Prev Sci
DOI 10.1007/s11121-017-0805-1

What *Doesn't* Work for Whom? Exploring Heterogeneity in Responsiveness to the Family Check-Up in Early Childhood Using a Mixture Model Approach

William E. Pelham III[1] · Thomas J. Dishion[1,2] · Jenn-Yun Tein[1] · Daniel S. Shaw[3] · Melvin N. Wilson[4]

Abstract This study applied latent class analysis to a family-centered prevention trial in early childhood to identify subgroups of families with differential responsiveness to the Family Check-Up (FCU) intervention. The sample included 731 families with 2-year-olds randomized to the FCU or control condition and followed through age 5 with yearly follow-up assessments. A two-step mixture model was used to examine whether specific constellations of family characteristics at age 2 (baseline) were related to intervention response across ages 3, 4, and 5. The first step empirically identified latent classes of families based on several family risk and adjustment variables selected on the basis of previous research. The second step modeled the effect of the FCU on longitudinal change in children's problem behavior in each of the empirically derived latent classes. Results suggested a five-class solution, where a significant intervention effect of moderate to large size was observed in one of the five classes—the class characterized by child neglect, legal problems, and parental mental health issues. Pairwise comparisons revealed that the intervention effect was significantly greater in this class of families than in two other classes that were generally less at risk for the development of child disruptive behavior problems, albeit still low-income. Thus, findings suggest that (a) the FCU is most successful in reducing child problem behavior in more highly distressed, low-income families, and (b) the FCU may have little impact for relatively low-risk, low-income families. Future directions include the development of a brief screening process that can triage low-income families into groups that should be targeted for intervention, redirected to other services, monitored prospectively, or left alone.

Keywords Intervention response · Moderation · Latent class analysis

Electronic supplementary material The online version of this article (doi:10.1007/s11121-017-0805-1) contains supplementary material, which is available to authorized users.

✉ William E. Pelham, III
wpelham@asu.edu

1 Department of Psychology, REACH Institute, Arizona State University, 950 S. McAllister Rd, Tempe, AZ 85287-1104, USA

2 Oregon Research Institute, 1776 Millrace Dr, Eugene, OR 97303, USA

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

4 Department of Psychology, University of Virginia, Gilmer Hall 306, Charlottesville, VA 22908, USA

Early-onset conduct problems entail substantial costs to society and to individuals. It has long been established that 5% of early-starting individuals commit 50% of crimes (Offord et al. 1991), and those children and teenagers who are affected by early-onset conduct problems often demonstrate impaired health, happiness, occupational outcomes, and family relationships as adults as well as considerable adverse consequences to their victims and society (Dishion and Patterson 2006). Thankfully, there are a growing number of early interventions that have been found to reduce these negative outcomes and prevent growth in conduct problems (O'Connell et al. 2009), including some that are offered preventively in community service settings. Yet Offering evidence-based services to every family in a public school, Head Start, pediatric service, or supplement program is expensive and inefficient if a large percentage of the targeted audience is healthy. Thus, in the context of prevention, the classic question of "What works for whom?" raised

Published online: 27 May 2017

by Gordon Paul (1967) almost 50 years ago instead becomes "For whom is the intervention needed?"

The Family Check-Up: a Public Health Intervention Model

The Family Check-Up (FCU) is an evidence-based approach for reducing the incidence of conduct problems that was developed to address some of the limitations of the traditional parent training model. The FCU is a brief preventive intervention based on motivational interviewing and modeled after the Drinker's Check-Up (Miller and Rollnick 2002) that seeks to motivate parents to engage in services that improve the quality of their parenting practices. The family and therapist meet for three sessions—an interview, assessment, and feedback—to explore potential areas of concern and promote engagement in follow-up services (e.g., parent training, cognitive behavioral therapy) that can address any identified issues. This framework was originally developed in the context of preventing substance use and abuse during adolescence (Dishion and Kavanagh 2003), with a randomized, controlled trial indicating that providing the FCU during middle school reduced rates of initiation of substance use from sixth to ninth grade, reduced growth in substance use over later adolescence, and reduced symptoms of substance abuse in young adulthood (Connell et al. 2007; Dishion et al. 2002), in addition to having positive effects in other domains (e.g., Connell and Dishion 2008).

Based on the success of the FCU framework during early adolescence, the FCU was subsequently applied to another period of developmental transition—the toddler period. Toddlerhood is marked by an increase in children's physical mobility without a concomitant increase in the cognitive facilities required to appreciate the consequences of their behavior on the physical safety of themselves and others (Shaw and Bell 1993). During this age period, the FCU aims to prevent growth in aggressive and oppositional behavior that may lead to more severe conduct problems. A pilot study with 120 indigent families from an urban community with male 2-year-olds using Women, Infant, and Children (WIC) Nutritional Supplement centers indicated that one dose of the FCU at age 2 was associated with reduced child disruptive behavior and greater maternal involvement with children at ages 3 and 4 (effect sizes in the medium range; Shaw et al. 2006). This initial study with toddlers was followed by a much larger, multisite trial including male and female toddlers and families living in urban, rural, and suburban communities, in which substantial reductions in growth in children's externalizing behavior were found when their caregivers were offered *annual* FCUs from ages 2 to 5 (effect sizes in the small-to-medium range; Dishion et al. 2008). Follow-up of this sample into primary school has indicated that these reductions continue through age 5, and that teacher ratings at age 7.5 show significant effects of the FCU on reducing aggressive and oppositional behavior in the school context (effect size in the small range; Dishion et al. 2014). However, these analyses indicate that not all families in need of intervention changed, nor did all families seem to need family-based support. Thus, it is important to understand both the lack of responsiveness to the FCU and the characteristics of families in which intervention is unlikely needed. This knowledge would facilitate the adaptation of the FCU to meet the unique needs of specific subpopulations less likely to respond and provide the opportunity to redirect families to more or less intensive interventions.

Potential Moderators of Response to the FCU

Statistical moderation analysis (Aiken and West 1991) is one way to examine child and family characteristics that might limit or enhance the FCU's effects on various child outcomes. Three previous studies have addressed this question: first, Gardner et al. (2009) tested potential moderators of the effects of the FCU on growth in externalizing problems during early childhood. Gardner and colleagues identified teen parent status and single parenthood as family characteristics that *limited* intervention effectiveness, and lower parental education as a characteristic that *enhanced* intervention effectiveness. Second, Shaw et al. (2017) examined the moderating effects of family's neighborhood deprivation on FCU effects into late elementary school. Direct intervention effects were observed only for those two thirds of the sample experiencing moderate levels of neighborhood deprivation, rather than extreme, although indirect intervention effects of the FCU were found on teacher-reported conduct problems by successfully improving parenting during early childhood even among those living in extreme neighborhood deprivation. Third, Shelleby et al. (under review) examined the moderating effects of baseline child problem behavior on growth in parent-rated aggressive and oppositional behavior from age 2 to age 9.5. Results indicated a larger intervention effect for those children high in problem behavior at age 2 than for those low in problem behavior at baseline.

In addition to these three FCU studies, there have been many studies of moderation of other parent training-based interventions. Reyno and McGrath (2006) conducted a meta-analysis of baseline moderators of parent training efficacy and identified 31 studies examining 15 different moderators. They found lower family income, more severe child behavior, higher maternal psychopathology, lower parental education, and more barriers to treatment to have medium to large ($r \sim .30$–$.50$) associations with poorer treatment outcome. Greater number of siblings, single parenthood, and higher maternal depression were also associated with poorer treatment outcome, albeit less strongly so ($r \sim .20$). Lundahl et al. (2006) also meta-analyzed moderators of parent training efficacy, but examined fewer moderators in a wider literature base (63 studies). Despite the fact that only three studies were included in both reviews (i.e., there was little overlap), these authors concurred with Reyno and McGrath's (2006) finding that lower socioeconomic status and single parenthood limited treatment efficacy. However, whereas Reyno and McGrath found more severe child misbehavior to be associated with poorer outcomes, Lundahl et al. (2006) reached the opposite conclusion: families with children in the clinical range received *greater* effects of parent training relative to families with children in the non-clinical range.

As discussion of Reyno and McGrath (2006) and Lundahl et al. (2006) has indicated, the FCU-specific moderation findings do not map perfectly onto those from the more general parent training literature. In addition, the two reviews sometimes reached contradicting conclusions (e.g., regarding severity of child behavior), and both conducted statistical tests that indicated substantial heterogeneity in the included studies. Together, these results suggest that moderation of response to the FCU may be more nuanced and warrant a different analytic approach.

A Person-Centered Approach to Identifying Moderation of Intervention Effects

Most existing studies examine moderation of intervention effects using a variable-centered approach, modeling covariation among variables in what is presumed to be a homogenous sample. Indeed, all of the findings reviewed above used this method, typically by including a series of treatment × moderator interaction terms in a multiple regression equation. However, it may be that a particular *constellation* of family conditions presents a context that affects response to the intervention, rather than any single variable in isolation. This possibility could be investigated by incorporating several higher-order interaction terms into a variable-centered method, but this approach is fraught with elevated type I error rates and reduced power (Lanza and Rhoades 2013). These issues have motivated the development of person-centered analytic approaches (e.g., latent class analysis) that may be able to surmount these limitations by summarizing a complex multivariate pattern of characteristics via a small number of underlying groups that comprise it. In the context of moderation analysis, these approaches can offer reduced type I error rates and improved power when probing for higher-order (e.g., three- or four-way) interactions among several variables of interest (Lanza and Rhoades 2013).

Herman et al. (2007) provide an early example of this latent class approach to moderation in modeling latent profiles of co-occurring symptomology (e.g., anxiety, oppositional defiance) in the Treatment for Adolescents With Depression Study (TADS). They examined treatment × latent class interactions to determine if the latent profiles moderated intervention effectiveness. None of these interactions were significant, perhaps owing to the modest sample size ($N = 423$, partitioned into five classes). More recently, Cooper and Lanza (2014) applied a person-centered moderation methodology to the Head Start Impact Study (3-year-old cohort, $N = 2449$), conducting a latent class analysis on the sample and then examining intervention effects in each latent class. Their results provide a more compelling illustration of the ability of this quantitative approach to illuminate the critical nuances affecting intervention effects. Five latent classes were identified, two of which experienced mostly positive intervention effects, two of which experienced no intervention effects, and one of which possibly experienced iatrogenic effects. The most robust effects were observed for a class characterized as married, English-language learners with lower education, whereas Head Start appeared to have little effect in a class characterized as married, lower-risk families. Cooper and Lanza's (2014) results painted a very different picture than would have a traditional, variable-centered method, illustrating the potential of the person-centered approach to clarify response to intervention.

Present Study

A person-centered approach might complement traditional means of identifying families more or less likely to respond to the FCU, enabling implementers to preserve finite resources and ensure the receipt of appropriate services. The present study applied this person-centered methodology to the Early Steps Multisite Trial, a large randomized, controlled trial of the FCU in early childhood (Dishion et al. 2014; Shaw et al. 2017). A two-step mixture model was used to examine whether specific constellations of family characteristics at age 2 (baseline) could be identified and whether the effects of the FCU on parent-reported child conduct problems over time (ages 2, 3, 4, and 5) varied across these constellations.

Methods

Participants Seven hundred thirty-one at-risk parents and their toddlers were recruited from the Woman, Infants, and Children (WIC) Nutritional Supplement program in three different types of communities: Eugene, OR (suburban), Charlottesville, VA (rural), and Pittsburgh, PA (urban). Parents were invited to participate if they had a 2-year-old child and possessed two of the three following risk factors for future behavior problems: current child behavior problems, family problems (e.g., maternal depression), and sociodemographic risk. Primary caregivers were almost universally mothers (16 fathers, 2.2%). Racial and ethnic background was as follows: 50% European American, 28% African American, 13% biracial, 9% other, and 13% Hispanic. Sixty-six percent of the sample had an income below \$20,000, and 41% had a high school diploma. See Dishion et al. (2008) for more details about the recruiting process and sample characteristics.

Design

Families were randomly assigned to either a control or an intervention condition when the child was age 2. Participants in the control condition did not receive any intervention services from the study, but were free to seek support services in the community. Participants in the intervention condition gained access to services implementing the Family Check-Up (FCU) model. The FCU comprised three sessions: (1) *assessment session*—the interviewer went to the home and videotaped the parent and child while they engaged in various tasks selected to evaluate parent-child interactions; (2) *initial interview*—the interviewer explored parent concerns and stage of change and encouraged parents to participate in an in-home assessment of family functioning; (3) *feedback session*—the interviewer provided feedback based on the assessment while seeking to promote reflection on behavior change and on potential engagement in further intervention services. This ordering of sessions deviated from typical FCU procedure (i.e., begin with interview) so that home visitors could be blinded to the intervention status of the family during the assessment phase (for more details about the FCU intervention, see Dishion and Kavanagh 2003). After completing the FCU, parents were able to engage in as-desired follow-up parenting support services such as parent training (e.g., Dishion et al. 2011).

Intervention-assigned families were re-contacted annually at ages 3, 4, and 5 and were offered the same FCU plus follow-up services package. The present analysis focuses on the effects of FCU in early childhood, using data from ages 2 through 5. When defining *engagement* in the intervention to require completion of (at least) the FCU feedback session, 76% of families engaged in the intervention at age 2, 69% at age 3, 70% at age 4, and 66% at age 5. In total, over 90% of families engaged in at least one feedback between ages 2 and 5.

Baseline Measures

Ten different variables were collected at baseline (age 2) and entered as indicators in a latent class analysis. These specific variables were selected for meeting three requirements: (a) they were among those indicated by the extant literature to be potential moderators of the intervention's effectiveness, (b) they were collected at baseline in the present study, and (c) they were commonly collected demographics or easily collectable further data (e.g., no direct observation-based scores). Descriptives for all ten variables appear in Table 1.

Child Externalizing Behavior Primary caregiver completed the age 1.5 to 5 version of the Child Behavior Checklist (Achenbach and Rescorla 2001). The 24-item raw total score on the Externalizing subscale was used as a broadband measure of disruptive behavior ($\alpha = 0.86$).

Family Income Primary caregiver reported monthly household income (including child support and other financial aid) on an approximately linear categorical scale where answers ranged from "\$415 or less" (coded as 1) to "\$7500 or more" (coded as 13). This variable was treated as continuous for these analyses.

Number of Children in Household Primary caregiver reported the number of children currently living in the household.

Parental Depression Primary caregiver reported on their depressive symptoms using the Center for Epidemiologic Studies Depression scale (CES-D; Radloff 1977; $\alpha = 0.76$).

Child Gender Child gender was coded as 0 = female, 1 = male.

Parental Education Primary caregiver reported his or her educational history. This was used to form a categorical variable scored as a 1 (less than high school), 2 (high school graduate through partial college), and 3 (junior college degree or more).

Single Parent Status Primary caregiver reported whether he or she currently had a live-in partner; this formed a binary indicator of single parent status.

Household Law Problems Primary caregiver reported whether persons living in the home had had trouble with the law since the child was born; this formed a binary indicator of household law problems.

Household Child Abuse Primary caregiver reported whether persons living in the home had been reported for child abuse since the child was born; this formed a binary indicator of household child abuse.

Household Mental Health Treatment Received Primary caregiver reported whether persons living in the home had been treated by a mental health professional since the child

Table 1 Descriptives for all variables

Variable	Control	Intervention
Number of participants	364	367
Baseline variables for latent class analysis		
Parent CBCL externalizing behavior	20.6 (7.0)	20.8 (7.6)
Family income	4 [2–5]	4 [2–5]
Number of children in household	2 [2–3]	2 [2–3]
Parental depression (CES-D)	14 [9–22]	15 [9–23]
Parental education	25/64/11%	22/66/12%
Child male	51%	50%
Single parent status	42%	38%
Household law problems	35%	34%
Household child abuse	7%	8%
Household mental health treatment received	39%	38%
Parent ratings for growth model		
Mean score CBCL aggressive/oppositional items at age 2	0.64 (0.32)	0.67 (0.35)
Mean score CBCL aggressive/oppositional items at age 3	0.56 (0.35)	0.55 (0.35)
Mean score CBCL aggressive/oppositional items at age 4	0.52 (0.35)	0.47 (0.36)
Mean score CBCL aggressive/oppositional items at age 5	0.47 (0.36)	0.43 (0.35)

Note. Where numbers are followed by brackets, they are in this form: median [25th–75th percentiles]. Where they are followed by parentheses, they are in this form: mean (standard deviation). Baseline balance was achieved. See "Methods" section for description of variable measurement scales (e.g., for family income). All values are complete-case, with Ns per variable as indicated in the "Handling of Missing Data" section.

was born; this formed a binary indicator of household mental health problems.

Dependent Measure

Parent Ratings of Aggressive/Oppositional Behavior Primary caregiver completed the CBCL at ages 2, 3, 4, and 5. Eight items previously used to map onto symptoms of oppositional defiant and conduct disorder (Brennan et al. 2015) from the CBCL were averaged to create a score ranging from 0 to 2 (0 = not true; 1 = somewhat true; 2 = very true) for each child at each age. Alpha reliability of this score was 0.71 at age 2, 0.75 at age 3, 0.78 at age 4, and 0.80 at age 5. Descriptives at each age appear in Table 1.

Analytic Plan

A two-step mixture model was implemented in MPLUS 7.3 (Muthén and Muthén 2012). A two-step procedure was necessary in order to incorporate latent class analysis, latent growth modeling, and missing data handling into the model and still attain convergence.

Step 1: Latent Class Analysis In step 1, a mixture model was fit to identify latent classes of families using baseline variables thought to be potential youth, parent, and family risk and protective factors. The variables listed above under "Baseline Measures" were all included as indicators of the latent class. Continuous variables were standardized and modeled as normally distributed. The binary variables (e.g., gender) and parental education were all modeled as categorical or ordinal.

Selecting the number of latent classes in a mixture model remains a subjective process, as various fit statistics perform differently in simulations and often contradict each other (Tein et al. 2013). In the present study, we based this decision on the theoretical plausibility of the solution and on three fit indices: the sample-adjusted Bayesian information criterion (saBIC; Sclove 1987), the Lo-Mendell-Rubin (LMR) likelihood ratio test (Lo et al. 2001), and the bootstrapped likelihood ratio test (BLRT; McLachlan and Peel 2000). Solutions with one to six latent classes were produced sequentially, and the results were evaluated according to these criteria.

Step 2: Multiple-Group Latent Growth Model In step 2, the latent classes identified in step 1 were treated as observed by assigning each family to its most likely class. A multiple-group latent growth model was then fit to examine whether the FCU has differential effects on the trajectories of child outcomes across the latent classes identified in step 1. The "groups" were the latent classes from the mixture model, and the "growth" was in aggressive and oppositional behavior from ages 2 to 5 (CBCL). We adopted the analytic procedure of the primary outcome paper (Dishion et al. 2014), specifying linear growth. The latent linear slope factor was regressed on intervention status and allowed to vary across the multiple groups to evaluate the FCU's effect in each of the latent classes (model depicted in Fig. 1). The latent intercept factor was also regressed on intervention status because (despite randomization) there were sometimes non-negligible intervention/control differences in baseline child behavior within the smaller classes (standardized differences ranged from 0.02 [class 2] to 0.52 [class 4]). The model-estimated intervention effect size in each latent class was computed by multiplying the coefficient relating intervention status to the slope factor by 3 (the number of time intervals) and dividing the result by the full-sample standard deviation in aggressive and oppositional behavior at baseline (SD = 0.34; Eq. 7 in Feingold 2015). Finally, to examine whether the effect of the FCU was *moderated* by latent class, the MODEL CONSTRAINT command was used to test the significance of *differences* in intervention status coefficient across each of the latent classes.

Handling of Missing Data Baseline family characteristics at age 2 all had less than 2% missing data. Primary caregiver ratings of aggressive/oppositional behavior ranged from 84 to 100% complete: age 2 (100%), age 3 (90%), age 4 (85%), and age 5 (84%). At the participant level, 539 of 731 (74%) participants had complete data for all the variables in both steps of the analysis. We used full-information maximum likelihood (FIML) estimation with auxiliary variables to address missing data (more details available online), assuming a Missing at Random (MAR) mechanism (Rubin 1976).

Results

Step 1: Latent Class Analysis

Fit statistics for the various *k*-class solutions suggested a three-, four-, five, or six-class solution was viable (see Table 2). Adding a fourth class separated a small class of distinct families (class 4 below), and all three fit statistics indicated that this addition produced statistical improvement. Adding a fifth class drew from the two largest classes in the four-class solution to produce a sizeable class (class 5 below) with dramatic differences from all other classes on the categorical indicators. Adding a sixth class yielded only marginal improvement in BIC, and the LMR likelihood ratio test did *not* indicate significantly improved model fit. Thus, we settled on a five-class solution.

The five identified latent classes can be roughly characterized as follows:

- *Class 1 (N = 181)*—relatively high income, low-risk
- *Class 2 (N = 105)*—low income, lower education, very high maternal depression, high single parenthood
- *Class 3 (N = 323)*—low income, lower education, high single parenthood, otherwise low-risk
- *Class 4 (N = 29)*—lower education, high child behavior problems, very high number of kids, high parental neglect, high maternal depression
- *Class 5 (N = 93)*—high legal problems, very high neglect, extremely high parental mental health treatment

Fig. 1 Diagram of latent growth model fit within each latent class. Note. *Agg./Opp.* aggressive and oppositional behavior. The model was fit within each of the five classes identified, and all free parameters were allowed to vary across classes

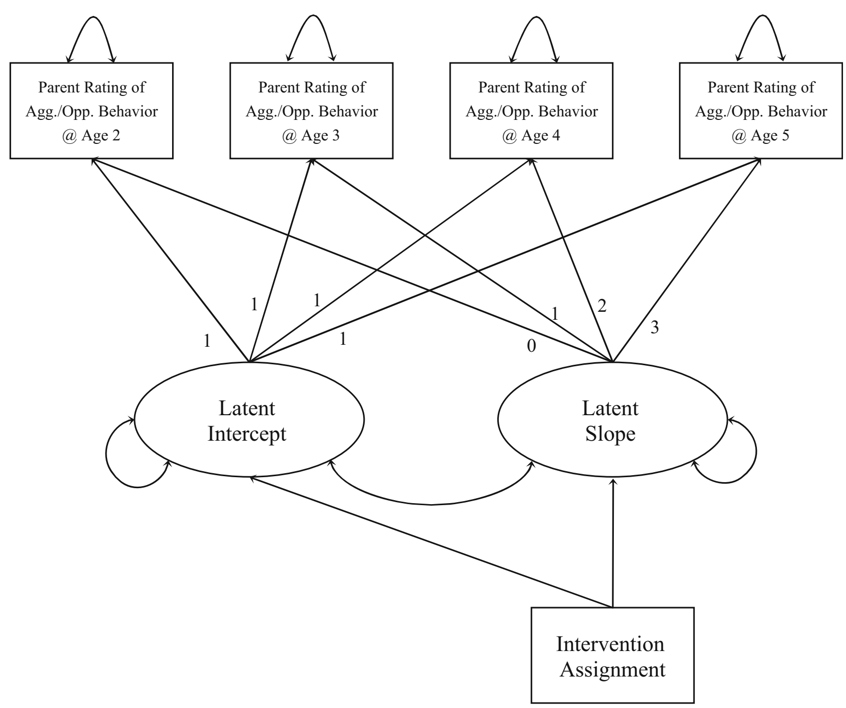

Table 2 Fit statistics for latent class analysis solutions

Fit indicator	1-class solution	2-class solution	3 class-solution	4-class solution	5-class solution	6-class solution
Sample size-adjusted BIC	13,875	13,676	13,581	13,533	13,506	13,491
Lo-Mendell-Rubin LR test	–	$p < .0001$	$p < .01$	$p < .05$	$p < .10$	*ns*
Bootstrapped LR test	–	$p < .0001$	No convergence	$p < .0001$	$p < .0001$	$p < .0001$
Sizes of classes (*N*s)	731	488, 243	211, 125, 395	388, 109, 205, 29	181, 105, 323, 29, 93	76, 240, 163, 104, 118, 30

The five-class solution was selected.

BIC Bayesian information criterion, *LR* likelihood ratio

Note that descriptors such as "low" and "high" are *relative to the rest of the present sample*—for example, the "high income" of class 1 corresponded to just \$25,000–30,000 per year (vs. ~\$10,000–15,000 for "low income" of classes 2 and 3). The exact profile of each of the identified classes across the ten baseline (age 2) family characteristics is depicted in Fig. 2 (continuous indicators in upper panel, categorical indicators in lower panel). Entropy for this solution was 0.74, suggesting there was some uncertainty in the process of assigning individual families to classes.

Step 2: Multiple-Group Latent Growth Model

We next examined the effects of assignment to the Family Check-Up on growth in aggressive-oppositional behavior within each of these five latent classes. Fit statistics indicated adequate model fit (RMSEA = 0.079, CFI = 0.97, SRMR = 0.078), and estimates of effects are reported in Table 3. A significant intent-to-treat effect of randomization to the Family Check-Up was observed in class 5 ($p < .01$; $d = -0.63$), which was characterized by high rates of parental neglect, legal problems, and mental health issues. Pairwise comparisons indicated that the effect in class 5 was significantly greater than the effect in either class 1 ($p < .05$; $d = -0.01$), which consisted of higher-income, lower-risk families, or class 3 ($p < .05$; $d = -0.08$), which consisted of low-income, single-parent families that were otherwise at low risk. Note that the effect size for class 4 was larger than for class 5, but its tiny size ($n = 29$) limited power to detect an effect. Thus, results suggested that the effects of random assignment to the FCU were more pronounced in distressed families compared to those characterized as low risk.

Post Hoc Analysis

Based on the uncertainty for individual latent class membership in the distressed group (class 5), we next formulated three groups using simplified, researcher-specified definitions based on the pattern of findings revealed in the latent class analysis. These definitions separated the sample into three classes—(A) low-risk, (B) demographic risk, and (C) demographic plus parental mental health risk—on the basis of five of the indicator variables (parental depression, history of mental health treatment, history of legal problems, single parent status, and income). The exact class criteria are presented in Table 3. We then conducted multigroup modeling and fit the same latent growth model shown in Fig. 1 within each of these three researcher-specified classes: results are reported in Table 3. Consistent with previous findings, a significant intervention effect was observed only in the class with both demographic and mental health risk (class C; $p < .01$; $d = -0.56$), and the effect in this class was marginally significantly greater than that in either of the two classes without both types of risk factors (classes A and B; $p < .10$; $d = -0.15$ and -0.04).

Discussion

Five different latent classes of families were identified, and the effect of random assignment to the FCU in early childhood was examined in each. Results indicated that the intervention had a moderate-to-large effect size in reducing parent-rated problem behavior in the class of families characterized by high rates of parental neglect, legal problems, and mental health issues. Pairwise comparisons among the classes indicated that the intervention effect was significantly greater in this class of distressed families than in two other classes that were generally less at-risk for the development of children's disruptive behavior problems. Post hoc analyses also indicated a moderating role of mental health issues. We now discuss these results and their implications.

Family Support for Distressed Families with Young Children

Note that this study involved a large group of low-income families seeking nutritional support for their children through the WIC program—*not* a group of families seeking intervention services for their children. Within that context, our results suggest that those low-income families with high rates of legal problems, child neglect, and mental health treatment were *more* responsive to the FCU. The outcome of this study mirrors those from earlier, variable-centered analyses of this dataset showing that families with more risk factors benefited

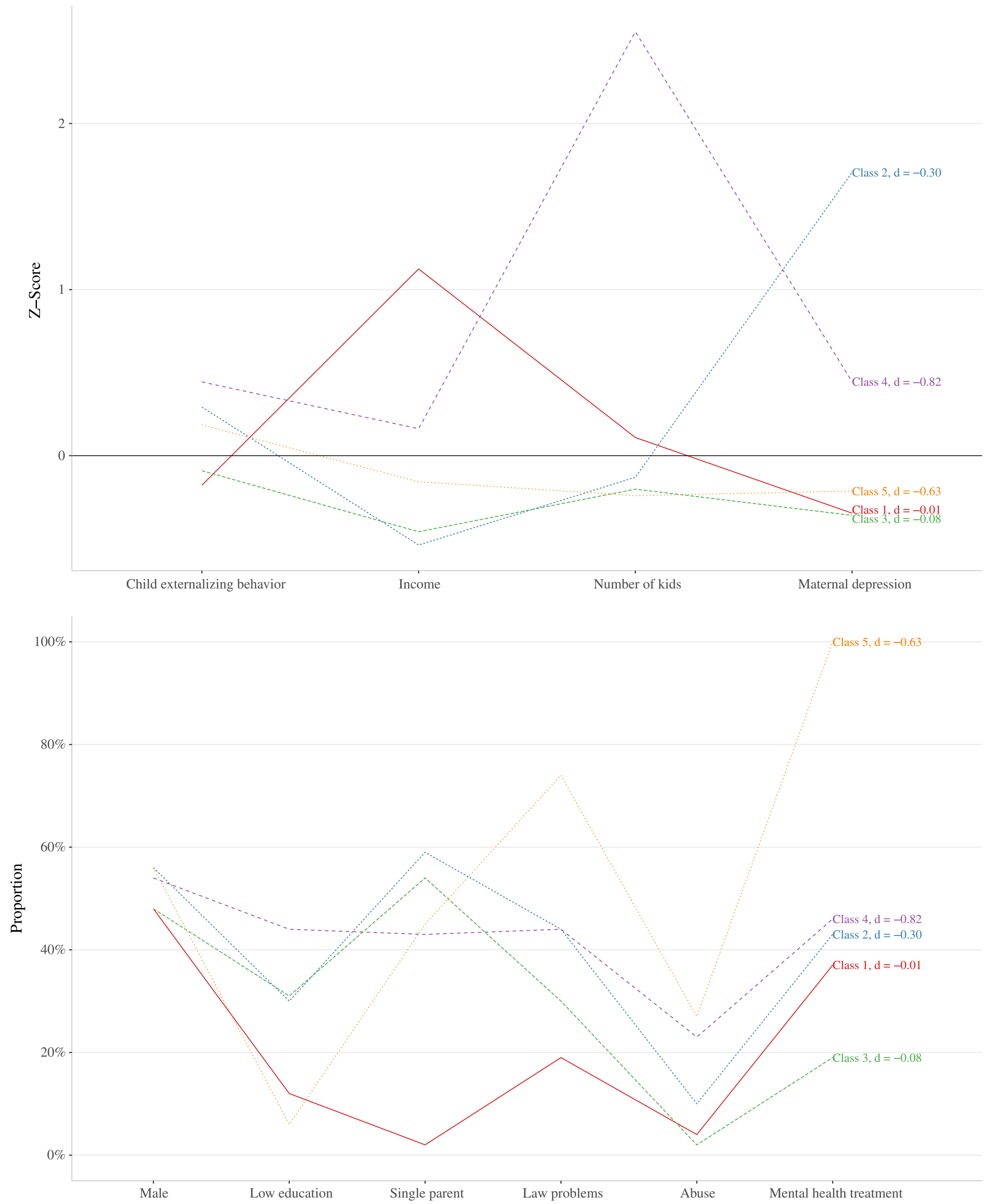

Fig. 2 Profiles of latent classes on continuous and categorical indicators. Note. In the *top panel*, the *y-axis* represents the mean *z*-score of each class on each *continuous* indicator variable. In the *bottom panel*, the *y-axis* represents the proportion of each class endorsing the *categorical* indicator variables. Values calculated based on most likely class membership. "d" indicates intervention effect size, as reported in Table 2

more from the intervention (Gardner et al. 2009; Shelleby et al. under review). This pattern was also seen in recent analyses showing that FCU-based reductions in neglectful parenting were greatest for those families with greater family adversity (Dishion et al. 2015), and that parents with greater perceived parenting stress were considerably more likely to engage in the intervention (Smith et al. 2017). The fact that families at relatively low risk (i.e., classes 1 and 3) did not appear responsive to the FCU in this study may indicate that these children are less likely to develop the problem behavior to be prevented, or are less likely to have the poor parenting practices that can be improved via intervention. Indeed, a desirable feature of any preventive intervention is that it reaches and benefits the most in-need families.

Table 3 Subgroup intervention effects on parent-rated aggressive and oppositional behavior

	Number	Est. (SE)	Estimated effect size
Latent classes			
Class 1—very high income, low-risk	181	−.001 (.014)	$d = -0.01$
Class 2—low income, very high maternal depression, high single parenthood	105	−.034 (.026)	$d = -0.30$
Class 3—low income, high single parenthood, otherwise low-risk	323	−.009 (.014)	$d = -0.08$
Class 4—high behavior problem, very high number of kids, high neglect, high maternal depression	29	−.092 (.075)	$d = -0.82$
Class 5—high law problems, very high neglect, extremely high mental health treatment**	93	−.070 (.026)	$d = -0.63$
Researcher-defined classes			
Class A low risk			
Did not meet criteria of either other classes	493	−.017 (.011)	$d = -0.15$
Class B demographic risk			
Either single parent or lower tercile income (<$1250/month) CES-D ≤15, neither mental health treatment nor legal problems	105	−.005 (.021)	$d = -0.04$
Class C demographic risk + mental health risk**			
Either single parent or lower tercile income (<$1250/month) CES-D >15, either or both of mental health treatment or legal problems	133	−.063 (.022)	$d = -0.56$

Note that descriptors (e.g., low, high) are relative to the rest of the sample. "Est." is the coefficient of latent slope regressed on dummy-coded intervention status (see Fig. 1). "Estimated effect size" reflects the total effect across the age 2 to 5 span, as described in the "Methods" section. Negative effect sizes indicate advantage of intervention over control. For the latent classes, pairwise tests indicated significant differences in effects between class 1 and class 5 ($p < .05$) and between class 3 and class 5 ($p < .05$). For the research-defined classes, pairwise tests indicated that the effect in class C was nearly significantly different from that in class A ($p = .054$) or class B ($p = .055$).

† $p < .10$; * $p < .05$; ** $p < .01$, by z-tests

The intervention appeared to be most effective for those families with high rates of child abuse (classes 4 and 5). In conjunction with recent results indicating that the FCU can reduce neglectful parenting during directly observed parent-child interaction (Dishion et al. 2015), this finding suggests potential utility of the FCU in the child welfare setting. Families at risk for neglect may benefit from receiving the FCU during early childhood *before* the child has been removed from the home, and more intensive services are needed (Dishion et al. 2016). Moreover, the FCU could prove useful in reducing the rate of placement failure for children identified as at risk for disruptions because of problem behavior (Chamberlain et al. 2006)—this possibility should be explored in a future study.

The Latent-Class-as-Moderator Approach for Prediction

Moderation analysis is useful for understanding intervention processes, but it can also be used to estimate the likelihood a specific family will respond to the intervention. The present model can calculate a predicted effect size of the FCU for each family in the sample, accounting for uncertainty in class membership. This quantity is of interest because families for whom the predicted effect is quite large could be especially targeted for intervention, while families for whom the predicted effect is quite small could be redirected to other services, monitored prospectively, or left alone. Indeed, the present results suggest that there may be substantial variability in responsiveness to the FCU among low-income families in the WIC program.

This variability in responsiveness to the FCU is an important consideration for real-world implementation. Approximately 50 items are needed to yield all of the baseline characteristic scores that were included in the current latent class model (8 individual items, 24 items from the CBCL-externalizing subscale, and 20 items from the CES-D; further analyses may support the use of an even briefer assessment [e.g., an abbreviated version of the CBCL]). With this information, straightforward arithmetic is needed to produce estimated probabilities of membership in each class and thus a predicted effect size for a specific family. Administration of the items and instantaneous calculation of the predicted effect size could be accomplished via a simple web application. The application could then display a predicted effect size and/or recommended action (e.g., "probably helpful," "maybe helpful," "probably not helpful") that is customized to audience (i.e., parent, therapist, physician). Thus, in fewer than 10 min, families could be evaluated for their potential need and responsiveness to a potential FCU, and parents and providers could receive tailored, practical advice.

The implementation outlined above is straightforward, but it is still aspirational. Future work must address several issues. First, the present analyses were conducted using data from a sample of families in WIC demonstrating multiple risk factors for the development of child conduct problems, and thus, the prediction model was fit in this context. The extent to which the prediction equation would generalize to populations that are more (e.g., families seeking treatment for behavioral problems) or less (e.g., primary care) at-risk is unknown. Second, and most important, cross-validation of the prediction model is needed to determine its accuracy out of sample (Hastie et al. 2009). How to best cross-validate predicted causal effects is still an active area of research (e.g., Athey and Imbens 2016). At present, we have no guarantee that our predicted effect sizes are accurate; this is obviously of paramount importance.

Limitations

A limitation of the present study is its definition of response to intervention exclusively through parent (i.e., largely mother) ratings of aggressive and oppositional behavior. Although this was the primary outcome of the multisite trial, other studies have demonstrated ancillary effects of the FCU in the domains of maternal depression (Shaw et al. 2009), positive parenting (Dishion et al. 2008), and teacher rating of problem behavior (Dishion et al. 2014), among others. Thus, families we have presently identified as *not* benefiting from the intervention (e.g., class 1) may in fact have seen positive effects in one of these other domains. Future work could repeat the present methodology but define response to intervention through a broader, composite measure.

A second limitation is that our analysis conflated assignment to treatment with receipt of treatment. Each year, approximately 25 to 35% of those assigned to the intervention did *not* engage in a feedback session or follow-up services (Dishion et al. 2014). Thus, a specific class of families may herein be identified as not responding (a) because they fail to engage in the FCU or (b) because they do engage in the FCU, but do not benefit from engagement. Comparison of the number of annual feedbacks completed across classes revealed that families in class 3 ($M = 2.2$ feedbacks) and class 4 ($M = 2.1$) engaged less than those in class 1 ($M = 2.9$), class 2 ($M = 2.9$), or class 5 ($M = 2.7$), suggesting that the difference in observed FCU effect between classes 3 and 5 is potentially related to engagement rates. Future work should investigate the effects of levels of engagement on outcomes using methods that can yield causal inferences in the presence of non-engagement (Imbens and Rubin 2015).

Finally, other limitations stem from our mixture model approach to moderation. First, because the classes varied across all of the baseline characteristics, it is unknown whether the differential effectiveness of the FCU across classes is indeed attributable to the prominent discrepancies we identified (i.e., legal problems, neglect, mental health problems). It may be that the differential effectiveness is because of other variables upon which the classes differed less (e.g., child externalizing behavior), or to confounding variables that were entirely absent from the model (Hernan and Robins 2016). Second, mixture models will identify multiple classes when the indicator variables depart substantially from normality (McLachlan and Peel 2000), and the specific pattern of the present results is consistent with the expected methodological artifacts (i.e., the classes differ most substantially on the highly skewed variables). This concern emphasizes the need to avoid reifying the classes and instead conceptualize them as a potential tool for predicting intervention response (Sterba and Bauer 2010). Third, our research question could also have been approached using a single-step growth mixture model in which the latent classes are simultaneously indicated by potential moderators and characterized by differential responsiveness to the FCU, and these results may have complemented or contradicted the current ones. Unfortunately, we were not able to achieve convergence in models using this approach.

Conclusion

The present results used a latent-class-as-moderator approach to identify a class of highly distressed families for whom the effect of the FCU was substantial and to identify non-trivial subsamples for which the effect on problem behavior appears to be limited. Critically, these latent classes were indicated by characteristics of the families at the baseline assessment. If implementers of the FCU can indeed identify non-responsive families *before* initiating the intervention, they can reduce costs and increase efficacy (i.e., be more efficient). A similar approach might apply to home visitation programs more widely (e.g., Nurse-Family Partnership; Olds 2006). Thus, in the prevention context, the classic question of "What works for whom?" might fruitfully be reframed as "What *doesn't* work for whom?" Future work should address the limitations of the present study and seek actionable answers to this question.

Compliance with Ethical Standards

Funding This research was supported by the National Institute on Drug Abuse (NIDA) grant DA036832-01A1 to the second, fourth, and fifth authors. The first author was supported during preparation of this manuscript by a NIDA T32 DA039772.

Conflict of Interest Dr. Dishion is the developer of the Family Check-Up program. The remaining authors declare that they have no conflicts of interest.

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

Informed Consent Informed consent/assent was obtained from all participants in this study.

References

Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA School-age Forms & Profiles: Child Behavior Checklist for Ages 6-18, Teacher's Report Form, & Youth Self-Report*. Burlington: University of Vermont, Research center for Children, Youth & Families.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks: SAGE.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. Proceedings of the National Academy of Sciences, 113(27), 7353–7360. doi:10.1073/pnas.1510489113.

Brennan, L. M., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2015). The predictive utility of early childhood disruptive behaviors for school-age social functioning. *Journal of Abnormal Child Psychology, 43*, 1187–1199.

Chamberlain, P., Price, J. M., Reid, J. B., Landsverk, J., Fisher, P. A., & Stoolmiller, M. (2006). Who disrupts from placement in foster and kinship care? *Child Abuse & Neglect, 30*, 409–424. doi:10.1016/j.chiabu.2005.11.004.

Connell, A. M., & Dishion, T. J. (2008). Reducing depression among at-risk early adolescents: Three-year effects of a family-centered intervention embedded within schools. *Journal of Family Psychology, 22*, 574–585. doi:10.1037/0893-3200.22.3.574.

Connell, A. M., Dishion, T. J., Yasui, M., & Kavanagh, K. (2007). An adaptive approach to family intervention: Linking engagement in family-centered intervention to reductions in adolescent problem behavior. *Journal of Consulting and Clinical Psychology, 75*, 568–579. doi:10.1037/0022-006X.75.4.568.

Cooper, B. R., & Lanza, S. T. (2014). Who benefits most from Head Start? Using latent class moderation to examine differential treatment effects. *Child Development, 85*, 2317–2338. doi:10.1111/cdev.12278.

Dishion, T. J., & Kavanagh, K. (2003). *Intervening in adolescent problem behavior: A family-centered Approach*. New York: Guilford Press.

Dishion, T. J., & Patterson, G. R. (2006). The development and ecology of antisocial behavior. In: *Developmental psychopathology: vol. 3, risk, disorder, and adaptation*. (2nd ed., pp. 503–541). Hoboken: Wiley.

Dishion, T. J., Kavanagh, K., Schneiger, A., Nelson, S., & Kaufman, N. K. (2002). Preventing early adolescent substance use: A family-centered strategy for the public middle school. *Prevention Science, 3*, 191–201. doi:10.1023/A:1019994500301.

Dishion, T. J., Shaw, D., Connell, A., Gardner, F., Weaver, C., & Wilson, M. (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*, 1395–1414. doi:10.1111/j.1467-8624.2008.01195.x.

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

Dishion, T. J., Brennan, L. M., Shaw, D. S., McEachern, A. D., Wilson, M. N., & Jo, B. (2014). Prevention of problem behavior through annual family check-ups in early childhood: Intervention effects from home to early elementary school. *Journal of Abnormal Child Psychology, 42*, 343–354. doi:10.1007/s10802-013-9768-2.

Dishion, T. J., Mun, C. J., Drake, E. C., Tein, J.-Y., Shaw, D. S., & Wilson, M. (2015). A transactional approach to preventing early childhood neglect: The Family Check-Up as a public health strategy. *Development and Psychopathology, 27*, 1647–1660.

Dishion, T. J., Forgatch, M., Chamberlain, P., & Pelham III, W. E. (2016). The Oregon model of behavior family therapy: From intervention design to promoting large-scale system change. *Behavior Therapy*. doi:10.1016/j.beth.2016.02.002.

Feingold, A. (2015). Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *Journal of Consulting Clinical Psychology, 83*, 157–168.

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. *Journal of Consulting and Clinical Psychology, 77*, 543–553.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer Science & Business Media.

Herman, K. C., Ostrander, R., Walkup, J. T., Silva, S. G., & March, J. S. (2007). Empirically derived subtypes of adolescent depression: Latent profile analysis of co-occurring symptoms in the Treatment for Adolescents with Depression Study (TADS). *Journal of Consulting and Clinical Psychology, 75*, 716–728.

Hernan, M., & Robins, J. (2016). Causal Inference. Boca Raton, FL: Chapman & Hall / CRC.

Imbens, G. W., & Rubin, D. B. (2015). Causal Inference in Statistics, Social, and Biomedical Sciences. Cambridge University Press.

Lanza, S. T., & Rhoades, B. L. (2013). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science, 14*, 157–168.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767–778. doi:10.1093/biomet/88.3.767.

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review, 26*, 86–104.

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). New York: Guilford Press.

Muthén, L. K., & Muthén, B. (2012). *Mplus User's Guide, Seventh Edition*. Los Angeles: Muthén & Muthén.

O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington D.C.: National Academies Press.

Offord, D. R., Boyle, M. H., & Racine, Y. A. (1991). The epidemiology of antisocial behavior in childhood and adolescence. In K. H. Rubin & D. J. Pepler (Eds.), *The development and treatment of childhood aggression* (pp. 31–54). Hillsdale: Erlbaum.

Olds, D. L. (2006). The nurse–family partnership: An evidence-based preventive intervention. *Infant Mental Health Journal, 27*, 5–25. doi:10.1002/imhj.20077.

Paul, G. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31*, 109–118. doi:10.1037/h0024436.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems—A meta-analytic review. *Journal of Child Psychology and Psychiatry, 47*, 99–111. doi:10.1111/j.1469-7610.2005.01544.x.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*, 581–592.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52*, 333–343. doi:10.1007/BF02294360.

Shaw, D. S., & Bell, R. Q. (1993). Developmental theories of parental contributors to antisocial behavior. *Journal of Abnormal Child Psychology, 21*, 493–518.

Shaw, D. S., Dishion, T. J., Supplee, L., Gardner, F., & Arnds, K. (2006). Randomized trial of a family-centered approach to the prevention of early conduct problems: 2-year effects of the family check-up in early childhood. *Journal of Consulting and Clinical Psychology, 74*, 1–9. doi:10.1037/0022-006X.74.1.1.

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. *Development and Psychopathology, 21*, 417–439.

Shaw, D. S., Sitnick, S. L., Brennan, L. M., Choe, D. E., Dishion, T. J., Wilson, M. N., & Gardner, F. (2017). The long-term effectiveness of the Family Check-Up on school-age conduct problems: Moderation by neighborhood deprivation. *Development and Psychopathology, FirstView, 28*, 1471–1486 1–16. doi:10.1017/S0954579415001212.

Shelleby, E. C., Shaw, D. S., Dishion, T. J., Wilson, M. N., & Gardner, F. (under review). Effects of the Family Check-Up intervention on reducing growth in conduct problems from toddlerhood through school age: An analysis of moderated mediation.

Smith, J. D., Berkel, C., Hails, K. A., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2017). Predictors of participation in the Family Check-Up program: A randomized trial of yearly services from age 2 to 10 years. *Prevention Science*, 1–11.

Sterba, S. K., & Bauer, D. J. (2010). Matching method with theory in person-oriented developmental psychopathology research. *Development and Psychopathology, 22*, 239–254. doi:10.1017/S0954579410000015.

Tein, J.-Y., Coxe, S., & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 20*, 640–657. doi:10.1080/10705511.2013.824781.