Causal Effects of Completing the *Parenting Wisely* Online Parenting Intervention on Recidivism Among Youth Involved in the Juvenile Justice System

Timothy J. Ozechowski[1], Donald Gordon[2], Holly B. Waldon[1], and Hyman Hops[1]

[1]Oregon Research Institute

[2]Family Works, Inc.

Methods

Al study procedures were approved by the Institutional Review Board of the Oregon Research Institute. The study was preregistered with ClinicalTrials.gov (NCT01861158).

Participants

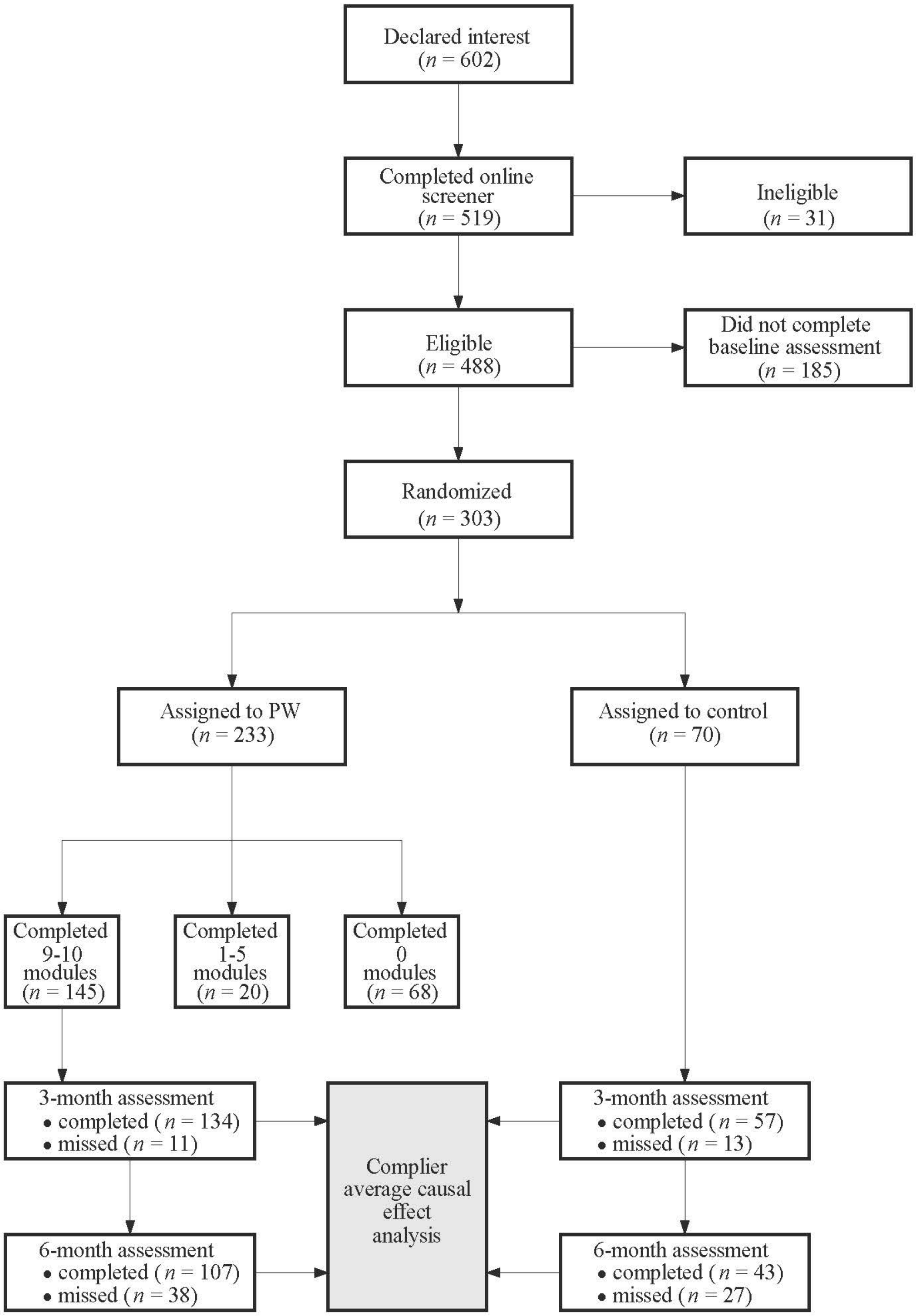
Participants in the clinical trial were parents of youth aged 12 to 17 years referred by law enforcement to receive juvenile justice services (excluding incarceration or other types of out-of-home placement). Parents were eligible to participate provided they lived with the target adolescent at least three days per week, had stable internet and e-mail access, were able to read and speak either English or Spanish.

Parents were referred to the study by JJS personnel in ten counties within four states including Oregon, New Mexico, Nevada, and Georgia. JJS personnel distributed brochures to parents providing information about the study along with the project website address and a phone number to contact the research office. Parents could enroll in the study directly via the project website or by contacting the study coordinator. A total of 602 parents declared interest in participating in the study by accessing the study website and creating a user account (see Figure 1). Of these, 519 parents (86.2%) completed the online screening questions to determine eligibility. Of those screened, 488 (94.0%) were eligible to participate. Eligible parents were provided with credentials to access the online consent forms and baseline assessment battery. Three hundred three eligible participants (62.1%) completed the online consent forms and the baseline assessments and were randomly assigned to either the experimental or control condition.

Seventy-six percent of parental participants were biological mothers, 12.5% were biological fathers, and the remaining were extended or step family members. Of the biological parental participants, 45% were non-Hispanic White, 28.6% were Mexican or New Mexican,

Figure 1

Participant Flowchart



11.8% were Black, 5.5% were Native American, and 9.1% were some other race. The mean age among parental participants was 40.5 years (*SD* = 6.8). Seventeen percent of parental participants had less than a high school education, 27.4% completed high school, and 43.1% attended some college, and 12.4% graduated college. Forty-five percent of parents reported and annual household income of $25,000 or less, 21.5% reported an annual household income between $25,000 and $55,000, and the remaining reported greater than $55,000. The mean age of adolescents as reported by parents was 15.1 years (*SD* = 1.5). Seventy-three percent of adolescents were male and 27% were female.

Procedures

Participants assigned to the experimental condition were provided with immediate access to the *Parenting Wisely* (PW) intervention whereas those assigned to the control condition were offered PW after a six-month delay. To ensure a sufficient number of participants for statistically modeling the effects of PW, a 3:1 imbalanced randomization protocol was implemented (see Chandereng et al., 2020). The final sample included 233 participants assigned to the PW condition and 70 assigned to the delayed condition.

In addition to the baseline assessment, all study participants were asked to complete a follow-up assessment battery three months and six months post-baseline. All questionnaires were completed online. Assessment data were entered directly into an encrypted database maintained on a secure project server. Parents were compensated $25 after completing the baseline assessment, $50 after completing the 3-month assessment, and $75 after completing the 6-month assessment.

Intervention

The experimental intervention in this study was an online version of PW which is a parent training intervention for adolescent behavioral problems (Feil et al., 2018; Gordon, 2000; Gordon & Rolland Stanar, 2003). The program consists of ten self-administered video modules which can be viewed on any electronic device with an internet connection. Each module presents an enacted vignette depicting parents and teens engaged in a conflictual interaction regarding a typical problem including drugs, curfew, friends, school work, household chores, etc. After each vignette, parents are asked to select a parenting response for dealing with each situation from a list of options spanning a range of parenting styles from assertive/authoritative, lax/permissive, to punitive/authoritarian. Next, parents are presented with an additional narrated vignette demonstrating and explaining the likely result of the selected parenting response. Parents are prompted to continue selecting response strategies until the assertive/authoritative one is viewed. In addition to the vignettes, the PW program provides interactive quizzes and other learning checks as well an accompanying workbook with practice exercises to complete between modules. Additionally, parents can communicate with other parents and trained professionals via an online forum. The PW website tracks parents' progress and completion of each module. The entire program can be completed in three to five hours. Parents receive an official certificate of completion after working through all the modules.

Measures

***Recidivism*.**

The primary outcome was youth recidivism which was defined as any police-generated report of an arrest or parole violation issued during the 6-month study period (Robertson et al., 2020). JJS programs in each county provided recidivism data for the target adolescent of each parental participant. Recidivism was coded as a binary variable where 1 indicated at least one police-generated report during the 6-month study period and 0 indicated no police generated reports during the 6-month study period (see Baglivio et al., 2014).

Eyberg Child Behavior Inventory (ECBI).

The ECBI is a 36-item scale measuring parents' perceptions of the frequency of a set of problematic child behaviors (Eyberg & Pincus, 1999). Each item is responded to on a 7-point ordinal scale ranging from 1 = *never* to 7 = *always*. A total score is obtained by summing the individual item scores. The ECBI has been widely utilized in child and adolescent behavioral research and has well established psychometric properties (Hukkelberg, 2016). The ECBI has been used successfully in numerous studies evaluating the effectiveness of parenting interventions (Leijten et al., 2017; Timmer et al., 2011; Webster-Stratton et al., 2011).

Strengths and Difficulties Questionnaire (SDQ).

The SDQ is a 25-item parent report measure of child mental and behavioral health and functioning (A. Goodman & Goodman, 2009; R. Goodman, 2001). Each item on the SDQ is scored on a 3-point ordinal scale indicating the degree to which a set of behavioral characteristics is true of the target child. The SDQ contains five subscales including *emotional symptoms, conduct problems, hyperactivity, peer problems*, and *prosocial behavior*. Each subscale score is obtained by summing the corresponding five items. The SDQ is one of the most widely used and well-validated child behavioral assessment instruments (Achenbach et al., 2008; Niclasen et al., 2013).

Parenting Scale (PS).

The PS is a 30-item self-report measure of parental behavioral management and discipline strategies in response to child problem behavior (Arnold et al., 1993). Each item is responded to on a 7-point ordinal scale on which 1 signifies a high probability of using an effective strategy and 7 indicates a high probability of utilizing an ineffective strategy. In the current study, we used the *laxness* and *overreactivity* subscales both of which are obtained by summing the corresponding items. The factor structure and psychometric properties of the PS are well established (Karazsia et al., 2008; Lorber et al., 2014).

Parenting Sense of Competence Scale (PSOC).

The PSOC is a 16-item measure of parents' perceptions of their parenting ability (Johnston & Marsh, 1989). Each item on the PSOC is responded to on a 6-point ordinal sale ranging from 1 = *strongly disagree* to 6 = *strongly agree*. The PSOC contains two subscales, *satisfaction* and *efficacy*, each of which are obtained by summing the corresponding items. The PSOC has been widely utilized and well validated in parenting research (Gilmore & Cuskelly, 2009; Jones & Prinz, 2005; Karp et al., 2015). The PSOC has also been used successfully in research on the effects of PW (Cotter et al., 2013; Stalker et al., 2018).

McMaster Family Assessment Device (FAD).

The FAD is a 60-item self-report measure of the quality of family functioning (Epstein et al., 1983). Each item is responded to on a 4-point ordinal scale ranging from 1 = *strongly agree* to 4 = *strongly disagree*. The FAD contains seven subscales including *problem-solving, communication, roles, affective responsiveness, affective involvement, behavioral control* and *general functioning*. Each subscale score is obtained as the sum of the corresponding items. The FAD is one of the most widely utilized and well validated measures of family functioning (Mansfield et al., 2015; Ryan et al., 2005). The FAD has been used in family-based research with JJS-involved youth (Folk et al., 2020; Tolou-Shams et al., 2018; Yonek et al., 2019) as well as on the effects of PW (Cotter et al., 2013; Stalker et al., 2018).

Parenting knowledge (PKN).

Developed specifically for this study, the PKN scale is a 16-item multiple choice test of parents' knowledge and comprehension of material covered in the PW program. Each item has four to six response options one of which is correct. For each item, one point is credited for a correct response and zero points for an incorrect response. A total score on the PKN scale is computed at the number of correct responses across the 16 items.

Statistical analyses.

A shown in Figure 1, 37.8% of parents assigned to the PW condition completed 5 or fewer modules including 29.2% who completed zero modules. Given this relatively high rate of intervention attrition, we opted to restrict the focus the analysis of the effects of PW to intervention "completers" or those completing 9 or 10 modules. Of all parents assigned to the PW condition, 145 (62.2%) were designated as completers. Data from PW "non-completers" were discarded from all analyses of PW effects.

Causal estimand

The appropriate estimand of intervention effects among intervention completers is known as the *complier average causal effect* (CACE: Jo, 2002; Stuart et al., 2008). The CACE estimator is the difference between the average outcome among intervention compliers (completers, in this case) and that among counterparts on the control condition who would have complied with the intervention had they been assigned to receive it. A fundamental challenge of CACE estimation is identifying "would-be" compliers in the control condition given that intervention compliance is unobserved among control participants (Follmann, 2000). One approach is to treat unobserved compliance status is missing data which may be jointly modeled along with the CACE using the expectation-maximization (EM) algorithm or Bayesian methods with noninformative priors (Imbens & Rubin, 1997; Little & Yau, 1998; Yau & Little, 2001). However, both of these approaches rely on large samples for accurate estimation (Little & Yau, 1998, p.157). Alternatively, unobserved intervention compliance status may be estimated using propensity scores (Follmann, 2000; Jo & Stuart, 2009). The propensity score approach estimates compliance status and intervention outcomes in separate models and therefore does not necessitate joint parametric assumptions about associations between intervention compliance and outcomes – both of which may be influenced by covariates. Accordingly, the propensity score approach to CACE modeling is methodologically simpler than joint modeling approaches and is more feasible to implement given modest sample sizes.

We conducted the CACE analysis using the propensity score approach detailed in Jo and Stuart (2009). Specifically, using data only from participants assigned to the PW condition we modeled the probability of PW completion using a series of simple and multiple logistic regression models which included an exhaustive set of predictor variables measured at baseline. We used the point estimates from the final multiple logistic regression model to estimate the conditional probability of PW completion P_i given the values of the baseline predictors for each study participant. Estimated values of P_i were used to compute individual analysis weights W_i which then were used in the CACE analysis. Specifically, we set $W_i = 1$ for all PW completers and $W_i = 0$ for all PW non-completers. For participants assigned to the control condition we set $W_i = \frac{P_i}{(1-P_i)}$. Accordingly, control group participants with relatively large estimates of P_i were weighted more heavily than those with smaller estimates of P_i and thereby were more influential in the estimation of the CACE. Under this propensity score weighting approach, the CACE estimator for a given treatment outcome Y was defined as the difference between the unweighted average of Y among PW completers and the weighted average of Y among participants assigned to the control condition. Assumptions underlying the CACE estimator are detailed in Jo (2002) and Stuart et al. (2008).

Missing Data

In addition to intervention attrition, there was a substantial amount of missing data primarily due to missed assessments at three months and six months (see Figure 1). We addressed missing data using multiple imputation (MI) which is a well-established technique for replacing missing values with plausible estimates given information about the missing values embedded in the observed data (Enders, 2010; Graham, 2012; Little & Rubin, 2002; Schafer & Graham, 2002). MI produces Bayesian estimates for missing values based on a joint probability model for the missing and observed data, which typically is a multivariate normal model by default. The MI procedure is set to execute m times, where m is specified by the analyst, resulting in m separate versions of the complete data. Statistical analyses are performed on each imputed data set separately and the collective results are averaged using formulas commonly known as "Rubin's rules" (Schafer & Graham, 2002).

MI yields valid results under the assumption that data are missing completely at random (MCAR) or missing at random (MAR) given a set of measured covariates. In randomized clinical trials a common source of non-randomly missing data is participant drop-out (Diggle & Kenward, 1994; Little, 1995). In the current study we defined dropouts as participants who missed the 6-month assessment. By this criterion the overall study dropout rate was 21.4% including 26.2% of PW completers and 38.6% of control group participants. Potential violations of random data missingness due to dropout can be mitigated by incorporating predictors of dropout into MI models (Enders, 2017). Accordingly, we identified a set of baseline variables that were significant predictors of dropout status separately within the PW and control conditions. In discriminant function analyses these sets of predictors correctly classified dropout status in 99.3% of PW completers and 100% of control group participants. Given these high dropout classification accuracy rates, we are confident that the MAR assumption was adequately upheld by including the dropout predictors in the imputation models for the PW and control conditions. We ran the MI procedure in the PW and control conditions separately. Based on guidelines provided by Graham et al. (2007), we set the number of imputations *m* to 200. We used the SAS PROC MI software to perform the imputations and SAS PROC MIANALYZE to obtain final estimates averaged across separate analyses of the 200 complete data sets.

Statistical Models

Intervention Effects on Recidivism. We estimated intervention effects on the primary outcome variable, youth recidivism, using a simple logistic regression model with a single binary predictor variable indicating study condition (1 = *PW*, 0 = *control*). We did not include any demographic variables as covariates because we did not find significant associations between youth recidivism and any parent or adolescent demographic variable in preliminary analyses.

Intervention Effects on Parent-Report Measures. We modeled intervention effects on change in the parent-reported measures of parenting and youth behavior across the baseline, 3-month, and 6-month assessments using piecewise latent growth models (PLGM: Bollen & Curran, 2006; Duncan et al., 2006; Flora, 2008). The parameters of PLGMs parse an entire longitudinal trajectory into discrete segments and estimate change separately within each segment. We specified a PLGM containing three factors including an intercept representing the value of the dependent variable Y at the baseline assessment, and two slope parameters representing change in Y between the baseline and 3-month assessments and between 3-month and 6-month assessments, respectively. Each PLGM factor is summarized by a mean and a variance parameter. Additional parameters include residual variance terms for *Y* at each measurement point. With only three longitudinal measurements, the PLGM with two slope factors amounts to a simpler parameterization of the latent change score model developed by McArdle (2009).

For each parent-reported measure, we fit separate unconditional PLGMs in the PW and control conditions within a multi-group modeling framework. The intercept mean and variance parameters were constrained to be equal between conditions due to random assignment. We tested intervention effects on change in each outcome *Y* by testing for differences between conditions in the slope factor means at both three months and six months.

Mediational Effects. We evaluated the mediational effects of change in parented-reported measures of parent/family functioning on the probability of youth recidivism at six months by combining the logistic regression and PLGMs described above and regressing the binary recidivism outcome variable on the PLGM intercept and slope factors (see Smid et al. 2020). We restricted the mediational modeling to include only those parent-reported measures of skill and functioning that produced significantly different slope factor means between conditions in the PLGM analyses. We fit the combined PLGM and logistic models separately in the PW and control conditions using a multiple-group modeling framework. We evaluated mediational effects by testing for differences in the effects of the PLGM factors on the probability of recidivism between the PW and control conditions. Significant differences in these effects between conditions suggested a mediational effect as per the "joint significance test" criterion (Fritz & MacKinnon, 2007; MacKinnon, 2008; MacKinnon et al., 2002).

Estimation and Inference

Given the relatively small sample size in this study, estimating the PLGMs and joint PLGM-logistic regression models described above using conventional maximum likelihood (ML) methods is likely to be unstable (McNeish, 2018). An alternative approach is Bayesian estimation and inference which has gained increasing interest among applied methodologists for addressing the statistical challenges posed by small samples (Baldwin & Fellingham, 2013; Lee & Song, 2004; D. McNeish, 2016, 2019; Smid, McNeish, et al., 2020; Song & Lee, 2012). Bayesian parameter estimates are anchored and stabilized in small samples by prior distributions which are probability density functions conveying known information about the parameter vector $\boldsymbol{\theta}$. Typically, prior parameters are specified a priori based on previous research or expert judgement. In most cases, a priori information about $\boldsymbol{\theta}$ is not available, especially for complex multivariate models, in which case vague or diffuse priors are utilized which have little impact on the analytic results. Consequently, the small-sample benefits of Bayesian methods accrue only if prior distributions are informative, that is, if they convey substantive information about $\boldsymbol{\theta}$. The combination of small samples and non-informative priors in Bayesian analyses can produce results that are highly unstable and inaccurate (Baldwin & Fellingham, 2013; McNeish, 2016, 2019).

In absence of prior information about $\boldsymbol{\theta}$, an alternative approach is to estimate prior parameters from the sample data in which case the analysis is no longer fully Bayesian but rather *empirical Bayes* (Carlin & Louis, 2009). Although empirical Bayes methods are objectionable to many Bayesian purists, the use of data-derived prior parameters affords other computational advantages of Bayesian estimation over ML, namely, the ability to obtain parameter estimates for complex models using Markov Chain Monte Carlo (MCMC) simulation and the capacity to constrain such estimates to fall within the valid parameter space though thoughtful prior specification (see Gelman et al., 2014; Kruschke, 2015).

To capitalize on the computational advantages of empirical Bayes methods, in the present study we estimated all statistical models using Bayesian methods with data-derived prior parameters. We followed methodological guidelines for empirical Bayes latent growth modeling provided by Ozechowski (2014) and (McNeish, 2016). Specifically, we obtained initial parameter estimates for all statistical models using ML estimation based on multiply imputed data. The final estimates were then used as informative prior parameters in a subsequent empirical Bayes estimation using MCMC simulation. In all empirical Bayes models, regression parameters were assigned a normal prior with mean equal to the corresponding ML point estimate and standard deviation equal to the corresponding ML standard error. All variance parameters were assigned element-specific gamma priors with shape parameter $\alpha = 2$ and scale parameter θ equal to the corresponding ML point estimate divided by 2 based on guidance provided by Chung and colleagues (Chung et al., 2013, 2015). Regarding model parameters for which ML did not yield valid estimates (e.g., negative estimates of slope factor variances in the PLGMs), we computed corresponding ordinary least squares (OLS) estimates directly from the imputed data and used the OLS estimates as informative prior parameters in the empirical Bayes estimation. We executed the empirical Bayes estimation for all statistical models using the SAS PROC MCMC software program (see Baldwin & Fellingham, 2013; McNeish, 2017; Ozechowski, 2014). We used the unimputed data in the empirical Bayes analysis. Missing data were estimated in the empirical Bayes analysis using a selection model containing the same dropout predictors as were included in the multiple imputation model (Daniels & Hogan, 2008; Mason et al., 2012).

Inference in the empirical Bayes analysis was based on posterior means and 95% credible intervals (CI). Support for a given hypothesized model was inferred if the CI for the relevant model parameter excluded the null value (Kruschke, 2015, ch. 12).

Results

Effect of PW Completion on Youth Recidivism

The effect of PW completion on youth recidivism at six months was tested using a simple logistic regression model with a single binary independent predictor indicating study condition (1 = *PW*, 0 = *control*). The posterior mean for the effect of PW completion relative to the control condition on the log-odds of youth recidivism was -0.64 (95% CI: -1.01 − -0.28). The 95% CI for the difference in the log-odds between conditions did not include the null value of 0 suggesting a lower rate of recidivism at 6-months among youth of PW completers compared to those of parents in the control condition. To help clarify this result, we translated the posterior estimates of the log-odds into posterior probabilities of recidivism for each condition. Among PW completers, the posterior mean probability of youth recidivism was 0.23 (95% credible interval: 0.15 − 0.32) compared to 0.39 (95% credible interval: 0.32 − 0.47) among youth whose parents were assigned to the control condition. The non-overlapping credible intervals for these estimates further indicates a significant beneficial effect of PW completion on youth recidivism at 6-months.

Effects of PW Completion on Parent-reported Skills and Functioning as Youth Behavior

We tested the effects of PW completion on all parent-reported measures of parenting skill and functioning, family functioning, and youth behavior using the PLGM strategy described above. As discussed previously, all PLGMs were estimated separately in the PW and control conditions within a joint modeling framework. Inferences regarding the effects of PW completion were based on the posterior difference in the slope factor estimates at three months and six month between conditions. We obtained evidence of a significant beneficial effect of PW completion on parental over-reactivity measured on the PS at three months (posterior mean difference between conditions: -4.10, 95% CI: -5.73 − -2.37), and on parenting knowledge at three months (posterior mean difference between conditions: 1.27, 95% CI: .73 – 1.84). The exclusion of 0 in the credible intervals for the difference between conditions on these outcomes indicates a significant beneficial effect of PW completion. The beneficial effect of PW completion on parental over-reactivity was also evident at the 6-month assessment (posterior mean difference between conditions: -2.48, 95% CI: -4.96 – 0.03), although the effect was not as strong as the 3-month effect (i.e. the upper end of the credible interval included 0).

Aside from these beneficial effects of PW completion, we obtained several results in favor of the control condition over PW including higher average increases on the behavioral control subscale of the FAD at both three months (posterior mean difference between conditions: 0.95, 95% CI: 1.82 − 0.11) and six months (posterior mean difference between conditions: 1.37, 95% CI: 2.59 − 0.17). These results indicate that mean levels of parent-reported behavioral control within the family increased among parents assigned to the control condition but did not increase among parents completing PW. In addition, we obtained evidence of a greater average increase in parental self-efficacy at three months among parents assigned to the control condition compared to those completing PW (posterior mean difference between conditions: 1.79, 95% CI: 3.17 − 0.41).

Mediational Effects

As discussed previously, we evaluated the mediational effects of change in the parent-reported measures of parent/family functioning on which there were significant differences between the PW and control conditions on the probability of youth recidivism at six months. Mediational effects were tested by combining the logistic regression and PLGMs described above and regressing the binary recidivism outcome variable on the PLGM intercept and slope factors. Results indicated that among parents assigned to the control condition increases in family behavioral control at both three months and six months were predictive of a higher probability of youth recidivism and six months. These associations were not evident among parents completing PW. The difference in the association between family behavioral control and youth recidivism was significant between conditions at both three months (posterior mean difference between conditions: -0.31, 95% CI: -0.55 − -0.08) and six months (posterior mean difference between conditions: -0.36, 95% CI: -0.66 − -0.06). Likewise, we obtained evidence of a positive association between parent-reported over-reactivity at baseline and the probability of youth recidivism at six months in the control condition, but a negative association between baseline parental over-reactivity and youth recidivism among parents completing PW. The difference in these effects between conditions was significant (posterior mean difference between conditions: -0.15, 95% CI: -0.20 − -0.01).

Summary

The main finding of this study was that parental completion of the PW program produced a significantly lower rate of recidivism among non-incarcerated JJS-involved youth compared to youth of parents assigned to the control condition. Empirical Bayes point estimates of youth recidivism at six months were 0.23 among completers in the PW condition and 0.39 in the control condition which amounts to an estimated reduction in recidivism of 41% on average due to PW completion. We also found evidence for significant improvements in parental knowledge and over-reactivity at three months among parents completing PW. Additionally, we found that parents assigned to the control condition on average reported higher levels of family behavioral control at three and six months and higher levels of parenting self-efficacy at three months compared to parents completing PW.

The results of our mediational modeling indicated that increases in family behavioral control at both three and six months reported by parents in the control condition predicted an increased risk of youth recidivism at six months. This association was not evident among parents completing PW thereby suggesting that increased family behavioral control may be a risk factor for recidivism among JJS-involved youth. Moreover, our results suggest that PW may counteract this apparently iatrogenic effect of JJS (see Gatti et al., 2009) in that mean levels of family behavioral control remained stable at three and six months among parents completing PW and were not associated with youth recidivism at six months. Along similar lines, our mediational modeling results suggested that higher baseline levels of parental over-reactivity among parents assigned to the control condition predicted an increased risk of youth recidivism at six months. The opposite effect was evident among parents completing PW. These results suggest that high levels of parental over-reactivity may be an additional risk factor for recidivism among youth receiving JJS, and that PW may neutralize this risk by lowing levels of parental over-reactivity.

In conclusion, our results suggest that PW completion reduces the risk of youth recidivism by counteracting potentially iatrogenic elements of JJS promoting increased behavioral control particularly among parents exhibiting relatively high levels of over-reactivity. Further PW research and development should focus on strategies for increasing rates of PW completion. Our results suggest that nearly 40% of parents offered PW complete fewer than half of the modules and nearly 30% complete zero modules. Nonetheless, the majority of parents offered PW do complete the program and demonstrate increases in parental knowledge and skill that appear to significantly reduce the risk of youth recidivism.

References

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H.-C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*(3), 251–275. https://doi.org/10.1111/j.1469-7610.2007.01867.x

Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). *The Parenting Scale: A Measure of Dysfunctional Parenting in Discipline Situations*. 8.

Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of Multisystemic Therapy and Functional Family Therapy Effectiveness: A Multiyear Statewide Propensity Score Matching Analysis of Juvenile Offenders. *Criminal Justice and Behavior*, *41*(9), 1033–1056. https://doi.org/10.1177/0093854814543272

Baldwin, S. A., & Fellingham, G. W. (2013). Bayesian methods for the analysis of small sample multilevel data with a complex variance structure. *Psychological Methods*, *18*(2), 151–164. https://doi.org/10.1037/a0030642

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Wiley.

Carlin, B. P., & Louis, Th. A. (2009). *Bayesian methods for data analysis.* (3rd ed.). Chapman & Hall/CRC.

Chandereng, T., Wei, X., & Chappell, R. (2020). Imbalanced randomization in clinical trials. *Statistics in Medicine*, *39*(16), 2185–2196. https://doi.org/10.1002/sim.8539

Chung, Y., Gelman, A., Rabe-Hesketh, S., Liu, J., & Dorie, V. (2015). Weakly Informative Prior for Point Estimation of Covariance Matrices in Hierarchical Models. *Journal of*

Educational and Behavioral Statistics, *40*(2), 136–157. https://doi.org/10.3102/1076998615570945

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A Nondegenerate Penalized Likelihood Estimator for Variance Parameters in Multilevel Models. *Psychometrika*, *78*(4), 685–709. https://doi.org/10.1007/s11336-013-9328-2

Cotter, K. L., Bacallao, M., Smokowski, P. R., & Robertson, C. I. B. (2013). Parenting Interventions Implementation Science: How Delivery Format Impacts the Parenting Wisely Program. *Research on Social Work Practice*, *23*(6), 639–650. https://doi.org/10.1177/1049731513490811

Daniels, M. J., & Hogan, J. W. (2008). *Missing data in longitudinal studes: Strategies for Bayesian modeling and sensitivity anaysis.* Chapman & Hall/CRC.

Diggle, P., & Kenward, M. G. (1994). Informative Drop-Out in Longitudinal Data Analysis. *Applied Statistics*, *43*(1), 49. https://doi.org/10.2307/2986113

Duncan, T. E., Duncan, S. C., & Stryker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Erlbaum.

Enders, C. (2010). *Applied missing data analysis*. Guilford.

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy*, *98*, 4–18. https://doi.org/10.1016/j.brat.2016.11.008

Feil, E. G., Gordon, D., Waldron, H., Jones, L. B., & Widdop, C. (2018). *Development and Pilot Testing of an Internet-based Version of Parenting Wisely*. 5.

Flora, D. B. (2008). Specifying Piecewise Latent Trajectory Models for Longitudinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 513–533. https://doi.org/10.1080/10705510802154349

Folk, J. B., Brown, L. K., Marshall, B. D. L., Ramos, L. M. C., Gopalakrishnan, L., Koinis-Mitchell, D., & Tolou-Shams, M. (2020). The Prospective Impact of Family Functioning and Parenting Practices on Court-Involved Youth's Substance Use and Delinquent Behavior. *Journal of Youth and Adolescence*, *49*(1), 238–251. https://doi.org/10.1007/s10964-019-01099-8

Follmann, D. A. (2000). On the Effect of Treatment among Would-Be Treatment Compliers: An Analysis of the Multiple Risk Factor Intervention Trial. *Journal of the American Statistical Association*, *95*(452), 1101–1109. https://doi.org/10.1080/01621459.2000.10474306

Fritz, M. S., & MacKinnon, D. P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychological Science*, *18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Gatti, U., Tremblay, R. E., & Vitaro, F. (2009). Iatrogenic effect of juvenile justice. *Journal of Child Psychology and Psychiatry*, *50*(8), 991–998. https://doi.org/10.1111/j.1469-7610.2008.02057.x

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). CRC Press.

Gilmore, L., & Cuskelly, M. (2009). Factor structure of the Parenting Sense of Competence scale using a normative sample. *Child: Care, Health and Development*, *35*(1), 48–55. https://doi.org/10.1111/j.1365-2214.2008.00867.x

Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(4), 400–403. https://doi.org/10.1097/CHI.0b013e3181985068

Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Gordon, D. A. (2000). Parent Training via CD-ROM: Using Technology to Disseminate Effective Prevention Practices. *The Journal of Primary Prevention*, 25.

Gordon, D. A., & Rolland Stanar, C. (2003). Lessons learned from the dissemination of parenting wisely, a parent training CD-ROM. *Cognitive and Behavioral Practice*, *10*(4), 312–323. https://doi.org/10.1016/S1077-7229(03)80049-4

Graham, J. (2012). *Missing data: Analysis and design*. Springer.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. *Prevention Science*, *8*(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Hukkelberg, S. (2016). The Eyberg Child Behavior Inventory: Factorial invariance in problem behaviors across gender and age. *Scandinavian Journal of Psychology*, *57*(4), 298–304. https://doi.org/10.1111/sjop.12290

Imbens, G. W., & Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Annals of Statistics*, *25*(1), 305–327.

Jo, B. (2002). Model misspecification sensitivity analysis in estimating causal effects of interventions with non-compliance. *Statistics in Medicine*, *21*(21), 3161–3181. https://doi.org/10.1002/sim.1267

Jo, B., & Stuart, E. A. (2009). On the use of propensity scores in principal causal effect estimation. *Statistics in Medicine*, *28*(23), 2857–2875. https://doi.org/10.1002/sim.3669

Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical Psychology Review*, *25*(3), 341–363. https://doi.org/10.1016/j.cpr.2004.12.004

Karazsia, B. T., van Dulmen, M. H. M., & Wildman, B. G. (2008). Confirmatory Factor Analysis of Arnold et al.'s Parenting Scale Across Race, Age, and Sex. *Journal of Child and Family Studies*, *17*(4), 500–516. https://doi.org/10.1007/s10826-007-9170-1

Karp, S. M., Lutenbacher, M., & Wallston, K. A. (2015). Evaluation of the Parenting Sense of Competence Scale in Mothers of Infants. *Journal of Child and Family Studies*, *24*(11), 3474–3481. https://doi.org/10.1007/s10826-015-0149-z

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). Elsevier.

Lee, S.-Y., & Song, X.-Y. (2004). Evaluation of the Bayesian and Maximum Likelihood Approaches in Analyzing Structural Equation Models with Small Sample Sizes. *Multivariate Behavioral Research*, *39*(4), 653–686. https://doi.org/10.1207/s15327906mbr3904_4

Leijten, P., Raaijmakers, M. A. J., Orobio de Castro, B., van den Ban, E., & Matthys, W. (2017). Effectiveness of the Incredible Years Parenting Program for Families with Socioeconomically Disadvantaged and Ethnic Minority Backgrounds. *Journal of Clinical Child & Adolescent Psychology*, *46*(1), 59–73. https://doi.org/10.1080/15374416.2015.1038823

Little, R. J. A. (1995). Modeling the Drop-Out Mechanism in Repeated-Measures Studies. *Journal of the American Statistical Association*, *90*(431), 1112–1121. https://doi.org/10.1080/01621459.1995.10476615

Little, R. J., & Yau, L. H. Y. (1998). Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model. *Psychological Methods*, *3*(2), 147–159.

Little, R., & Rubin, D. B. (2002). *Statistical analysis with missing data.* (2nd ed.). Wiley.

Lorber, M. F., Xu, S., Slep, A. M. S., Bulling, L., & O'Leary, S. G. (2014). A New Look at the Psychometrics of the Parenting Scale Through the Lens of Item Response Theory. *Journal of Clinical Child & Adolescent Psychology*, *43*(4), 613–626. https://doi.org/10.1080/15374416.2014.900717

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Erlbaum.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*(1), 83–104. https://doi.org/10.1037/1082-989X.7.1.83

Mansfield, A. K., Keitner, G. I., & Dealy, J. (2015). The Family Assessment Device: An Update. *Family Process*, *54*(1), 82–93. https://doi.org/10.1111/famp.12080

Mason, A., Richardson, S., Plewis, I., & Best, N. (2012). Strategy for Modelling Nonrandom Missing Data Mechanisms in Observational Studies Using Bayesian Methods. *Journal of Official Statistics*, *28*(2), 279–302.

McArdle, J. J. (2009). Latent Variable Modeling of Differences and Changes with Longitudinal Data. *Annual Review of Psychology*, *60*(1), 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McNeish, D. (2016). On Using Bayesian Methods to Address Small Sample Problems. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(5), 750–773. https://doi.org/10.1080/10705511.2016.1186549

McNeish, D. (2017). Fitting Residual Error Structures for Growth Models in SAS PROC MCMC. *Educational and Psychological Measurement*, *77*(4), 587–612. https://doi.org/10.1177/0013164416652441

McNeish, D. (2018). Brief Research Report: Growth Models With Small Samples and Missing Data. *The Journal of Experimental Education*, *86*(4), 690–701. https://doi.org/10.1080/00220973.2017.1369384

McNeish, D. (2019). Two-Level Dynamic Structural Equation Models with Small Samples. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 948–966. https://doi.org/10.1080/10705511.2019.1578657

McNeish, D. M. (2016). Using Data-Dependent Priors to Mitigate Small Sample Bias in Latent Growth Models: A Discussion and Illustration Using M *plus*. *Journal of Educational and Behavioral Statistics*, *41*(1), 27–56. https://doi.org/10.3102/1076998615621299

Niclasen, J., Skovgaard, A. M., Andersen, A.-M. N., Sømhovd, M. J., & Obel, C. (2013). A Confirmatory Approach to Examining the Factor Structure of the Strengths and Difficulties Questionnaire (SDQ): A Large Scale Cohort Study. *Journal of Abnormal Child Psychology*, *41*(3), 355–365. https://doi.org/10.1007/s10802-012-9683-y

Ozechowski, T. J. (2014). Empirical Bayes MCMC estimation for modeling treatment processes, mechanisms of change, and clinical outcomes in small samples. *Journal of Consulting and Clinical Psychology*, *82*(5), 854–867. https://doi.org/10.1037/a0035889

Robertson, A. A., Fang, Z., Weiland, D., Joe, G., Gardner, S., Dembo, R., Mcreynolds, L., Dickson, M., Pankow, J., Dennis, M., & Elkington, K. (2020). Recidivism Among Justice-Involved Youth: Findings From JJ-TRIALS. *Criminal Justice and Behavior*, 009385482092289. https://doi.org/10.1177/0093854820922891

Ryan, C. E., Epstein, N. B., Keitner, G. I., Miller, I. W., & Bishop, D. S. (2005). *Evaluating and treating families: The McMaster approach.* Taylor & Francis Group.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Smid, S. C., Depaoli, S., & Van De Schoot, R. (2020). Predicting a Distal Outcome Variable From a Latent Growth Model: ML versus Bayesian Estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(2), 169–191. https://doi.org/10.1080/10705511.2019.1604140

Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(1), 131–161. https://doi.org/10.1080/10705511.2019.1577140

Song, X.-Y., & Lee, S.-Y. (2012). *Basic and advanced Bayesian structural equation modeling: With applications in the medical and behavioral sciences.* Wiley.

Stalker, K. C., Rose, R. A., Bacallao, M., & Smokowski, P. R. (2018). Parenting Wisely Six Months Later: How Implementation Delivery Impacts Program Effects at Follow-Up. *The Journal of Primary Prevention*, *39*(2), 129–153. https://doi.org/10.1007/s10935-017-0495-2

Stuart, E. A., Perry, D. F., Le, H.-N., & Ialongo, N. S. (2008). Estimating Intervention Effects of Prevention Programs: Accounting for Noncompliance. *Prevention Science*, *9*(4), 288–298. https://doi.org/10.1007/s11121-008-0104-y

Timmer, S. G., Ho, L. K. L., Urquiza, A. J., Zebell, N. M., Fernandez y Garcia, E., & Boys, D. (2011). The Effectiveness of Parent–Child Interaction Therapy with Depressive Mothers: The Changing Relationship as the Agent of Individual Change. *Child Psychiatry & Human Development*, *42*(4), 406–423. https://doi.org/10.1007/s10578-011-0226-5

Tolou-Shams, M., Brogan, L., Esposito-Smythers, C., Healy, M. G., Lowery, A., Craker, L., & Brown, L. K. (2018). The role of family functioning in parenting practices of court-involved youth. *Journal of Adolescence*, *63*, 165–174. https://doi.org/10.1016/j.adolescence.2017.12.016

Webster-Stratton, C., Rinaldi, J., & Reid, J. M. (2011). Long-Term Outcomes of Incredible Years Parenting Program: Predictors of Adolescent Adjustment*: Long-Term Follow-Up. *Child and Adolescent Mental Health*, *16*(1), 38–46. https://doi.org/10.1111/j.1475-3588.2010.00576.x

Yau, L. H. Y., & Little, R. J. (2001). Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment for the unemployed. *Journal of the American Statistical Association*, *96*(456), 1232–1244.

Yonek, J. C., Dauria, E. F., Kemp, K., Koinis-Mitchell, D., Marshall, B. D. L., & Tolou-Shams, M. (2019). Factors Associated With Use of Mental Health and Substance Use Treatment Services by Justice-Involved Youths. *Psychiatric Services*, *70*(7), 586–595. https://doi.org/10.1176/appi.ps.201800322.