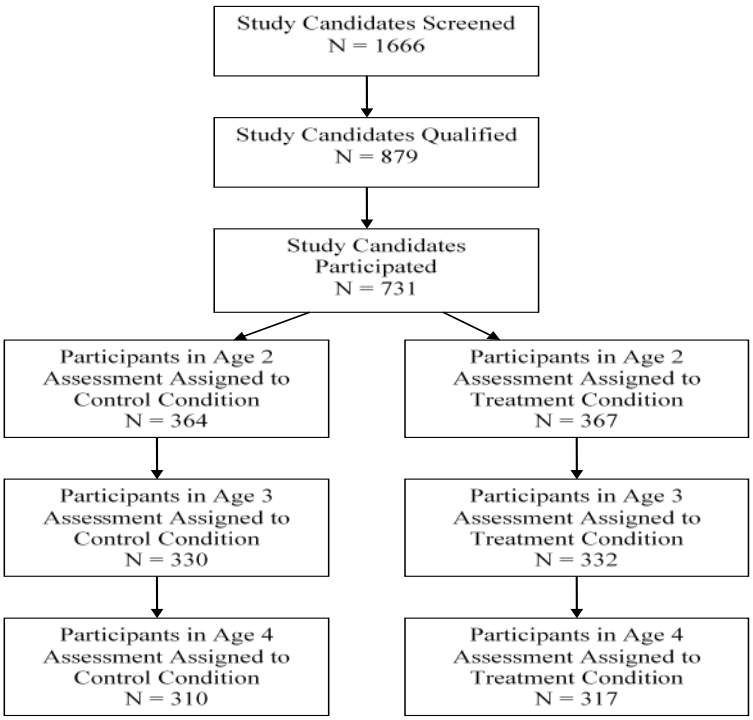
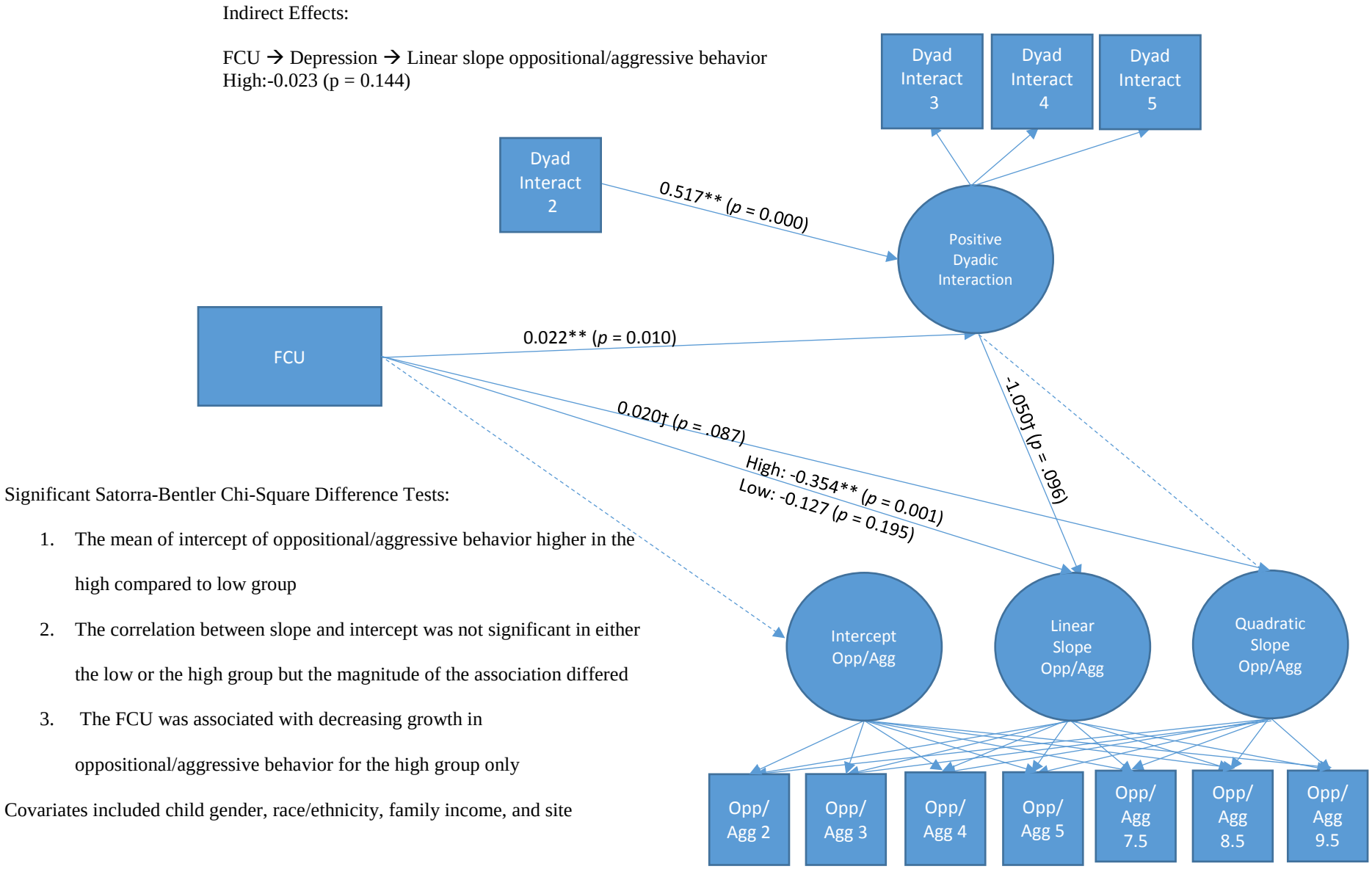


Figure 2: Multi-group parallel process growth model of positive dyadic interaction and parent-reported oppositional/aggressive behavior



Significant Satorra-Bentler Chi-Square Difference Tests:

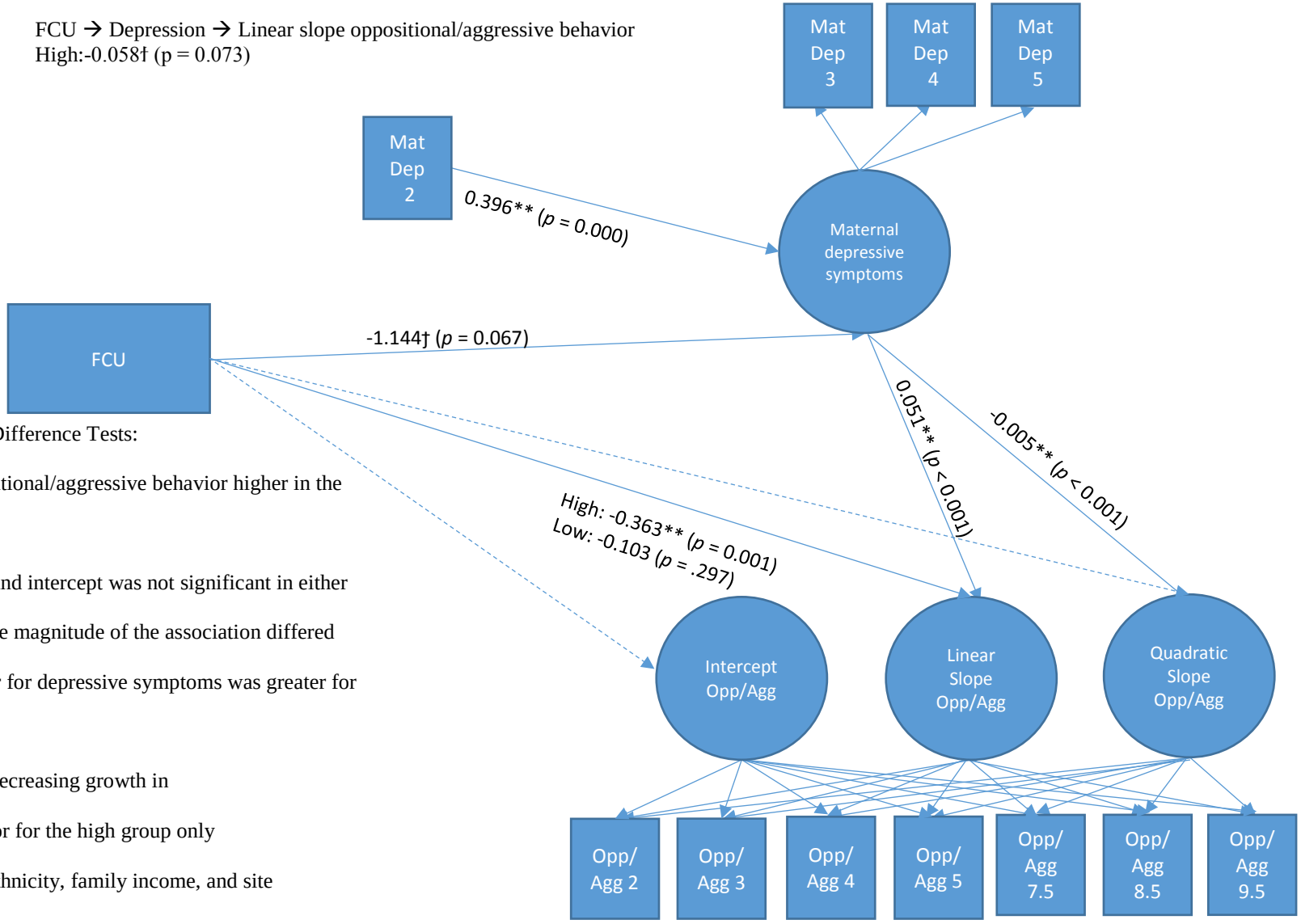
1. The mean of intercept of oppositional/aggressive behavior higher in the high compared to low group
2. The correlation between slope and intercept was not significant in either the low or the high group but the magnitude of the association differed
3. The FCU was associated with decreasing growth in oppositional/aggressive behavior for the high group only

Covariates included child gender, race/ethnicity, family income, and site

Figure 3: Multi-group parallel process growth model of maternal depressive symptoms and parent-reported oppositional/aggressive behavior

Indirect Effects:

FCU → Depression → Linear slope oppositional/aggressive behavior
High:-0.058† (p = 0.073)



Significant Satorra-Bentler Chi-Square Difference Tests:

1. The mean of intercept of oppositional/aggressive behavior higher in the high compared to low group
2. The correlation between slope and intercept was not significant in either the low or the high group but the magnitude of the association differed
3. The variance of the latent factor for depressive symptoms was greater for the high group
4. The FCU was associated with decreasing growth in oppositional/aggressive behavior for the high group only

Covariates included child gender, race/ethnicity, family income, and site