Development and Psychopathology (2024), 1–13
doi:10.1017/S095457942300144X

CAMBRIDGE UNIVERSITY PRESS

Regular Article

Collateral effects of Coping Power on caregiver symptoms of depression and long-term changes in child behavior

Lissette M. Saavedra[1], John E. Lochman[2], Antonio A. Morgan-López[1], Heather L. McDaniel[3], Catherine P. Bradshaw[3], Nicole P. Powell[2], Lixin Qu[2], Alexa Budavari[3] and Anna C. Yaros[1]

[1]Community Health Research Division, RTI International, Research Triangle Park, NC, USA, [2]Center for Youth Development and Intervention, University of Alabama, Tuscaloosa, AL, USA and [3]School of Education and Human Development, University of Virginia, Charlottesville, VA, USA

Abstract

A large body of research demonstrates positive impacts of the Coping Power Program as a preventive intervention for youth behavioral outcomes, but potential collateral effects for caregivers is less known. The current study examined whether the youth-focused Coping Power Program can have a secondary impact on caregiver self-reported symptoms of depression and in turn result in longer-term impacts on child disruptive behavior problems including aggression, conduct problems and hyperactivity. Data from 360 youth/caregiver pairs across 8 waves of data (grades 4 through 10) were analyzed. We used two methodological approaches to (a) assess indirect effects in the presence of potential bidirectionality using timepoint-to-timepoint dynamic effects under Autoregressive Latent Trajectory modeling and (b) estimate scale scores in the presence of measurement non-invariance. Results showed that individually delivered Coping Power (ICP) produced greater direct effects on conduct problems and indirect effects on general externalizing and hyperactivity (through reductions in caregiver self-reported symptoms of depression), compared to group Coping Power (GCP). In comparison to GCP, ICP produced similar direct effects on reductions in caregiver depression. Child-focused prevention interventions can have an indirect impact on caregiver depression, which later shows improvements in longer-term reductions for child disruptive problems.

Keywords: Coping Power Program; caregiver depression; disruptive behavior problems

(Received 1 November 2022; revised 12 October 2023; accepted 13 October 2023)

Coping Power (CP) is a widely used and established evidence based cognitive-behavioral school-based program for working with at-risk youth from elementary to high school. Evidence from the past 25 years support the use of CP for targeting externalizing behaviors in selective populations of youth with elevated aggressive, disruptive behavior. CP can be administered in different modalities including individual, group, and virtual/internet formats, and with child and parent or a child-only focus, each showing efficacy (Lochman et al., 2019). Moreover, there is evidence that the group format does not have the iatrogenic effects previously associated with group interventions for at-risk youth with elevated aggression and disruptive behaviors. However, studies examining the relative impact of group compared to individual modality favors the individual format (Lochman et al., 2015).

Intervention effects on child behavior and caregiver depression

There is growing interest in the identification of interventions that focus on reducing child disruptive behavior problems but also yield benefits for caregiver mental health. For example, there is evidence that cognitive behavioral family intervention reduces maternal depression and concurrently child disruptive behavior (Sanders & McFarland, 2000), and a behavioral parent training intervention delivered to parents of toddlers (Family Check-Up) has led to reductions in parents' depression and produced indirect effects through reduced maternal depression on children's subsequent externalizing (Shaw et al., 2009) and internalizing (Reuben et al., 2015; Shaw et al., 2009, 2016) behaviors. Other studies have reported improvements in posttreatment caregiver depression following modular intervention that included child psychosocial intervention plus a parenting intervention (ages 6–11) and found reductions in child behavior problems were maintained at a 3-year follow-up (but no linkage between change in maternal depression and later change in child behavior) (Shaffer et al., 2013). Much of the work to date demonstrating collateral effects has included interventions that provide simultaneous programming related to parent–child relationships and parents' depressed mood. These skills-based interventions typically involve direct parent training focused on reducing disruptive behaviors. There is also evidence that improvements in child and adolescent behavior problems following multisystemic therapy are associated with decreases in maternal depressive symptoms (Grimbos & Granic, 2009). It should also be noted that all these interventions had a specific parental component (Furlong et al., 2012).

Corresponding author: Lissette M. Saavedra; Email: lsaavedra@rti.org

Cite this article: Saavedra, L. M., Lochman, J. E., Morgan-López, A. A., McDaniel, H. L., Bradshaw, C. P., Powell, N. P., Qu, L., Budavari, A., & Yaros, A. C. (2024). Collateral effects of Coping Power on caregiver symptoms of depression and long-term changes in child behavior. *Development and Psychopathology*, 1–13, https://doi.org/10.1017/S095457942300144X

School-based interventions that use an individual approach to address child disruptive behavior problems also appear to have the potential to yield positive impacts for caregivers; importantly, these programs do so without intervening with caregivers (usually mothers) (Boyd & Gillham, 2009; Powell et al., 2011). For example, there is emerging evidence that interventions focused on child emotion regulation, perspective-taking, and social problem-solving (Lochman & Grescham, 2008) can influence parental stress as well as caregiver mental health in general but depression specifically. Although focused on reducing child disruptive behaviors, components of childhood preventive interventions, like CP, may indirectly reduce maternal/caregiver depression by enhancing child skills and child behavior patterns. It is hypothesized that due to the potential bidirectional nature of the parent–child relationship, child-focused interventions may in turn influence other aspects of the child's life, including family functioning and caregiver depression.

Conversely, studies that have *directly* attended to maternal functioning and show improvements in child and adolescent behavior problems following intervention also show improvement in maternal functioning, and this improvement is associated with decreases in maternal depressive symptoms (Bagner & Eyberg, 2003; Grimbos & Granic, 2009). It is hypothesized that the individual approach of working one-on-one with a child can permit the therapist–child relationship to serve as a model for communication and other relationship skills that may generalize to the child's interpersonal schemas and relationships with other adults, including their parents (Lochman et al., 2015, 2019). Taken together, these studies suggest a complex and potentially bidirectional association between child and caregiver mental health, which may include collateral effects of youth-focused programming on maternal mental health. But the directionality of these associations is not well understood. Moreover, much of the work to date has used traditional measurement approaches (e.g., use of total scores for outcomes) and analytic approaches, such as cross-lagged panel models (CLPMs), without considering measurement non-invariance (MNI). Recent methodological advances that help us specifically attend to both issues are needed to elucidate the direction and impact child-focused interventions have on caregiver outcomes.

The current study

As summarized above, the full Coping Power Program is effective in reducing aggressive children's aggressive and general externalizing behaviors through post-intervention, and 1- and 3-year follow-ups (Lochman & Wells, 2002, 2003, 2004; Lochman et al., 2013, 2014). Despite these favorable outcomes, little is known about the potential collateral effects for their caregivers. Therefore, we sought to explore whether CP had a secondary impact on caregiver depression and in turn resulting in longer-term impacts on child disruptive behavior problems including aggression, conduct problems, and hyperactivity. We further sought to better understand the potential direction of these collateral effects by exploring whether addressing childhood behavior problems via a youth-focused cognitive-behavioral intervention implemented at school can have a secondary impact on maternal depression (at the end of grade 4). We were also interested in whether these direct effects on maternal depression in turn contributed to longer-term impacts on child disruptive behavior problems (at the end of grade 10). We had the unique opportunity in this study to leverage data from a randomized controlled trial testing (only) the child component of the Coping Power Program, delivered in two different formats: the traditional group Coping Power (GCP) vs. the individual Coping Power (ICP), where the program was delivered individually. In both cases, the program was delivered by a trained clinician to fourth graders. Data from this same randomized controlled trial previously demonstrated that the individual format of CP produced stronger reductions in teacher-rated youth externalizing behavior problems than does the group format in a short-term follow-up period (Lochman et al., 2015).

To examine our primary research aims we first assessed the indirect effects in the presence of potential bidirectionality using timepoint-to-timepoint dynamic effects under the Autoregressive Latent Trajectory (ALT) modeling framework (Bollen & Curran, 2004), which allows for the separation of two simultaneous changes processes: (a) a dynamic process in which the prediction of parallel outcomes at a given timepoint (i.e., maternal depression and child externalizing) are a function of the same variable at previous timepoints (i.e., autoregression) and time-varying prediction of each outcome's influence on the other based on previous timepoints that are a part of the parallel outcome process (i.e., cross-lagged effects) and (b) individual-level trajectories estimated under the latent growth curve framework. The ALT also addresses recent criticisms of the more popular CLPM, which makes an inherent assumption that there is no individual variability in the stability or "trait" of the outcome, which can lead to bias in estimates of "state-like" dynamic processes (Hamaker et al., 2015; King et al., 2018; Littlefield et al., 2021). Here, the interest is in estimating the dynamic interplay of whether ICP (compared to group CP) impacts greater reductions in maternal depression, which in turn, leads to later reductions in child externalizing and/or the impacts greater reductions in child externalizing which, in turn, lead to later reductions in maternal depression.

Second, from a measurement perspective, criticisms of the use of total scores for outcome measures have been long-standing (Campbell, 1960; McNeish & Wolf, 2020), yet their use in outcomes analysis across most-if-not-all domains of psychology and psychiatry persists despite several criticisms (Dorans, 2007; McNeish & Wolf, 2020; Sijtsma, 2009). Most applications treat the creation of sum scores for scaling as simply "a data management problem" (Morgan-López, Hien et al., 2022) but McNeish and Wolf (2020) in particular note that there is a *de facto* psychometric model underlying sum scores: the presumption that each item is an equally weighted indicator of the underlying construct that is measured without error. The "equal weights" assumption is testable as a psychometric model (e.g., confirmatory factor analysis model with equality constraints on factor loadings, one-parameter logistic model in item response theory; Andrich, 1978; Muthén, 1978) and, in the rare cases where it is tested explicitly, often fails to fit psychiatric outcome data (He et al., 2014; Morgan-López et al., 2020).

Therefore, for this second aim, we estimated scale scores in the presence of MNI, also known as differential item functioning (DIF). More specifically, differences in item parameters (i.e., MNI/DIF) can manifest in differences in threshold/difficulty parameters and/or factor loading/discrimination parameters across populations and/or across development and, if not mitigated, can lead to distorted estimates of scale scores that contribute to bias in estimating intervention effects (Brincks et al., 2018; Howe et al., 2019; Morgan-López, McDaniel et al., 2022). Moderated nonlinear factor analysis (MNLFA; Bauer & Hussong, 2009; Bauer, 2017) is quickly becoming the most flexible approach to item parameter estimation and scale score estimation that takes into account MNI/

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

DIF across both sets of item parameters across multiple, simultaneous predictors of MNI/DIF, addressing limitations of other approaches to scale score estimation (e.g., multiple indicator-multiple cause [MIMIC] models, IRT models).

Method

Participants

Fourth-grade teachers from 20 schools completed the Aggressive Behavior Screener (Dodge et al., 1997) on all students in their classrooms to identify those children who were at risk for aggressive behavior and thus eligible for recruitment into the CP intervention. At the first screening gate, children whose scores fell at or above the set 25th percentile cutoff were considered eligible for participation. For each of the 20 schools, six students were recruited from each annual cohort. At the second screening gate, children with average or above average parent ratings on the Behavior Assessment System for Children (BASC) Aggression Scale remained eligible for enrollment. The three annual cohorts resulted in a total sample size of 360 parent–child pairs (see Table 1 for sample information). Random assignment to either the ICP of the GCP condition (no other control condition) was made at the school level. Schools were paired based on size, ethnic distribution, and eligibility for free and reduced meals. One school from each pair was randomly assigned to each condition, resulting in 10 schools in the ICP condition and 10 schools in the GCP condition. At GCP schools, each intervention unit of six children participated in a CP group together. Students participated in an average of 28.75 sessions (range = 0–34). Students who received ICP had an average of 28.96 sessions (range = 3–34); students who received GCP attended an average of 28.54 sessions (range = 0–34).

Procedure

Preintervention (Time 1) measures were completed with children and parents at the time of enrollment during the spring semester of students' fourth-grade year. The CP intervention was delivered during the end of fourth grade and throughout fifth grade. Mid-intervention assessments (Time 2) occurred in the summer after fourth grade, post-intervention assessments (Time 3) occurred in the summer after fifth grade, and 1-year follow-up assessments (Time 4) took place during the summer after students completed sixth grade. Assessments at Times 5–8 took place each summer thereafter through the completion of tenth grade for most participants. Children and parents were interviewed separately, typically in their homes, by research staff. The study was approved by the university's Institutional Review Board.

Coping Power intervention

Because this study focuses on the effects of the delivery format of the CP child intervention, only 32 sessions of the CP child component (Lochman et al., 2008) were offered, not the parent component. Children in GCP participated in small groups of five to six children, with two coleaders, and group sessions were 50–60 min. Children in GCP were exposed to all the same intervention content as children in ICP, as well as some additional group-level activities (e.g., role-playing with peers, generating a group name or cheer, opportunity to earn group-level points and rewards). In addition to the group sessions, children in GCP also received the brief monthly individual sessions typically included in the CP intervention. GCP leaders did not receive explicit training about deviancy training in groups, but they did receive routine training in setting and enforcing group behavior rules. Children in ICP met with a CP leader individually for each of the 32 planned 30-min sessions. Children in ICP completed role-plays and other interactive activities with their CP leader rather than with their peers.

Intervention fidelity and quality

Each leader served a similar number of GCP and ICP participants. To ensure high fidelity of implementation, two doctoral-level psychologists who had substantial experience implementing CP met with the interventionists weekly to monitor and provide feedback on program implementation. The interventionists also received detailed supervisory feedback on video-recorded GCP and ICP sessions on a monthly basis to ensure that program implementation remained consistent. GCP leaders and ICP leaders rated that they completely or partially completed 91.07% and 86.43% of objectives, respectively.

Measures

Demographic Variables

Demographic variables were used as predictors of MNI/DIF as well as covariates in the longitudinal model. Demographic variables utilized in the current study include: student gender (1 = male; 0 = female), student race (1 = African American; 0 = Other race), student age (in years), family income level, and repeated grade status (1 = repeated a grade at least once; 0 = did not repeat a grade).

Teacher-Reported Externalizing via the BASC

The study used the 37 items from the Conduct Problems, Aggression, and Hyperactivity subdimensions of the Externalizing factor from teachers' ratings on the BASC (Kamphaus et al., 1999; Reynolds & Kamphaus, 1992), which has demonstrated strong reliability under Classical Test Theory (Cronbach's alpha of .80–.89) and construct validity. All items are rated on a four-point scale ranging from "Never" to "Almost Always". This instrument was administered in each of the timepoints (T1, T3–T8).

Caregiver Depression

The Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) was administered to caregivers to assess their own levels of depression symptoms. The BDI consists of 21 items that are rated on a 4-point (0–3) scale in which adults rate levels of depressive symptoms. Scores obtained on the BDI have been found to correlate significantly with clinicians' ratings of depression as well as objective behavioral measures of depression (Beck et al., 1988). We also recorded the BDI respondent's relationship with the target child as a dummy code (i.e., 1 = biological mother; 0 = Other). This instrument was administered in each of the timepoints (T1–T8).

Analysis plan

The initial set of analyses involve tests of dimensionality, first using the existing factor structures in the literature for (a) the BDI items (single factor) and (b) the BASC Externalizing items (3 factors), conducted using means-and-variance-adjusted weighted least squares estimation for categorical indicators in Mplus (Version 8; Muthén & Muthén, 1998–2017). For the BDI, a single factor model was fit, in addition to a restricted factor model with equality constraints on factor loadings; the latter model was fit per the recommendations of McNeish and Wolf (2020) to formally test

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

Table 1. Descriptives at baseline

	GCP		ICP		
Variable	M	SD	M	SD	p-Value
Age	10.20	0.51	10.24	0.49	0.43
Gender	0.69		0.63		0.23
Race/ethnicity					0.48
Black	82.4%		76.7%		
White	14.1%		18.7%		
Other	3.5%		4.7%		
Family income					0.98
None	4.9%		4.0%		
<$15,000	24.7%		24.5%		
$15,000–$29,999	33.3%		31.8%		
$30,000–$49,999	21.8%		23.2%		
>$50,000	15.5%		16.6%		
BDI reporter relationship to child					
Biological mother	87.8%		87.8%		
BDI MNLFA scores	−0.05	0.84	−0.02	0.86	0.71
BASC MNLFA scores					
General externalizing	−0.01	0.86	0.15	0.77	0.3
Aggressive behavior	−0.1	0.82	0.07	0.87	0.3
Hyperactivity	0.08	0.92	−0.06	0.88	0.22
Conduct problems	−0.11	0.8	0.12	0.84	0.09

Notes: GCP = group Coping Power. ICP = individual Coping Power; *M* = mean; SD = standard deviation; BDI = Beck Depression Inventory; MNLFA = moderated nonlinear factor analysis; BASC= Behavior Assessment System for Children.

whether the psychometric model that is assumed when using total scores actually fits the data (He et al., 2014; Morgan-López et al., 2020; Morgan-López, Hien et al. 2022). For the BASC, a series of factor models were fit to the data, as described in more detail below, to discern the best model-data fit. We aimed to retain the three-factor structure that was suggested by the BASC developers (Reynolds & Kamphaus, 2004).

Next, for the caregiver depression and child externalizing constructs, a series of MNLFA models were fit under Mplus robust maximum likelihood (with standard error adjustment for school-level clustering), examining MNI/DIF in separate models for each item, such that each set of demographic variables (and dummy variables coding timepoint) was assessed for whether they contributed to significant MNI/DIF on each item (above-and-beyond their effects on "true" latent underlying caregiver depression and externalizing dimensions, respectively). The MNI/DIF parameters that were significant at $p \leq .001$ were then retained for a global MNLFA model. Of those that remained significant in the global model, these parameters were retained for the final MNLFA scale score estimation model. From the final MNLFA models, the MNI/DIF parameters that were significant are shown in Table 2 for the BDI and Table 3 for the BASC. Predictors of underlying severity are described in the *Results* section. Scale scores estimated under MNLFA were then output using the Mplus SAVEDATA command for later analysis. Additional detail on MNLFA measurement modeling and scale score estimation is reported in Bauer (2017).

Estimating dynamic mediation effects via autoregressive latent trait (ALT) modeling

Dynamic mediation models (Cole & Maxwell, 2003; Maxwell & Cole, 2007) have traditionally been estimated under the CLPM framework. More recently, criticisms of the CLPM have emerged in that CLPMs make an inherent assumption that there is no individual variability in the stability or "trait" of the outcome, which can lead to bias in estimates of "state-like" dynamic processes (Hamaker et al., 2015; King et al., 2018; Littlefield et al., 2021). Models that address this state versus trait dichotomy include the random intercept CLPM (RI-CLPM; Hamaker et al., 2015) and the ALT model, which adds a random individual-level slope-over-time to the RI-CLPM (Bollen & Curran, 2004). The ALT model, fit under robust ML (for school-level cluster-adjusted standard errors and missingness under the missing-at-random assumption) is a hybrid model that synthesizes two longitudinal modeling traditions: (a) a dynamic process in which the prediction of an outcome at a given timepoint is a function of the same variable at previous timepoints (i.e., autoregression) and time-varying predictor(s) at previous timepoints that are a part of a parallel outcome process (i.e., cross-lagged effects) and (b) individual-level trajectories estimated under the latent growth curve framework. The ALT model provides a framework for estimation of the dynamic interplay by which ICP (compared to GCP) on externalizing (i.e., general externalizing and/or externalizing subdimensions) are mediated by "state" shifts in caregiver depression and vice versa (i.e., dynamic mediation) separate and apart from interindividual or "trait" differences in the trajectories of child externalizing and caregiver depression over time.

Results

Preliminary tests of model fit

BDI Caregiver Depression

The model for the conventional test for unidimensionality of the BDI fit adequately, with results showing that a single factor underlies the items, comparative fit index (CFI) = .99, root mean square error of approximation (RMSEA) = .060, 95% CI [.058, .063], meeting the standard for essential unidimensionality (Millsap & Kwok, 2004). The total score "analog" (TSA) model, where factor loadings were constrained to equality, failed to fit the data in absolute terms (e.g., RMSEA = .109 [.107, .111]) and, more importantly, comparatively against the conventional single factor model ($\Delta\chi^2(20) = 2308.724$, $p < .0001$). Thus, total scores for the BDI would be highly biased and mischaracterize caregiver depression severity.

BASC Externalizing

The model test for unidimensionality of the items on the BASC externalizing scales (i.e., all items on the BASC Aggression, Conduct Problems, and Hyperactivity scales loading on a single factor) suggested that this model did not adequately fit the data (RMSEA = .10, 90% CI [.100, .104]; CFI = .89). Next, we estimated a correlated, three-factor model where each set of items originally designated to correspond to the Aggression, Conduct, and Hyperactivity scales loaded on each respective scale. Fit indices suggest that estimation of this model did not yield adequate model-data fit (RMSEA = .09, 90% CI [.084, .088]; CFI = .92). Finally, we estimated a bifactor model where all items loaded on a general factor, theoretically representing externalizing problems, as well as specific factors that corresponded to the BASC developers' original

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

Table 2. Final BDI MNLFA item parameters

Item	Factor loading (λ)	Threshold (τ; 0 to 1)	Threshold (1 to 2)	Threshold (2 to 3)	Gender λ DIF	Race λ DIF	Income λ DIF	Gender τ DIF	Race τ DIF	Income τ DIF
Sadness	1.97	2.04	4.73	5.85						
Pessimism	1.71	2.21	5.29	6.09						
Past failure	2.45	2.97	5.28	8.10					−0.53	
Loss of satisfaction	1.82	0.27	4.34	5.82			0.06			
Guilt	1.90	1.90	4.40	6.92						
Punishment	2.01	2.34	4.00	4.15						
Self-dislike	2.32	2.18	6.47	8.70						
Self-criticism	1.52	0.89	3.23	4.66						
Suicidal thoughts	2.11	4.46	8.78	X						
Crying	1.56	1.31	3.03	3.29						
Agitation	1.23	−0.22	2.41	2.75						
Loss of interest	1.55	0.74	3.49	5.76					0.38	
Indecision	1.83	1.66	4.16	7.45						
Feelings of worthlessness	1.74	1.55	3.86	5.60		0.78			−1.49	0.11
Loss of energy	1.40	0.86	2.86	5.40	−0.51			0.37		
Changes in sleep	1.36	−0.30	1.76	3.02						
Irritability	1.58	−0.59	3.03	4.74						0.04
Changes in appetite	1.39	0.78	2.94	4.52					0.52	−0.05
Concentration problems	0.29	0.88	1.80	2.38						
Tiredness	1.47	0.46	3.28	4.90			−0.06			
Low libido	1.97	0.62	2.34	3.43			0.11			−0.04

scales of Aggression, Conduct Problems, and Hyperactivity, which yielded adequate fit (RMSEA = .07, 90% CI [.065, .068]; CFI = .95). This model fit the data significantly better than the single factor ($\Delta\chi^2$ (40) = 2924.698, $p < .0001$) and three-factor models ($\Delta\chi^2$ (37) = 1742.068, $p < .0001$). Some of the loadings in this solution demonstrated were nonsignificant or demonstrated nonsignificant and/or small loadings on either the specific (i.e., Aggression, Hyperactivity, Conduct Problems) or general factor (i.e., Externalizing Problems) (See Supplemental materials). We estimated a final model in which these nonsignificant and/or small factor loadings (i.e., standardized factor loadings < .2) were constrained to zero. This model also yielded adequate model-fit (RMSEA = .06, 90% CI [.062, .066]; CFI = .96). The TSA model, where factor loadings were constrained to equality on each specific factor and the general factor, did not yield adequate model-data fit (e.g., RMSEA = .10 [.097, .101]) and, more importantly, comparatively against the conventional bifactor model ($\Delta\chi^2(59) = 2893.958$, $p < .0001$).

MNLFA

BDI Caregiver Depression

To model MNI/DIF across demographic variables (e.g., gender, race/ethnicity, income) in the externalizing scale scores, MNLFA models were fit, with results presented in Table 2. The item parameters from the final BDI MNLFA model are presented in Table 2, with all demographic predictors (i.e., gender, race, family income, age, repeating a grade) of MNI/DIF and severity centered so that all comparisons are against the sample average values. Time was centered at baseline. Given the predictors of MNI/DIF examined, 11 of the 21 BDI items showed no DIF across any predictor (i.e., "empirical anchor" symptoms). For the thresholds/difficulty parameters, key predictors that showed statistically significant MNI/DIF on at least 3 items included race and family income. For the factor loading/discrimination parameters, key predictors that showed statistically significant MNI/DIF (i.e., variation in item "weights") included race, gender, and family income. African American caregivers demonstrated significantly greater-than-average thresholds for Items 12 and 18 ("Loss of Interest", "Changes in Appetite") and lower-than-average threshold for Item 3 ("Past Failure"). A larger-than-average loading/discrimination parameter was observed for African American caregivers for Item 14 ("Feelings of Worthlessness").

BASC Externalizing

Item parameters from the final MNLFA model are presented in Tables 3 and 4. Of the 37 externalizing items, 21 showed no DIF across any predictor (i.e., "empirical anchor" symptoms). For the thresholds/difficulty parameters (Table 4), key predictors that showed statistically significant MNI/DIF on at least 4 items (compared to the overall sample average thresholds/difficulty parameters in the three "Threshold" columns) included gender and time. For the factor loading/discrimination parameters (Table 3), key predictors that showed statistically significant MNI/DIF (i.e., variation in item "weights") included race and

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

Table 3. Final BASC MNLFA parameters: loading DIF

Specific factor/item	General factor loading (λ)	Specific factor loading (λ)	Gender	Race	Age	Income	Repeat grade	T3	T4	T5	T6	T7	T8
Conduct													
Has to stay after school for punishment	1.11	0.58						0.63					
Steals/steals at school	1.09	0.77											
Cheats in school	1.11	1.03											
Uses foul language	1.47	1.62											
Shows a lack of concern for others' feelings	1.71	0.69							0.58				
Skips classes at school	0.00	2.13											
Complains about police or other law enforcement officers	1.07	1.33											
Is truant	0.52	1.16											
Has been suspended from school	1.29	1.58											
Has friends who are in trouble	1.18	0.80											
Aggression													
Argues when denied own way	2.18	0.53											
Threatens to hurt others	2.14	2.19											
Blames others	2.21	0.85											
Bullies others	2.48	2.09											
Breaks other children's things	1.66	0.89											
Talks back to teachers	2.25	0.88											
Orders others around	2.01	0.96		0.56									
Is critical of others	2.22	0.75											
Calls other adolescents names	2.43	1.68											
Shows off	1.95	0.00											
Teases others	2.85	1.77											
Complains about rules	1.94	0.51											
Hits other adolescents	1.75	1.66											
Is a "sore loser"	1.78	0.56											
Hyperactivity													
Rushes through assigned work	2.21	3.78											
Bothers other children when they are working	2.46	0.50											
Talks too loud	2.18	0.00											
Seeks attention while doing schoolwork	2.19	0.00											
Taps foot or pencil	1.13	0.00											
Acts without thinking	2.35	0.62											
Calls out in class	2.74	0.00											
Interrupts others when they are speaking	2.90	0.00											
Makes loud noises when playing	1.62	0.00											
Hurries through assignments	1.69	2.62											
Acts silly	1.34	0.00											
Is overly active	2.00	0.00											
Cannot wait to take turn	2.11	0.00											

Note. BASC = Behavior Assessment System for Children; MNLFA = moderated nonlinear factor analysis; DIF = differential item functioning.

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

Table 4. Final BASC MNLFA parameters: threshold DIF

Specific factor/item	Threshold (τ; 0 to 1)	Threshold (τ; 1 to 2)	Threshold (τ; 2 to 3)	Gender	Race	Age	Income	Repeat grade	T3	T4	T5	T6	T7	T8
Conduct														
Has to stay after school for punishment	0.61	3.18	5.42							1.07				
Steals/steals at school	1.80	3.89	5.37											
Cheats in school	0.37	2.97	4.79											
Uses foul language	−0.06	3.02	5.39									0.68		1.14
Shows a lack of concern for others' feelings	−1.88	1.40	3.32											
Skips classes at school	3.65	5.65	7.23											
Complains about police or other law enforcement officers	2.98	5.05	6.64											
Is truant	1.53	3.21	4.60											
Has been suspended from school	0.44	3.33	5.56	0.59						1.30				
Has friends who are in trouble	−2.08	0.90	2.95			0.34						−1.30	−0.89	
Aggression														
Argues when denied own way	−3.61	−0.38	1.56											
Threatens to hurt others	−1.21	2.82	5.39											
Blames others	−3.88	−0.57	1.56									−1.56		
Bullies others	−1.95	1.71	4.16											
Breaks other children's things	0.54	3.32	5.13											
Talks back to teachers	−2.79	0.60	2.53											1.26
Orders others around	−1.78	1.14	3.53	−1.47			0.08							
Is critical of others	−2.38	1.06	3.32	−0.81										
Calls other adolescents names	−3.84	0.37	3.26											
Shows off	−1.87	0.68	2.83											
Teases others	−4.27	0.45	3.46											
Complains about rules	−2.03	1.30	3.68									1.61	1.45	1.59
Hits other adolescents	−1.20	2.70	5.28											
Is a "sore loser"	−1.33	1.63	3.77											
Hyperactivity														
Rushes through assigned work	−5.26	0.33	4.05											
Bothers other children when they are working	−3.18	0.51	3.07	0.49										
Talks too loud	−2.13	0.38	2.33	−0.85	0.57								1.14	
Seeks attention while doing schoolwork	−2.28	0.57	2.54											
Taps foot or pencil	−0.66	1.53	3.20	0.68										
Acts without thinking	−3.89	−0.09	2.31					0.53						
Calls out in class	−2.95	0.31	2.58									0.87		
Interrupts others when they are speaking	−3.23	0.85	3.45											
Makes loud noises when playing	−0.19	1.68	3.33											
Hurries through assignments	−4.17	0.26	3.15											
Acts silly	−2.77	0.03	1.81											
Is overly active	−1.23	1.03	2.73	0.43										
Cannot wait to take turn	−1.08	1.94	3.98											

Note. BASC = Behavior Assessment System for Children; MNLFA = moderated nonlinear factor analysis; DIF = differential item functioning.

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

time. African American youth demonstrated a significantly greater-than-average loading/discrimination parameter on the item "orders others around". There were significantly larger-than-average loading/discrimination parameters at time three for the item "Has to stay after school for punishment" and time four for the item, "Shows a lack of concern for others' feelings".

Dynamic mediation under autoregressive latent trajectory models

ICP/GCP Effects on General Externalizing

To address our primary research question, we modeled the dynamic effects between caregiver depression and general externalizing. The overall model for dynamic effects between caregiver depression and general externalizing fit well, CFI = .953, RMSEA = .040, [CI:.029, .049]. All models included gender, race/ethnicity, age, family income and whether the participant had repeated a grade as covariates. All model results are presented independent of the growth modeling portion of the ALT models; growth modeling results with this sample have been presented elsewhere (Lochman et al., 2015). Significant first-order autoregression was observed for both caregiver depression ($b = .140$ (.044), $t = 3.18$, $p = .001$) and general externalizing ($b = .135$ (.063), $t = 2.13$, $p = .033$). Dynamic effects on caregiver depression were observed favoring ICP (compared to GCP) ($b = -.220$ (.036), $t = -6.08$, $p < .001$) but no ICP/GCP differences were observed on general externalizing ($p = .20$). Reductions in caregiver depression at time t led to reductions in general externalizing in the following year ($b = .113$ (.057), $t = 1.99$, $p = .046$); the reverse effect of general externalizing at time t did not have an impact on caregiver depression the following year ($p = .59$). Formal tests of mediation used the "traditional" estimate for mediation under Asymmetric Confidence Interval estimation (Tofighi & MacKinnon, 2011), given tests for intervention × mediator (caregiver depression) interactions were nonsignificant (*p*s all > .17); these intervention × mediator interaction effects are required for newer potential outcomes mediation analyses (MacKinnon et al., 2020). The tests of whether ICP's effects on general externalizing were mediated by changes in caregiver depression in the previous year were nonsignificant (ab = −.025 [CI: −.053, .001]), despite both the "a" path and "b" path each being significant.

ICP/GCP Dynamic Effects on Hyperactivity

The overall model for dynamic effects between caregiver depression and hyperactivity fit well, CFI = .962, RMSEA = .034 (CI: .022, .044). Significant first-order autoregression was observed for both caregiver depression ($b = .141$ (.045), $t = 3.15$, $p = .002$) and hyperactivity ($b = .100$ (.043), $t = 2.31$, $p = .021$). Dynamic effects on caregiver depression were observed favoring ICP (compared to GCP) ($b = -.223$ (.037), $t = -6.01$, $p < .001$) but no ICP/GCP differences were observed on hyperactivity ($p = .26$). Reductions in caregiver depression at time t led to reductions in hyperactivity in the following year ($b = .152$ (.036), $t = 4.20$, $p < .001$); the reverse effect of hyperactivity at time t did not have an impact on caregiver depression the following year ($p = .27$). The tests of whether ICP's effects on hyperactivity were mediated by changes in caregiver depression in the previous year were significant (ab = −.034 [CI: −.052, −.013]).

ICP/GCP Dynamic Effects on Conduct Problems

The overall model for dynamic effects between caregiver depression and conduct problems fit well, CFI = .947, RMSEA = .044 [CI: .035, .054]. Significant first-order autoregression was observed for both caregiver depression ($b = .138$ (.044), $t = 3.11$, $p = .002$) and conduct problems ($b = .204$ (.063), $t = 3.22$, $p = .002$). Dynamic effects on caregiver depression were observed favoring ICP (compared to GCP) ($b = -.225$ (.036), $t = -6.25$, $p < .001$) in addition to direct effects favoring ICP on conduct problems ($b = -.525$ (.134), $t = -3.97$, $p < .001$). Neither cross-lagged effect was significant for caregiver depression effects on conduct problems ($p = .14$) nor conduct problem effects on caregiver depression ($p = .31$).

ICP/GCP Dynamic Effects on Aggression

Finally, the overall model for dynamic effects between caregiver depression and aggression fit well, CFI = .945, RMSEA = .044 [CI: .035, .053]. Significant first-order autoregression was observed for both caregiver depression and aggression (*p*s < .003); however, outside of the ICP effects on caregiver depression observed in the other models, no direct nor indirect effects on aggression were observed.

Discussion

The present study examined whether, CP, a youth-focused cognitive-behavioral intervention aiming to reduce childhood behavior problems delivered at school (only) had a secondary impact on maternal depression and, in turn, longer-term impacts on key child disruptive behavior problems (Aggression, Conduct Problems, and Hyperactivity), with a particular interest in whether those collateral effects were stronger when the program was delivered individually or in a group format (Lochman et al., 2015). Further, our approach is unique in that it uses two methodological tools to address improvements in the quality of (a) assessing indirect effects in the presence of potential bidirectionality and (b) scale score estimation in the presence of MNI, also known as DIF.

Effect of group versus individual format on children's conduct problems

Results showed an overall intervention effect on teacher-reported youth conduct problems for both the individual and group Coping Power conditions. Coping Power addresses mechanisms or common elements that are typical in cognitive-behavioral approaches to reducing externalizing behavior problems in youth. A prior study with this sample found that teacher-rated general externalizing behavior decreased through a 1-year follow-up for both the individual and group formats, but the positive effects for the individual format were significantly stronger (Lochman et al., 2015). Using this approach with (a) finer-grain component constructs within the overall externalizing construct (i.e., a bifactor model with a general externalizing factor and three specific factors: hyperactivity, conduct, and aggression) and (b) timepoint-to-timepoint dynamic effects under ALT modeling, current results also showed that the individual condition (ICP) produced greater direct effects on conduct problems, and indirect effects on general externalizing and hyperactivity (through reductions in caregiver depression), compared to the group condition (GCP). However, this was not the case for aggression. For example, the conduct problem items tended to include more disruptive behaviors, such as "uses foul language" and "has been suspended from school," whereas the aggressive behaviors involved more immediate interpersonal behaviors; this suggests that CP problem-solving training, and especially anticipation of longer-term consequences of choices, may be weaker in GCP than in ICP. This could have occurred because children may not have learned the concepts as

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

well and were affected by negative or deviant-reinforcing peer interactions within GCP.

It is also possible that in GCP, children may not attend as frequently, or learn the problem-solving skills as well, due to distraction and emotional dysregulation caused by peers in group. It should be noted, based on fidelity data, the intervention objectives and activities were delivered adequately in both conditions by the facilitators. It is possible that youth in GCP may not have attended as well to the material. Also, highly impulsive verbal and physical behaviors from an individual in an intervention group can interfere with the group leaders' ability to effectively lead activities and discussions, and with the ability of other group members to selectively attend to the relevant program material. Off-task interactive behaviors among other group members can have the same effect of disrupting children's ability to bring new social-emotional skills into their repertoire of problem solutions in their working memory.

In addition, in GCP, some peers may have overvalued short-term positive consequences for negative behavior, and through deviancy training, modeled and reinforced rule-breaking and conduct problems for the target child. Research suggests that deviancy training is a core mechanism that has been theorized to produce increased risk for adolescent problem behaviors (Dishion & Patterson, 2016) and negative outcome effects of some group interventions (Dishion & Andrews, 1995; Poulin et al., 2001). In this case, children may learn more adaptive and prosocial skills in the group but have greater incentive to enact the peer-reinforced negative behaviors. The risk of negative peer behaviors to enhance individual children's teacher-rated conduct problems within group interventions can be reduced, in a protective manner, when group therapists have higher rates of clinical skills (are more positive and flexible, and less angry, irritable and frustrated [Lochman et al., 2017]), but the potentially iatrogenic effect of peer reinforcement of negative child behavior in group can reduce children's motivation to acquire and use proactive and competent coping skills (Dodge et al., 2007). These complicated peer effects do not occur in individual therapy, simplifying the therapist's ability to present and practice alternative coping skills.

Effects of individual intervention format on caregiver depression and children's subsequent hyperactive behaviors

Individual intervention effects on caregiver depressive symptoms

In addition to its direct effects on children's reduction of conduct problems, the ICP condition, in comparison to GCP, produced similar direct effects on reductions in caregiver depression at the second timepoint (early intervention). It is notable that these effects were evident with the use of only the child component of CP; inclusion of the parent component of CP could enhance these child-only intervention effects on caretaker depression and could be a useful future area of research. Our conceptual model of the child component of CP, with its focus on emotion regulation and social-cognitive processing, predicted that CP would affect conduct problems (Lochman & Wells, 2002a). As noted above, the larger impacts of the individual version of CP on youth externalizing behavior than the group format were consistent with expectations and our theoretical model (Lochman et al., 2015, 2019). However, the ICP effect on caregiver depression was not explained by our conceptual model of the child component of the program. Moreover, the reductions in caregiver depression predicted greater decreases in children's hyperactivity at a subsequent timepoint, suggesting that changes in caregiver depression mediated changes in children's hyperactive behaviors.

There are several possible reasons why ICP has this beneficial effect on caregiver depression, and as well as possible reasons why changes in caregiver depression can then affect changes in hyperactive behaviors. For example, there is a robust literature that links caregiver depression and child aggression and conduct problems (Barry et al., 2005, 2018; Connell & Goodman, 2002; Singh et al., 2011). These associations have been found between children with externalizing problems and both with caregivers with chronic and severe depression (van der Waerden et al., 2015) and with caregivers with subclinical levels of depression (Barry et al., 2005; Conners-Burrow et al., 2015). Caregiver depression has been found to predict to subsequent externalizing behavior (Blatt-Eisengart et al., 2009), and, reciprocally in the other temporal direction, children's frequent problem behaviors can create more parental distress over time (Neece et al., 2012).

In the current study, as children's symptoms improved, according to their teacher-reported conduct problems through their involvement in the ICP program, caregivers may experience less distress and have fewer depressive symptoms because the caregivers experience less parenting stress (Morgan et al., 2002). As children experienced improvements in their behavior in school due to intervention, they are likely to improve in their academic progress (Lochman et al., 2012). Parents may in turn begin to receive better feedback from schools about their children's behavioral and academic progress, which in turn can lead them to feel more hopeful and less despondent over their children's behavior. In addition, children's behavioral improvement at school can generalize to their home behavior (e.g., Lochman et al., 2015), thus reducing the caregivers' time spent in disciplinary actions and their associated parenting stress.

Caregivers' social cognitions may also play a role. Depressed mothers have more negative causal attributions of children's misbehavior (White & Barrowclough, 1998), but as their depression declines due to intervention, their negative attributions about their children's behavior have been found to decline (Novick et al., 2022). As children's conduct problems decrease, caregivers may develop more benign and less hostile attributions of their children's behavior (Dix & Lochman, 1990), and as their hostile attributions decline, they in turn may become less depressed.

At the level of therapeutic process, children's involvement in ICP may lead to stronger therapeutic alliance with the therapist than might be the case in the group format of the intervention. A positive therapeutic working relationship has been found to be a critical part of the therapeutic process, across a broad range of interpersonal, psychodynamic and cognitive-based intervention models (Accurso et al., 2013). Both adolescent self-reports and clinician-rated therapeutic alliance have been found to be strongly associated with a reduction in posttraumatic stress symptoms (Ormhaug et al., 2014) and internalizing symptoms (Zorzella et al., 2015).

A recent series of studies of studies on the ICP condition, using the Therapeutic Process Observational Coding System (TPOCS-A; McLeod & Weisz, 2005) to code for therapeutic alliance, = indicated that strong emotional bonding between youth and therapist in the early sessions of ICP, predicted a reduction in teacher-reported externalizing behaviors (Mitchell et al., 2021). Moreover, positive, warm, non-irritable therapist behaviors have predicted greater improvement in teacher-rated externalizing behaviors in group-delivered CP (Lochman et al., 2017, 2019), suggesting that therapeutic alliance is also important for group intervention. However, it is not known if the actual observed

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

therapeutic alliance varies between the individual versus group forms of the intervention. The one-on-one format of the intervention may permit greater trust and emotional bonding to occur, resulting in corrective emotional experiences and stronger emotion regulation (Castonguay & Hill, 2012).

Experiencing stronger alliance and bonding with the therapist may begin to alter children's interpersonal schemas for their interactions with adults in general in their lives. Children's schemas and associated expectations for others' behavior and the outcomes of their own behavior can affect children's cognitive processing and their subsequent behavior, in a manner similar to internalized representations and working models in attachment therapy (Lochman & Dodge, 1998; Lochman & Lenhart, 1995). Working models are derived from experience with caregivers (Bretherton, 1995), and are open to change depending on subsequent interpersonal relationships (Mash & Dozois, 1996). However, because schemas are conservative and limit the encoding of new information because of preexisting beliefs (Fiske & Taylor, 1984), strong stable new relationships are necessary to influence schema change. In the case of ICP, the children's positive constructive relationship with the adult therapist can lead to modifications of their schemas about adults, including how they perceive, and what they expect from, their caregiver. If the children begin to alter their schemas about adults, and this generalizes to their perceptions and expectations of their caregiver, this has the potential to enhance their attachment to the caregiver and change the nature of their interactions with caregivers in a positive way. Such a change could affect the caregivers' attachment relationship to the child, and in turn contribute to more positive interactions; furthermore, this may have decreased caregiver distress and depressive symptoms and results in more positive emotions when interacting with their children.

Effects of caregiver depression on children's hyperactive behaviors

The caregivers of children with attention-deficit/hyperactivity disorder and hyperactive behavior problems have higher levels of stress and distress, even though it can be subclinical (Hudec & Mikami, 2017; Gerdes et al., 2021). The effect of caregiver depression on externalizing behavior in general is partially mediated by inconsistent parenting (Barry et al., 2009). Similar parenting difficulties have been noted for children with hyperactive behaviors, as parents have been found to be more negative and less supportive, responsive and engaged with overactive children (Chronis et al. 2007; Freitag et al., 2012). In the current study, as caregivers experienced fewer depressive symptoms, they may have become more engaged and supportive of their children's behavior and progress, and better able to consistently reinforce and consequate their behavior at home and, according to school report they receive, at school. As a result, children may have become less impulsive and uninhibited, more emotionally regulated, and better able to follow adult directions and focus their attention in the school setting. It should be noted that the Hyperactivity specific construct, as measured in the current study, consisted of only four indicator items (see Table 3). Items loading on the specific factor seemed to be specifically related to rushing through schoolwork (e.g., "Rushes through assigned work" demonstrated the highest factor loading). As such, these results predominately reflect schoolwork related hyperactivity.

Strengths and limitations

This study had several strengths, including the longitudinal design, which allowed examination of intervention condition specific effects. Along with the multiple timepoints, our unique approach gave us the opportunity to examine the bidirectional relation and assess indirect effects and scale score estimation in the presence of MNI. Second, care was taken in reporter selection as teachers were used as the main child behavior problem reporter. This is important particularly for the setting of interest- prevention in schools. However, a large body of literature recommends capturing child behavior problems via a multi-informant battery and argues for the use of teacher ratings of students in lieu of maternal/caregiver ratings. This is in part consistent with evidence across studies that maternal depression can distort maternal reporting of child behaviors leading to over-reporting of child behaviors.

Despite these strengths, several limitations should be considered. It is important to note that this was a school prevention sample of youth with elevated ratings of aggression and related problem behaviors. Therefore, caregivers in this study varied in extent of depressive symptoms, and most of them were not within the clinical limits. Generalizability of these results is limited to such samples and care should be taken when considering caregivers who report clinical levels of depression. Also, we collapsed all caregivers into one category, the majority being mothers. While the randomized design was a strength, we did not have a control condition to contrast to the two active intervention conditions. Additionally, the hyperactivity specific factor only had four indicator items, perhaps narrowly representing hyperactivity. Future research replicating these findings should include measures of hyperactivity more broadly operationalized. Despite these limitations, the current results provide encouraging evidence for the potential collateral impact that individual-focused prevention interventions can have on caregiver depression which later show improvements in longer-term reductions for child disruptive problems, particularly conduct and attention-deficit/hyperactivity disorder-related behaviors.

Implications of methodological advances for developmental psychopathology

There are several implications of these findings for prevention intervention research on problem behaviors for youth and adolescence and for the broader area of developmental psychopathology. For example, the use of bifactor MNLFA estimation and scale scoring (Bauer, 2017; Eid et al., 2016) yielded two features that would be of broad interest to developmental psychopathologists. The explicit testing of the psychometric model suggested a single externalizing dimension did not fit the data, while a bifactor MNLFA model fit the BASC externalizing items optimally. While it is not particularly surprising that a 1-factor model did not fit once explicitly tested, the total BASC externalizing score is what is commonly used in practice. The total score essentially assumes a 1-factor model with all factor loadings equated (Andrich, 1978; McNeish & Wolf, 2020), would fit the data worse, and would be indicative of biased scale scores. What would have been lost by using a score reflecting a single dimension, independent of the measurement bias in a single externalizing score, is the fact that the externalizing subdimensions each had very different pathways by which ICP achieved its effects.

A related point is the fact that, as part of MNLFA scoring, DIF across multiple variables (e.g., race, SES, developmental period) was observed that may be of scientific interest in and of itself. It has been noted for a long time that using unweighted total scores will produce biased estimates of an underlying construct if DIF is present but unmodeled (Campbell, 1960; McNeish & Wolf, 2020)

https://doi.org/10.1017/S095457942300144X Published online by Cambridge University Press

with particular implications for bias across developmental periods (Sandler et al., 2010) and race/ethnicity (Bettencourt et al., 2022). Different item parameters across race, SES and/or time in the MNLFA models also reflects variation in the relevance of some item content across populations and across development. In many ways, this measurement model, and subsequent externalizing scale score estimates, properly captures the intended spirit of reflecting changes in the relevance of item content across development within the major developmental psychopathology assessment platforms (ASEBA, Achenbach, 1966; BASC, Reynolds & Kamphaus, 2004).

Taken together, these findings provide promise for the potential collateral effectives of preventive interventions for caregivers, as well as the youth most directly impacted by the program. Methodological approaches such as those leveraged in the current study should be considered for future studies of youth-focused interventions to better understand the broader impact of the programming for children and their caregivers.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S095457942300144X.

Funding statement. The work presented in this paper was sponsored by the National Institute of Mental Health grant R01MH124438 (Antonio A. Morgan-López & John E. Lochman, MPIs).

Competing interests. The authors report no conflicts.

References

Accurso, E. C., Hawley, K. M., & Garland, A. F. (2013). Psychometric properties of the therapeutic alliance scale for caregivers and parents. *Psychological Assessment*, *25*(1), 244–252.

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied*, *80*(7), 1–37.

Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, *2*(4), 581–594. https://doi.org/10.1177/014662167800200413

Bagner, D. M., & Eyberg, S. M. (2003). Father involvement in parent training: When does it matter? *Journal of Clinical Child and Adolescent Psychology*, *32*(4), 599–605.

Barry, T. D., Dunlap, S. T., Cotten, S. J., Lochman, J. E., & Wells, K. C. (2005). The influence of maternal stress and distress on disruptive behavior problems in boys. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(3), 265–273. https://doi.org/10.1097/00004583-200503000-00011

Barry, T. D., Dunlap, S. T., Lochman, J. E., & Wells, K. C. (2009). Inconsistent discipline as a mediator between maternal distress and aggression in boys. *Child & Family Behavior Therapy*, *31*(1), 1–19. https://doi.org/10.1080/07317100802701186

Barry, T. D., Lindsey, R. A., Fair, E. C., & DiSabatino, K. M. (2018). Parent psychopathology. In J. E. Lochman, & W. Matthys (Eds.), *The wiley handbook of disruptive and impulse-control disorders* (pp. 275–289). Wiley Blackwell.

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, *22*(3), 507–526. https://doi.org/10.1037/met0000077

Bauer, D. J., & Hussong, A. M. (2009). Psychometric approaches for developing commensurate measures across independent studies: Traditional and new models. *Psychological Methods*, *14*(2), 101–125. https://doi.org/10.1037/a0015583

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck depression inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77–100. https://doi.org/10.1016/0272-7358(88)90050-5

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561–571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Bettencourt, A. F., Musci, R. J., Masyn, K. E., & Farrell, A. D. (2022). Latent classes of aggression and peer victimization: Measurement invariance and differential item functioning across sex, race-ethnicity, cohort, and study site. *Child Development*, *93*(2), e117–e134. https://doi.org/10.1111/cdev.13691

Blatt-Eisengart, I., Drabick, D. A., Monahan, K. C., & Steinberg, L. (2009). Sex differences in the longitudinal relations among family risk factors and childhood externalizing symptoms. *Developmental Psychology*, *45*(2), 491–502. https://doi.org/10.1037/a0014942

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research*, *32*(3), 336–383. https://doi.org/10.1177/0049124103260222

Boyd, R. C., & Gillham, J. E. (2009). Review of interventions for parental depression from toddlerhood to adolescence. *Current Psychiatry Reviews*, *5*(4), 226–235. https://doi.org/10.2174/157340009789542123

Bretherton, I. (1995). A communication perspective on attachment relationships and internal working models. In *Monographs of the society for research in child development* (pp. 310–329). https://doi.org/10.2307/1166187

Brincks, A., Montag, S., Howe, G. W., Huang, S., Siddique, J., Ahn, S., & Brown, C. H. (2018). Addressing methodologic challenges and minimizing threats to validity in synthesizing findings from individual-level data across longitudinal randomized trials. *Prevention Science*, *19*(1), 60–73. https://doi.org/10.1007/s11121-017-0769-1

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist*, *15*(8), 546–553.

Castonguay, L. G., & Hill, C. E. (2012). *Transformation in psychotherapy: Corrective experiences across cognitive behavioral, humanistic, and psychodynamic approaches*. American Psychological Association.

Chronis, A. M., Lahey, B. B., Pelham, W. E.Jr, Williams, S. H.,Baumann, B. L., Kipp, H., & Rathouz, P. J. (2007). Maternal depression and early positive parenting predict future conduct problems in young children with attention-deficit/hyperactivity disorder. *Developmental Psychology*, *43*(1), 70–82. https://doi.org/10.1037/0012-1649.43.1.70

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558–577. https://doi.org/10.1037/0021-843X.112.4.558

Connell, A. M., & Goodman, S. H. (2002). The association between psychopathology in fathers versus mothers and children's internalizing and externalizing behavior problems: A meta- analysis. *Psychological Bulletin*, *128*(5), 746–773. https://doi.org/10.1037/0033-2909.128.5.746

Conners-Burrow, N., McKelvey, L. M., Pemberton, J., Mesman, G., Holmes, K., & Bradley, R. H. (2015). Buffering the negative effects of maternal alcohol problems on child behavior.*Journal of Family Psychology*, *29*, 576–584, https://doi:10

Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology*, *63*(4), 538.

Dishion, T. J., & Patterson, G. R. (2016). The development and ecology of antisocial behavior: Linking etiology, prevention, and treatment. Developmental *Psychopathology*, 1–32.

Dix, T., & Lochman, J. E. (1990). Social cognition and negative reactions to children: A comparison of mothers of aggressive and nonaggressive boys. *Journal of Social and Clinical Psychology*, *9*(4), 418–438. https://doi.org/10.1521/jscp.1990.9.4.418

Dodge, K. A., Dishion, T. J., & Lansford, J. E. (2007). *Deviant peer influences in programs for youth: Problems and solutions*. Guilford Press.

Dodge, K. A., Lochman, J. E., Harnish, J. D., Bates, J. E., & Pettit, G. S. (1997). Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. *Journal of Abnormal Psychology*, *106*(1), 37–51. https://doi.org/10.1037//0021-843x.106.1.37

Dorans, N. J. (2007). Linking scores from multiple health outcome instruments. *Quality of Life Research*, *16*, 85–94.

Eid, M., Geiser, C., & Koch, T. (2016). Measuring method effects: From traditional to design-oriented approaches. *Current Directions in*

Psychological Science, *25*(4), 275–280. https://doi.org/10.1177/0963721416649624

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition: From brains to culture*. Sage.

Freitag, C. M., Haenig, S., Schneider, A., Seitz, C., Palmason, H., Retz, W., & Meyer, J. (2012). Biological and psychosocial environmental risk factors influence symptom severity and psychiatric comorbidity in children with ADHD. *Journal of Neural Transmission*, *119*(1), 81–94. https://doi.org/10.1007/s00702-011-0659-9

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Campbell Systematic Reviews*, *8*(1), 1–239. https://doi.org/10.1002/14651858.CD008225.pub2

Gerdes, A. C., Malkoff, A., Kapke, T. L., & Grace, M. (2021). Parental ADHD Knowledge in Latinx Families: Gender Differences and Treatment Effects. *Journal of Attention Disorders*, *25*(14), 1955–1961. https://doi.org/10.1177/1087054720951853

Grimbos, T., & Granic, I. (2009). Changes in maternal depression are associated with MST outcomes for adolescents with co-occurring externalizing and internalizing problems. *Journal of Adolescence*, *32*(6), 1415–1423. https://doi.org/10.1016/j.adolescence.2009.05.004

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hammon, C. (2002). Context of stress in families of children with depressed parents. In S. H. Goodman, & I. H. Gotlib (Eds.), *Children of depressed parents: Mechanisms of risk an implications for treatment* (pp. 175–199). : American Psychological Association.

He, Q., Glas, C. A., & Veldkamp, B. P. (2014). Assessing impact of differential symptom functioning on post-traumatic stress disorder (PTSD) diagnosis. *International Journal of Methods in Psychiatry Research*, *23*(2), 131–141. https://doi.org/10.1002/mpr.1417

Howe, G. W., Dagne, G. A., Brown, C. H., Brincks, A. M., Beardslee, W., Perrino, T., & Pantin, H. (2019). Evaluating construct equivalence of youth depression measures across multiple measures and multiple studies. *Psychological Assessment*, *31*(9), 1154–1167. https://doi.org/10.1037/pas0000737

Hudec, K. L. & Mikami, A.Y. (2017). Diagnostic Issues for ODD/CD with ADHD Comorbidity. In J.E. Lochman & W. Matthys (Eds.), *The Wiley Handbook of Disruptive and Impulse-Control Disorders*. https://doi.org/10.1002/9781119092254.ch4

Kamphaus, R. W., Reynolds, C. R., & Hatcher, N. M. (1999). Treatment planning and evaluation with the BASC: The Behavior Assessment System for Children.

King, K. M., Littlefield, A. K., McCabe, C. J., Mills, K. L., Flournoy, J., & Chassin, L. (2018). Longitudinal modeling in developmental neuroimaging research: Common challenges, and solutions from developmental psychology. *Developmental Cognitive Neuroscience*, *33*, 54–72. https://doi.org/10.1016/j.dcn.2017.11.009

Littlefield, A. K., King, K. M., Acuff, S. F., Foster, K. T., Murphy, J. G., & Witkiewitz, K. (2021). Limitations of cross-lagged panel models in addiction research and alternative models: An empirical example using project MATCH. *Psychology of Addictive Behaviors*, *36*(3), 271–283. https://doi.org/10.1037/adb0000750

Lochman, J. E., Baden, R. E., Boxmeyer, C. L., Powell, N. P., Qu, L., Salekin, K. L., & Windle, M. (2014). Does a booster intervention augment the preventive effects of an abbreviated version of the coping power program for aggressive children? *Journal of Abnormal Child Psychology*, *42*(3), 367–381. https://doi.org/10.1007/s10802-013-9727-y

Lochman, J. E., Boxmeyer, C. L., Powell, N. P., Qu, L., Wells, K., & Windle, M. (2012). Coping power dissemination study: Intervention and special education effects on academic outcomes. *Behavioral Disorders*, *37*(3), 192–205. https://doi.org/10.1177/019874291203700306

Lochman, J. E., Dishion, T. J., Boxmeyer, C. L., Powell, N. P., & Qu, L. (2017). Variation in response to evidence-based group preventive intervention for disruptive behavior problems: A view from 938 coping power sessions. *Journal of Abnormal Child Psychology*, *45*(7), 1271–1284. https://doi.org/10.1007/s10802-016-0252-7

Lochman, J. E., Dishion, T. J., Powell, N. P., Boxmeyer, C. L., Qu, L., & Sallee, M. (2015). Evidence-based preventive intervention for preadolescent aggressive children: One-year outcomes following randomization to group versus individual delivery. *Journal of Consulting and Clinical Psychology*, *83*(4), 728–735. https://doi.org/10.1037/ccp0000030

Lochman, J. E., & Dodge, K. A. (1998). Distorted perceptions in dyadic interactions of aggressive and nonaggressive boys: Effects of prior expectations, context, and boys' age. *Development and Psychopathology*, *10*(3), 495–512. https://doi.org/10.1177/019874291203700306

Lochman, J. E., Glenn, A. L., Powell, N. P., Boxmeyer, C. L., Bui, C., Kassing, F., Qu, L., Romero, D., & Dishion, T. (2019). Group versus individual format of intervention for aggressive children: Moderators and predictors of outcomes through four years after intervention. *Development and Psychopathology*, *31*(5), 1757–1775. https://doi.org/10.1017/S0954579419000968

Lochman, J. E., & Grescham, F. M. (2008). Intervention Development, Assessment, Planning, and Adaptation. *Cognitive-Behavioral Interventions for Emotional and Behavioral Disorders: School-Based Practice*, 29.

Lochman, J. E., & Lenhart, L. (1995). Cognitive behavioral therapy of aggressive children. In *Behavioral approaches for children and adolescents* (pp. 145–166). Springer.

Lochman, J. E., Wells, K., & Lenhart, L. A. (2008). *Coping power: Child group facilitator's guide* (vol. *2*). Oxford University Press.

Lochman, J. E., & Wells, K. C. (2002a). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the coping power program. *Development and Psychopathology*, *14*(4), 945–967. https://doi.org/10.1017/s0954579402004157

Lochman, J. E., & Wells, K. C. (2002b). The coping power program at the middle-school transition: Universal and indicated prevention effects. *Psychology of Addictive Behaviors*, *16*(4S), S40.

Lochman, J. E., & Wells, K. C. (2003). Effectiveness of the coping power program and of classroom intervention with aggressive children: Outcomes at a 1-year follow-up. *Behavior Therapy*, *34*(4), 493–515. doi:10.1016/S0005-7894(03)80032-1.

Lochman, J. E., & Wells, K. C. (2004). The coping power program for preadolescent aggressive boys and their parents: Outcome effects at the 1-year follow-up. *Journal of Consulting and Clinical Psychology*, *72*(4), 571–578. https://doi.org/10.1037/0022-006X.72.4.571

Lochman, J. E., Wells, K. C., & Lenhart, L. A. (2008). *Coping power: Child group program workbook*. Oxford University Press.

Lochman, J. E., Wells, K. C., Qu, L., & Chen, L. (2013). Three year follow-up of coping power intervention effects: Evidence of neighborhood moderation? *Prevention Science*, *14*(4), 364–376. https://doi.org/10.1007/s11121-012-0295-0

MacKinnon, D. P., Valente, M. J., & Gonzalez, O. (2020). The correspondence between causal and traditional mediation analysis: The link is the mediator by treatment interaction. *Prevention Science*, *21*(2), 147–157. https://doi.org/10.1007/s11121-019-01076-4

Malkoff, A. C., Grace, M., Kapke, T. L., & Gerdes, A. C. (2020). Family functioning in Latinx families of children with ADHD: The role of parental gender and acculturation. *Journal of Child and Family Studies*, *29*(4), 1108–1122. https://doi.org/10.1007/s10826-019-01673-

Mash, E. J., & Dozois, D. J. A. (1996). Child psychopathology: A developmental-systems perspective. In E. J. Mash, & R. A. Barkley (Eds.), *Child psychopathology* (pp. 3–60). The Guilford Press.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*(1), 23–44. https://doi.org/10.1037/1082-989X.12.1.23

McLeod, B. D., & Weisz, J. R. (2005). The therapy process observational coding system-alliance scale: Measure characteristics and prediction of outcome in usual clinical practice. *Journal of Consulting and Clinical Psychology*, *73*(2), 323–333. https://doi.org/10.1037/0022-006X.73.2.323

McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods*, *52*(6), 2287–2305. https://doi.org/10.3758/s13428-020-01398-0

Millsap, R. E., & Kwok, O.-M. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological Methods*, *9*(1), 93–115. https://doi.org/10.1037/1082-989X.9.1.93

Mitchell, Q. P., Younginer, S. T., Lochman, J. E., Vernberg, E. M., Powell, N. P., & Qu, L. (2021). Examining for the protective effects of therapeutic

engagement on child aggression. *Journal of Emotional and Behavioral Disorders*, *29*(3), 162–174. https://doi.org/10.1177/1063426620980701

Morgan-López, A. A., McDaniel, H. L., Bradshaw, C. P., Saavedra, L. M., Lochman, J. E., Kaihoi, C. A., . . . & Yaros, A. C. (2022). Design and methodology for an integrative data analysis of coping power: Direct and indirect effects on adolescent suicidality. *Contemporary Clinical Trials*, *115*, 106705. https://doi.org/10.1016/j.cct.2022.106705

Morgan, J., Robinson, D., & Aldridge, J. (2002). Parenting stress and externalizing child behaviour. *Child & Family Social Work*, *7*(3), 219–225. https://doi.org/10.1046/j.1365-2206.2002.00242.x

Morgan-López, A. A., Hien, D. A., Saraiya, T. C., Saavedra, L. M., Norman, S. B., Killeen, T. K., Simpson, T. L., Fitzpatrick, S., Mills, K. L., Ruglass, L. M., Back, S. E., López-Castro, T., & Consortium on Addiction, Stress and Trauma (CAST) (2022). Estimating posttraumatic stress disorder severity in the presence of differential item functioning across populations, comorbidities, and interview measures: Introduction to project harmony. *Journal of Traumatic Stress*, *35*(3), 926–940. https://doi.org/10.1002/jts.22800

Morgan-López, A. A., Killeen, T. K., Saavedra, L. M., Hien, D. A., Fitzpatrick, S., Ruglass, L. M., & Back, S. E. (2020). Crossover between diagnostic and empirical categorizations of full and subthreshold PTSD. *Journal of Affective Disorders*, *274*, 832–840.

Muthén, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika*, *43*(4), 551–560.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8thed.). Muthén & Muthén.

Neece, C. L., Green, S. A., & Baker, B. L. (2012). Parenting stress and child behavior problems: A transactional relationship across time. *American Journal on Intellectual and Developmental Disabilities*, *117*(1), 48–66. https://doi.org/10.1352/1944-7558-117.1.48

Novick, D. R., Lorenzo, N. E., Danko, C. M., & Tuscano, A. C. (2022). Evaluation of an integrated parenting intervention targeting maternal depression: Effects on parent attributions of child behaviors. *Journal of Child and Family Studies*, *31*(8), 1–14. https://doi.org/10.1007/s10826-022-02267-4

Ormhaug, S. M., Jensen, T. K., Wentzel-Larsen, T., & Shirk, S. R. (2014). The therapeutic alliance in treatment of traumatized youths: Relation to outcome in a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *82*(1), 52–64. https://doi.org/10.1037/a0033884

Poulin, F., Dishion, T. J., & Burraston, B. (2001). 3-year iatrogenic effects associated with aggregating high-risk adolescents in cognitive-behavioral preventive interventions. *Applied Developmental Science*, *5*(4), 214–224.

Powell, N. P., Boxmeyer, C. L., Baden, R., Stromeyer, S., Minney, J. A., Mushtaq, A., & Lochman, J. E. (2011). Assessing and treating aggression and conduct problems in schools: Implications from the coping power program. *Psychology in the Schools*, *48*(3), 233–242. https://doi.org/10.1002/pits.20549

Reuben, J. D., Shaw, D. S., Brennan, L. M., & Dishion, T. J. (2015). A family-based intervention for improving children's emotional problems through effects on maternal depressive symptoms. *Journal of Consulting and Clinical Psychology*, *83*(6), 1142–1148. https://doi.org/10.1037/ccp0000049

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children*. American Guidance Service.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children* (2nd ed.). American Guidance Service.

Sanders, M. R., & McFarland, M. (2000). Treatment of depressed mothers with disruptive children: A controlled evaluation of cognitive behavioral family intervention. *Behavior Therapy*, *31*(1), 89–112. https://doi.org/10.1016/S0005-7894(00)80006-4

Sandler, I. N., Ma, Y., Tein, J. Y., Ayers, T. S., Wolchik, S., Kennedy, C., & Millsap, R. (2010). Long-term effects of the family bereavement program on multiple indicators of grief in parentally bereaved children and adolescents. *Journal of Consulting and Clinical Psychology*, *78*(2), 131–143.

Shaffer, A., Lindhiem, O., Kolko, D. J., & Trentacosta, C. J. (2013). Bidirectional relations between parenting practices and child externalizing behavior: A cross-lagged panel analysis in the context of a psychosocial treatment and 3-year follow-up. *Journal of Abnormal Child Psychology*, *41*(2), 199–210. https://doi.org/10.1007/s10802-012-9670-3

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*(2), 417–439.

Shaw, D. S., Sitnick, S. L., Reuben, J., Dishion, T. J., & Wilson, M. N. (2016). Transactional effects among maternal depression, neighborhood deprivation, and child conduct problems from early childhood through adolescence: A tale of two low-income samples. *Development and Psychopathology*, *28*(3), 819–836.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107–120.

Singh, A. L., D'Onofrio, B. M., Slutske, W. S., Turkheimer, E., Emery, R. E., Harden, K. P., Heath, A. C., Madden, P. A. F., Statham, D. J., & Martin, N. G. (2011). Parental depression and offspring psychopathology: A children of twins study. *Psychological Medicine*, *41*(7), 1385–1395. https://doi.org/10.1017/S0033291710002059

Tofighi, D., & MacKinnon, D. P. (2011). RMediation: An R package for mediation analysis confidence intervals. *Behavior Research*, *43*(3), 692–700. https://doi.org/10.3758/s13428-011-0076-x

Tseng, H. W., Tsai, C. S., Chen, Y. M., Hsiao, R. C., Chou, F. H., & Yen, C. F. (2021). Poor mental health in caregivers of children with attention-deficit/hyperactivity disorder and its relationships with caregivers' difficulties in managing the children's behaviors and worsened psychological symptoms during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, *18*(18), 9745. https://doi.org/10.3390/ijerph18189745

van der Waerden, J., Galéra, C., Larroque, B., Saurel-Cubizolles, M. J., Sutter-Dallay, A. L., Melchior, M., & EDEN Mother-Child Cohort Study Group (2015). Maternal depression trajectories and children's behavior at Age 5 Years. *The Journal of Pediatrics*, *166*(6), 1440–8.e1. https://doi.org/10.1016/j.jpeds.2015.03.002

White, C., & Barrowclough, C. (1998). Depressed and non-depressed mothers with problematic preschoolers: Attributions for child behaviours. *British Journal of Clinical Psychology*, *37*(4), 385–398.

Zorzella, K. P., Muller, R. T., & Cribbie, R. A. (2015). The relationships between therapeutic alliance and internalizing and externalizing symptoms in trauma-focused cognitive behavioral therapy. *Child Abuse & Neglect*, *50*, 171–181. https://doi.org/10.1016/j.chiabu.2015.08.002