Exploring the Link Among Behavior Intervention Plans, Treatment Integrity, and Student Outcomes Under Natural Educational Conditions

Clayton R. Cook, G. Roy Mayer, Diana Browning Wright, Bonnie Kraemer, Michele D. Wallace, Evan Dart, Tai Collins and Alberto Restori

J Spec Educ 2012 46: 3 originally published online 13 May 2010

DOI: 10.1177/0022466910369941

Hammill Institute on Disabilities

The Journal of Special Education
46(1) 3–16

DOI: 10.1177/0022466910369941
http://journalofspecialeducation.sagepub.com
SAGE

Exploring the Link Among Behavior Intervention Plans, Treatment Integrity, and Student Outcomes Under Natural Educational Conditions

Clayton R. Cook[1], G. Roy Mayer[2], Diana Browning Wright[3], Bonnie Kraemer[4], Michele D. Wallace[2], Evan Dart[1], Tai Collins[1], and Alberto Restori[5]

Abstract

Several researchers have argued that the functional behavior assessment (FBA) and behavior intervention plan (BIP) mandates in the Individuals With Disabilities Education Act of 2004 have gone beyond the current research base. For instance, although BIPs have been shown to improve student outcomes when implemented with strict control and oversight by researchers, it is unclear whether these relationships hold true when implemented under real educational conditions. The purpose of this research was to conduct an initial study evaluating the relationship among the evidence-based quality of federally mandated BIPs, treatment integrity, and student outcomes under real-world educational conditions free from the help of researchers. Results indicated that the evidence-based quality of BIPs was significantly related to positive student outcomes. Results also supported the role of treatment integrity as a mediator of the relationship between the evidence-based quality of BIPs and student outcomes. The implications and limitations of this research as well as directions for future research are discussed.

Keywords

behavior intervention plans, effectiveness research, behavior problems, treatment integrity, functional behavior assessment

The importance of effectively and appropriately addressing behavior cannot be ignored, considering the impact that disruptive and destructive behaviors have on student learning as well as on the overall learning environment (Walker, Ramsey, & Gresham, 2004). In an attempt to improve special education students' outcomes and disciplinary practices, policy makers signed into law the landmark behavioral discipline provisions of the Individuals With Disabilities Education Act (IDEA, 1997), which were later renewed with the Individuals With Disabilities Education Improvement Act (IDEIA, 2004). According to these provisions, when a student's problem behaviors impede learning, the individualized education program team shall conduct a functional behavior assessment (FBA) and implement a behavior intervention plan (BIP). Case law has consistently shown that failure to adhere to the FBA-BIP mandates is tantamount to depriving the student of a free and appropriate public education (see Drasgow, Yell, Shriner, & Bradley, 1999; Etschdeit, 2006).

Despite the laudable intentions behind the inclusion of the FBA-BIP mandates, several research-based arguments have been put forth that suggest that the mandates were premature in the sense of not knowing (a) whether school personnel were capable of delivering such services and (b) whether such services actually lead to better outcomes for students (Conroy, Clark, Gable, & Fox, 1999; Gresham, Quinn, & Restori, 1999). Indeed, recent studies have shown that school personnel have made limited progress in adequately meeting the FBA-BIP mandates and fulfilling the intent of the law (Cook et al., 2007; C. R. Smith, 2000; Van Acker, Boreson, Gable, & Potterton, 2005). However, these same studies have identified that some educators are capable of developing high-quality BIPs that possess evidence-based components

[1]Louisiana State University, Baton Rouge, LA, USA
[2]California State University, Los Angeles, CA, USA
[3]Diana Browning Wright, Inc., CA, USA
[4]San Diego State University, San Diego, CA, USA
[5]California State University, Northridge, CA, USA

Corresponding Author:
Clayton R. Cook, Louisiana State University, 206 Audubon Hall, Baton Rouge, LA 70808, USA
E-mail: cook2142@lsu.edu

(Cook et al., 2007). What remains unclear from this scholarship is whether the BIPs that possess evidence-based components actually lead to improved student outcomes when implemented under natural educational conditions free from the influence of researchers.

When applying an evidence-based lens to the IDEIA (2004) statutes, it becomes apparent that additional research is needed to determine whether there is empirical support for certain legislative mandates. The FBA-BIP process is one mandate that has limited empirical support as to whether it translates into improved student outcomes when carried out in actual practice by everyday educators. Although a number of researchers have demonstrated the efficacy of FBA-based interventions to improve student behavior, this research has primarily been conducted under tight experimental conditions with individuals with developmental disabilities and with ongoing support provided to school personnel by researchers (Chandler, Dahlquist, Repp, & Feltz, 1999; Codding, Feinburg, Dunn, & Pace, 2005; DuPaul & Ervin, 1996; Ellingson, Miltenberger, Stricker, Galensky, & Garlinghouse, 2000; Gettinger & Stoiber, 2006; Ingram, Lewis-Palmer, & Sugai, 2005; Iwata, Dorsey, Slifer, Bauman, & Richman, 1994; Newcomer & Lewis, 2004; B. W. Smith & Sugai, 2000; Vollmer & Northup, 1996). This literature offers unclear guidance on whether the findings will transfer or generalize to everyday practice in the school. To provide evidentiary support for the FBA-BIP mandates, research is needed that evaluates the relationship between BIPs and student outcomes in naturalistic educational settings, using BIPs developed for a range of students and without ongoing support and guidance from researchers (i.e., effectiveness research).

The point is not to discount the importance of the previous research because there should be a logical progression of research from efficacy, or internal validity, to effectiveness, or external validity, with each part constituting an important aspect of empirical support for intervention procedures (Weisz, 2000). Rather, the point is to emphasize that it represents only half of the picture when considering the evidence-based status of BIPs. The other half, which may be more important than the efficacy research, is the data resulting from effectiveness research. To complete the missing half of the picture, effectiveness research must be conducted. The results of effectiveness types of research will help determine whether the findings from the efficacy research on BIPs translate into similar effects when employed with a broadly defined student population and with naturally varying implementation and adherence levels (Flay, 1986). The need for effectiveness research is not a novel concept considering that Horner and Carr (1997) called for this type of research some time ago:

> Existing strategies of personnel preparation will be valuable, but insufficient, to disseminate behavior support systems on a broad scale. . . . Greater attention is needed on the instruction of behavioral systems in real-world contexts rather than in simulations. (p. 108)

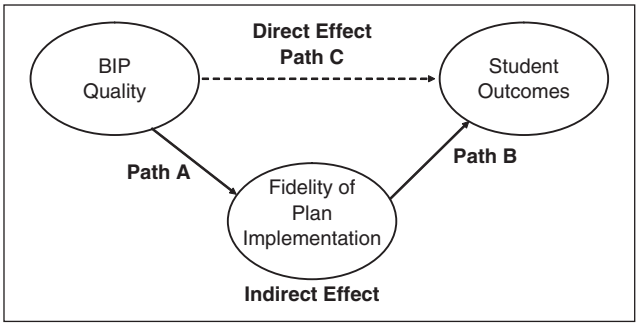

Figure 1. Hypothesized mediational model linking behavior intervention plan (BIP) quality, treatment integrity, and student outcomes.

The Purpose of This Study

The purpose of this study was to evaluate the relationship between the evidence-based quality of BIPs and student outcomes under natural educational conditions. In this way, this study was conducted in the spirit of effectiveness research to begin to uncover whether the efficacy research translates to actual practice in the schools. Several research questions of practical importance were addressed:

1. To what extent does the evidence-based, substantive quality of BIPs significantly predict positive student outcomes?
2. To what extent is the evidence-based quality of BIPs associated with treatment integrity under real-world conditions?
3. To what extent is there a significant relationship between the integrity with which BIPs are implemented and student outcomes?
4. To what extent does treatment integrity mediate the relationship between the evidence-based, substantive quality of BIPs and student outcomes (see Figure 1)?
5. To what extent are the above relationships cross-validated by a different informant?

Method

Participants

The participants in this study were educators in the state of California who are either directly involved in the Positive Environments Network of Trainers (PENT) organization or were identified by a PENT member because of their involvement as the primary or coimplementer of a BIP. Members of PENT were identified to participate by their special education

Table 1. PENT Member and Primary Implementer Demographic Information

	M		*SD*		Frequency		%	
Demographic variable	PENT	PI	PENT	PI	PENT	PI	PENT	PI
Number of BIPs developed	13.7	8.2	2.3	1.2				
Number of years in education	10.0	7.0	4.5	3.1				
Number of ABA courses	1.5	0.3	0.4	0.5				
Location								
Northern California					67	43	48	43
Southern California					72	57	52	57
Job								
School psychologist					57		41	
Behavior specialist					47		34	
Program specialist					35		25	
Special educator						85		85
General educator						5		5
Other						9		10
Student population								
High incidence					90	65	65	65
Low incidence					42	32	30	32
General ed.					7	3	5	3
Grade range								
Pre-K or elementary					56	42	40	43
Middle school					28	20	20	20
High school					41	34	30	34
Mixed					14	3	10	3

Note: PENT = Positive Environments Network of Trainers; PI = primary implementer; BIP = behavior intervention plan; ABA = applied behavior anaylsis.

local plan area directors for their work involving the development and implementation of BIPs. A total of 200 PENT members were initially asked to participate in this study as part of a larger training. Out of the 200 PENT members, 139 (70%) submitted BIPs for analysis. Of the 139 PENT members who submitted BIPs, 99 (71%) were able to successfully recruit primary implementers to participate in this study as cross-informants. The 99 primary implementers were a subset of educators who were linked to a PENT member via a BIP. In this way, the responses obtained from the primary implementers were capable of being compared to those of the PENT members to cross-validate findings. Thus, there were two samples of participants, albeit not independent, with complete data (139 PENT members and 99 primary implementers).

The demographic information gathered on the PENT members suggests that, on average, a PENT member was involved in education for 10 years, completed roughly two classes in applied behavior analysis, and developed nearly 14 BIPs (see Table 1). PENT members were predominantly school psychologists who worked with students with high-incidence disabilities. PENT members were the educators who were in charge of leading the BIP development process using data collected from an FBA given their expertise with this process and knowledge of the focal student. The demographic information gathered on the primary implementers suggests that, on average, a primary implementer was involved in education for 7 years, had not taken a course in applied behavior analysis, and had implemented approximately 8 BIPs (see Table 1). Primary implementers were primarily special education teachers who worked with students with high-incidence disabilities. Primary implementers were the educators who were most familiar with the students and their behavior problems and, therefore, were heavily involved with implementing the plan.

As for the students for whom the BIPs were written, the majority of students were receiving special education services under the eligibility category of emotional disturbance (34%), with the remaining breakdown as follows: 14% specific learning disability, 19% autism spectrum, 17% mental retardation, 5% general education, and 11% other. With regard to grade level, the majority of the students were in the elementary grades (43%), followed by high school (32%) and then middle school (25%; see Table 2).

Procedure

At the end of September 2006, 200 PENT members were sent an e-mail asking them to participate in the study. They were asked to identify and submit a BIP that met the following characteristics: (a) they had taken the lead role with developing the plan, (b) the plan was developed for a student within

Table 2. Classification and Grade of Students Included in Behavior Plans

Demographic variable	Frequency	%
Classification		
Learning disability	20	14
Emotional disturbance	47	34
Autism spectrum	27	19
Mental retardation	24	17
General education	7	5
Other		11
Grade range		
Pre-K or elementary	60	43
Middle school	35	25
High school	44	32

the previous week, and (c) the plan was currently being implemented. They were also asked to obtain consent from the student's parent and assent from the student.

The PENT members were also informed that they would receive a follow-up e-mail in a month with a link to a Web-based survey. They were also asked to provide the name and contact information of a fellow educator who was most familiar with the student's behavior problems and the primary or coimplementer of the submitted BIP. We waited a month before contacting the PENT member and the primary implementer to allow sufficient time for each plan to be implemented and the student to demonstrate a response. Out of the 200 original PENT members, 139 (70%) submitted BIPs and completed the Web-based survey. Of these 139 PENT members, 99 (71%) were able to recruit primary implementers to participate and complete the survey.

Research Design

A longitudinal research design was used to collect data at staggered time points to conduct proper meditational analyses. Data on the evidence-based quality of BIPs were collected near the time the plan was developed, whereas data on student outcomes and treatment integrity were collected after a month of plan implementation.

Measures

Behavior Support Plan-Quality Evaluation Scoring Guide. The evidence-based, substantive quality of BIPs was evaluated by rating each plan with the *Behavior Support Plan-Quality Evaluation Scoring Guide* (BSP-QE). The BSP-QE is a quantitative scoring rubric that was designed to objectively evaluate and rate the quality of the content specified in BIPs, including content consistency across items and presence of key concepts of effective behavior planning (Browning-Wright, Saren, & Mayer, 2003): Several articles have been published using the BSP-QE and found it to be a reliable and valid tool (Browning-Wright et al., 2007; Cook et al., 2007; Kraemer, Cook, Browning-Wright, Mayer, & Wallace, 2008). The BSP-QE is based on the six key, evidence-based concepts of BIPing identified from the literatures on applied behavior analysis, positive behavior support, team performance, and the law: (a) behavior function, (b) situational specificity, (c) behavior change: environmental alteration and teaching strategies, (d) reinforcement, (e) reactive strategies, and (f) team coordination and communication. In all, the BSP-QE includes 12 items that assess for the presence of the six key concepts and the consistency between items. The 12 items are rated on a 3-point Likert-type scale from 0 to 2 to produce a maximum score of 24.

A total of 56 advanced graduate students in applied behavior analysis from California State University, Los Angeles rated the 139 plans. Raters were extensively trained by experts in the area of BIPs on the six key concepts and how to score plans using the BSP-QE. Full, partial, and nonexamples were provided during training. Raters then practiced scoring plans and were provided feedback based on performance to improve their ability to rate plans successfully and accurately. Training continued until all raters scored the plans with 90% agreement with anchor scores derived from the developers of the instrument. Rating dyads were then created so all the BIPs were independently rated by two of the graduate students.

Evidence in support of the reliability of the BSP-QE was established by calculating item-total correlations, internal consistency, and interrater reliability (IRR) statistics. The item-total correlations ranged from .30 to .58, indicating adequate structure of overall plan quality. In terms of the internal consistency, the BSP-QE obtained a Cronbach's alpha of .82, which indicates adequate internal consistency across the rating items. In terms of IRR, all 139 (100%) plans were rated by two raters to generate estimates of IRR. Because of the 3-point Likert-type metric of the rating items, the most appropriate method of computing an estimate of IRR was to calculate a Pearson product–moment correlation between the pairs of scores for each item and the total plan score (Cohen, Cohen, West, & Aiken, 2003; Stemler, 2004). The IRR estimate for the total plan score was .79, which according to Stemler (2004) is an acceptable estimate of IRR.

BIP survey. A Web-based survey was constructed to assess student outcomes and aspects of plan implementation as well as to obtain demographic information on both the educators and the student involved with each plan. This survey involved obtaining raters' perception or judgment of student behavior and plan implementation. A principal components analysis was performed to reduce the items into factors representing student response or outcomes and treatment integrity. Interpretation of the scree plot and Eigenvalues indicated that there was a two-factor solution representing the constructs of student outcomes and treatment integrity. A simple structure was derived whereby the items loaded cleanly on their respective factors with factor loadings that exceeded 0.40. There were six items that assessed student response or

Table 3. Descriptive Statistics for BIP, Treatment Integrity, and Student Outcome Variables

	PENT member				Primary implementer			
Independent and dependent variables	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max
Total BIP score	16.04	4.10	4	24	16.04	4.10	4	24
Treatment integrity								
% of plan components implemented	68.15	19.89	5	100	75.06	23.99	10	100
BIP was implemented as written	1.63	0.68	0	3	1.73	0.79	0	3
Adult act more positively	2.06	0.63	0	3	2.19	0.58	0	3
Degree of adult behavior change	1.67	0.74	0	3	1.74	0.74	0	3
Student outcomes								
Replacement behavior(s) increase	1.74	0.70	0	3	1.78	0.64	0	3
Problem behavior(s) decrease	1.81	0.74	0	3	1.88	0.71	0	3
Positive behavior(s) increase	1.88	0.66	0	3	1.85	0.71	0	3
Academic performance improves	1.59	0.65	0	3	1.64	0.58	0	3
Overall behavior improvement	1.13	0.70	0	2	1.14	0.62	0	2
BIP goals met	1.20	0.56	0	2	1.19	0.58	0	2

Note: BIP = behavior intervention plan; PENT = Positive Environments Network of Trainers.

outcomes: (a) "What was the overall degree of improvement?" (3-point Likert-type scale), (b) "Were the behavioral goals of the interventions described in the plan met?" (3-point Likert-type scale), (c) "To what degree did *problem behaviors* change as a result of the implementation of the plan?" (4-point Likert-type scale), (d) "To what degree did *functionally equivalent replacement behaviors* change a result of the implementation of the plan?" (4-point Likert-type scale), (e) "To what degree did *positive behaviors* (other than functionally equivalent behaviors) change a result of the implementation of the plan?" (4-point Likert-type scale), and (f) "To what degree did the student experience improvements in *academic performance* as the result of the implementation of the plan?" (4-point Likert-type scale).

There were a total of four items that were used to assess treatment integrity: (a) "To what extent were the supports and strategies specified in the plan implemented as they were written?" (4-point Likert-type scale), (b) "What percentage of the supports and strategies included in the plan were implemented as intended?" (c) "What was the degree of adult behavior change that occurred as a result of the plan?" (4-point Likert-type scale), and (d) "What was the degree of change in the way that adults positively interacted with students as a result of the plan?" (4-point Likert-type scale). Reliability estimates in the form of Cronbach's alphas were computed for the each of the factors: student outcomes $\alpha = .84$, treatment integrity $\alpha = .87$.

Results

Descriptive Statistics

Descriptive statistics in the form of measures of central tendency (e.g., mean) and variability (i.e., standard deviation, range) were computed and are displayed in Table 3 for all continuous variables assessed in this study (e.g., BIP total score, treatment integrity variables, student outcome data). As one can see, the mean score for the sample of BIPs was 16 out of 24 possible points, which is consistent with previous research on samples of BIPs (Browning-Wright et al., 2007; Cook et al., 2007). The scores ranged from a minimum of 4 to a maximum of 24 points. Descriptive statistics were calculated separately for the PENT members' and primary implementers' ratings of treatment integrity and student outcomes. A comparison of the descriptive statistics across raters indicates that PENT members provided slightly lower ratings on most of the treatment integrity and student outcome variables than the primary implementers. Also, there was a discrepancy between the two groups' ratings of the percentage of BIP components implemented with integrity, with the PENT members (65%) providing lower ratings of treatment integrity than the primary implementers (75%).

Inferential Data Analytic Strategies

The primary data analytic techniques used in this research were correlational in nature. Specifically, bivariate correlations were calculated and structural equation modeling (SEM) was performed to address the proposed research questions. Assumptions of correlation analysis (i.e., independence of errors, normality, multicollinearity) and SEM (e.g., linearity, additivity, uncorrelated error terms, unidirectional causal flow; Bollen & Curren, 2006) were assessed and fully tested (e.g., scatterplots, box and whisker diagrams, plot observed versus predicted scores, normal probability plot of the residuals) prior to conducting analyses. Interpretation of the results was based on the obtained test statistics, associated *p* values, and effect sizes.

Table 4. Intercorrelation Matrix for BIP Quality and Positive Environments Network of Trainers Member Reports of Student Outcomes and Treatment Integrity

Independent and dependent variables	1	2	3	4	5	6	7	8	9	10	11
(1) BIP quality: Total plan score	1	—	—	—	—	—	—	—	—	—	—
(2) Student outcomes: Problem behavior(s) decrease	.47*	1	—	—	—	—	—	—	—	—	—
(3) Student outcomes: Replacement behavior(s) increase	.41*	.67*	1	—	—	—	—	—	—	—	—
(4) Student outcomes: Gen. positive behavior(s) increase	.31*	.60*	.58*	1	—	—	—	—	—	—	—
(5) Student outcomes: Academic performance improves	.24*	.59*	.46*	.66*	1	—	—	—	—	—	—
(6) Student outcome: Overall, behavior improvement	.31*	.71*	.42*	.49*	.44*	1	—	—	—	—	—
(7) Student outcome: BIP goals met	.33*	.48*	.42*	.40*	.34*	.47*	1	—	—	—	—
(8) Treatment integrity: BIP was implemented as written	.52*	.51*	.40*	.33*	.26*	.42*	.49*	1	—	—	—
(9) Treatment integrity: % of plan components implemented	.44*	.38*	.36*	.27*	.26*	.32*	.43*	.82*	1	—	—
(10) Treatment integrity: Adult act more positively	.48*	.58*	.46*	.46*	.41*	.44*	.40*	.48*	.42*	1	—
(11) Treatment integrity: Degree of adult behavior change	.41*	.55*	.34*	.42*	.29*	.50*	.48*	.65*	.56*	.59*	1

Note: BIP = behavior intervention plan.
*$p < 0.01$.

Correlation Analyses

Bivariate correlation matrices were separately calculated for the PENT members and the primary implementers. The first set of correlations was calculated to assess the interrelationships among the BIP total score, treatment integrity variables, and student outcome variables. Examination of the correlation matrix for the PENT members revealed that all variables significantly correlated with each other in the predicted directions (see Table 4). Most notably, the total BIP score positively correlated with all student outcome and treatment integrity variables. The BIP total score correlated most strongly with decrease in problem behaviors (student outcome $r = .47$) and BIP was implemented as written (treatment integrity $r = .52$). The BIP total score had the weakest correlations with academic performance improved (student outcome $r = .24$) and increase in general positive behaviors (student outcome $r = .31$). Treatment integrity variables significantly correlated with the student outcome variables. As a group, the treatment integrity variables correlated most strongly with decrease in problem behaviors and most weakly with academic performance improved.

The correlations for the primary implementers revealed similar findings to those for the PENT members (see Table 5). In particular, the BIP total score related significantly to both student outcome and treatment integrity items in the predicted direction. However, unlike the PENT members, not all the correlations were statistically significant below the $p < .05$ level. For example, the total BIP score did not significantly relate to academic performance improved (student outcome $r = .09$). Also, the academic performance improved rating item did not relate significantly with any of the items assessing treatment integrity. Overall, the results indicated that ratings of BIP quality correlated significantly and positively with treatment integrity and student outcome variables across both groups of raters. Also, the variables assessing treatment integrity and student outcomes correlated significantly and positively with each other across both groups of raters.

SEM Analyses

SEM was performed using maximum likelihood estimation with the AMOS 7.0 statistical software (Arbuckle, 1997). As suggested by Hu and Bentler (1999), three fit indices were used to assess the goodness of fit for the models: the normed fit index (NFI; values of .90 or greater indicate acceptable fit and .95 or greater indicate good fit), the root mean square error of approximation (RMSEA; values less than .08 indicate acceptable fit and .05 or less indicate good fit), and the χ^2 goodness of fit (nonsignificant chi-square statistics indicate the model provides an adequate fit).

SEM is particularly useful in assessing a mediational model. A meditational model is defined as a model that depicts the mechanism by which (i.e., mediator) an initial variable affects an outcome variable. Mediation analysis is a key part of conducting what has been called a process analysis

Table 5. Intercorrelation Matrix for BIP Quality and Primary Implementer Reports of Student Outcomes and Treatment Integrity

Independent and dependent variables	1	2	3	4	5	6	7	8	9	10	11
(1) BIP quality: Total plan score	1	—	—	—	—	—	—	—	—	—	—
(2) Student outcomes: Problem behavior(s) decrease	.61*	1	—	—	—	—	—	—	—	—	—
(3) Student outcomes: Replacement behavior(s) increase	.46*	.64*	1	—	—	—	—	—	—	—	—
(4) Student outcomes: Positive behavior(s) increase	.19+	.47	.49*	1	—	—	—	—	—	—	—
(5) Student outcomes: Academic performance improves	.09	.30*	.26*	.54*	1	—	—	—	—	—	—
(6) Student outcome: Overall, behavior improvement	.38*	.76*	.41*	.47*	.21*	1	—	—	—	—	—
(7) Student outcome: BIP goals met	.50*	.64*	.54*	.47*	.08	.47*	1	—	—	—	—
(8) Treatment integrity: BIP was implemented as written	.75*	.53*	.48*	.22*	-.02	.48*	.49*	1	—	—	—
(9) Treatment integrity: % of plan components implemented	.67*	.36*	.39*	.12	.12	.29*	.34*	.88*	1	—	—
(10) Treatment integrity: Adult act more positively	.57*	.50*	.34*	.25*	.17	.42*	.43*	.44*	.34*	1	—
(11) Treatment integrity: Degree of adult behavior change	.56*	.63*	.22*	.34*	.19+	.62*	.59*	.63*	.49*	.62*	1

Note: BIP = behavior intervention plan.
*$p < 0.01$.

(Kenny, Bolger, & Korchmaros, 2003). SEM is particularly useful in conducting a meditational analysis because two or more structural models with or without the indirect path (i.e., the mediator) can be readily compared to a null model.

One particular advantage of using SEM is to test the adequacy of model fit across groups using multiple-group SEM analysis. This type of analysis would potentially provide robust support for the hypothesized relationships. However, there was dependency in the observed data collected from the two groups because primary implementers were nested within or linked to the PENT members via a BIP. As a result, an intraclass correlation was calculated to determine whether the independence of observations assumption to perform a multiple-group SEM analysis was violated. The intraclass correlation between the PENT members' and the primary implementers' ratings was $r = .71$, indicating a high degree of dependency in the data across raters. A high intraclass correlation is an indicator that there is a violation of the independence of errors assumptions. Because of this violation, rather than performing a multiple-group SEM analysis, whereby the derived parameter estimates for the PENT members were held constant for the primary implementers, all models were first established with the data from the PENT members. Once a final model was identified, a separate SEM analysis was performed using data from the primary implementers to cross-validate the final model derived from the PENT members (Bollen & Curran, 2006).

The plan of analysis was to build the meditational model step by step to empirically demonstrate whether it provided the best account for the data. Per recommendations by Gerbing and Anderson (1988), confirmatory factor analyses were performed first to examine whether the measurement models provided an acceptable fit to the data. Once acceptable measurement models were developed, the structural model was then tested. Results indicated that each of the latent variable measurement models (BIP quality, treatment integrity, and student outcomes) provided an adequate fit to the data, as determined by examination of the three fit indices. All the models had a nonsignificant χ^2, an RMSEA less than .08, and an NFI greater than .95.

Analyses of the meditational model were conducted in light of recommendations by Baron and Kenny (1986) and MacKinnon, Lockwood, Hoffman, West, and Sheets (2002). Four sequential steps to model building were used to establish mediation. In each of the steps, the fit of both a null base model and the target model was estimated and corresponding chi-square statistics were compared to determine the adequacy of the target model.

Research Question 1: To what extent does the evidence-based, substantive quality of BIPs significantly predict positive student outcomes? Step 1 entails showing that the Y criterion variable (i.e., student outcomes) is correlated with the X predictor variable (i.e., BIP quality). This step establishes that there is an effect that may be mediated. Thus, the first step was to establish the presence of a direct effect between the latent variables of BIP quality and student outcomes. The results from this analysis suggested that the model depicting the direct effect of BIP quality on student outcomes significantly fit the data, $\chi^2(120) = 131.89$, $p = .22$; RMSEA = .05, 90% confidence interval (CI) = .02–.07; NFI = .96, suggesting that the evidence-based quality of a BIP is significantly related to positive student outcomes. This model fit significantly better than the null model without the path linking BIP quality to student outcomes, $\Delta\chi^2(1) = 245.21$, $p < .0001$.

Research Question 2: To what extent is the evidence-based quality of BIPs associated with treatment integrity under real-world conditions? Step 2 in establishing mediation requires that the X predictor variable (i.e., BIP quality) and Y criterion variable (i.e., student outcomes) are separately correlated with the M mediator variable (i.e., treatment integrity). The M mediator variable is, therefore, separately regressed on X and Y to establish M as a potential mediator of the effects

of X on Y. First, the results of the model examining the relationship between BIP quality (X) and treatment integrity (M) revealed that it provided a significant fit to the data, $\chi^2(96) = 107.52$, $p = .20$; RMSEA = .02, 90% CI = .00–.06; NFI = .96. This model provided a significantly better fit to the data than the null model without the direct path between BIP quality and student outcomes, $\Delta\chi^2(1) = 77.48$, $p < .001$. The results from this model suggest that the evidence-based quality of BIPs is significantly, positively related to PENT members' ratings of the degree to which the plans were implemented as intended.

Research Question 3: To what extent is there a significant relationship between the integrity with which BIPs are implemented and student outcomes? The results from the model testing the relationship between treatment integrity (M) and student outcomes (Y) indicated that it provided a significant fit to the data, $\chi^2(27) = 32.14$, $p = .23$; RMSEA = .04, 90% CI = .01–.09; NFI = .95. This model also fit the data significantly better than the null model, $\Delta\chi^2(1) = 92.16$, $p < .001$. The findings from this model suggest that PENT members reported that the better the plans were implemented as written, the more likely student outcomes improved.

Research Question 4: To what extent does treatment integrity mediate the relationship between the evidence-based, substantive quality of BIPs and student outcomes? Step 3 requires testing the actual meditational model because all the prerequisites to establishing mediation have been tested and supported by previous models. The results of this analysis suggested the mediation model including both the direct and indirect paths linking BIP quality to student outcomes significantly explained the data, $\chi^2(191) = 209.12$, $p = .18$; RMSEA = .05, 90% CI = .02–.08; NFI = .95. This model fit the data significantly better than the null model without the direct and indirect paths, $\Delta\chi^2(3) = 189.07$, $p < .001$. It is important to point out that the direct path between BIP quality and student outcomes reduced from .49 to .09 when the indirect path was included in the model. This path was not statistically significant.

The purpose of Step 4 is to establish whether M (i.e., treatment integrity) completely mediates the relationship between X (i.e., BIP quality) and Y (i.e., student outcomes). This requires comparing this model to the model including both the direct and indirect paths. The model without the direct path between BIP quality and student outcomes fit the data significantly well, $\chi^2(192) = 204.28$, $p = .26$; RMSEA = .03, 90% CI = .01–.07; NFI = 0.97, and it provided a significantly better fit to the data than the null base model, $\Delta\chi^2(2) = 193.91$, $p < .001$ (see Figure 2). This model was then compared to the model including both the direct and indirect path. Results suggested that the model without the direct path (i.e., treatment integrity as a full mediator) provided a significantly better fit to the data than the model without the direct path between BIP quality and student outcomes, $\Delta\chi^2(1) = -4.84$, $p = .027$. Thus, in light of the law of parsimony and model comparison logic, a final model was specified in which the direct path from BIP quality and student outcomes was dropped. To further test the significance of the indirect effect, Sobel's test was calculated (Sobel, 1982). The result from the Sobel test indicated that BIP quality was significantly indirectly related (.34) to student outcomes via treatment integrity ($z = 2.65$, $p < .01$).

Research Question 5: To what extent are the above relationships cross-validated by a different informant? As discussed earlier, the final model parameters derived from the PENT members were held constant for the primary implementers. When conducting this analysis, particular attention was paid to invariant structural relations (paths) among the latent variables. The results of the cross-validation analysis suggested that the model without the direct path between BIP quality and student outcomes provided a significant fit to the data, $\chi^2(192) = 213.01$, $p = .14$; RMSEA = .06, 90% CI = .02–.10; NFI = .94 (see Figure 3). The meditational model fit the data significantly better than the null model, $\Delta\chi^2(2) = 194.81$, $p < .001$. The path estimates were slightly weaker in magnitude than the path estimates from the PENT members but, nevertheless, were similar in magnitude, and structural relations among the latent variables were invariant across the groups. These results support two main conclusions. First, the evidence-based quality of BIPs was found to positively relate to improved student outcomes. Second, the results supported the role of treatment integrity as a full mediator of the relationship between BIP quality and student outcomes.

Discussion

This research represents an initial study examining the relationship among BIP quality, treatment integrity, and student outcomes under natural educational conditions free from the influence of researchers. In this way, this research attempts to provide the missing translational research needed to determine whether the positive findings from the efficacy research on BIPs generalize to actual practice in the schools. The results of this study are preliminary but nevertheless reveal several important findings that are worthy of discussion.

The data resulting from this study provide preliminary support for the federal BIP mandates (IDEA, 1997; IDEIA, 2004). That is, the relationships established by the well-controlled and methodologically rigorous research (i.e., efficacy research) were found to generalize to reported practices and outcomes in natural educational settings. The results from the correlation and SEM analyses indicated that there was a significant positive relationship between the evidence-based quality of BIPs and student outcomes when carried out under real-world conditions: (a) reductions in identified problem behaviors, (b) increases in general positive behaviors, (c) increases in appropriate replacement behaviors, and (d) improved academic performance. Although many of the

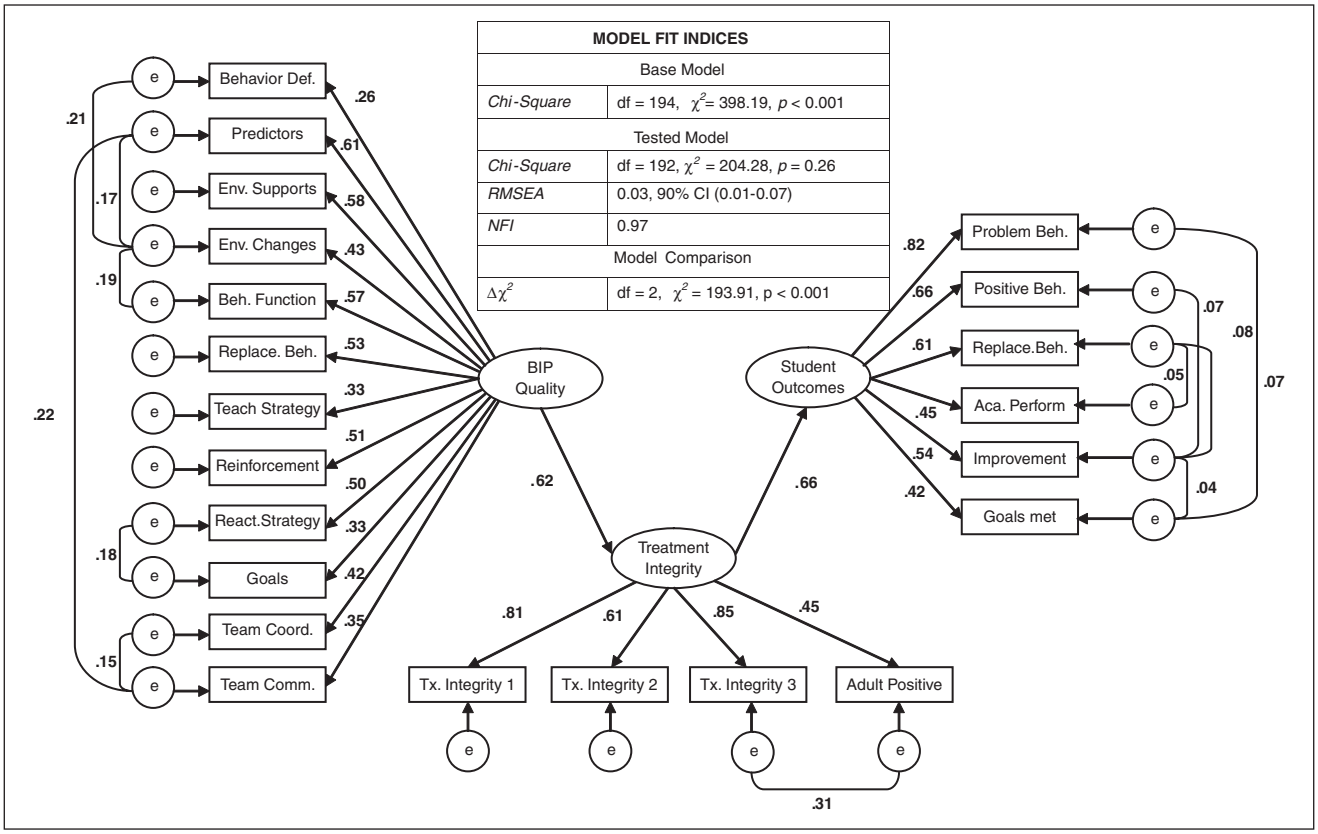

Figure 2. Step 4: Structural equation modeling results for mediational model without direct effect.

BIPs were lacking evidence-based components, those plans that did were associated with better reported improvements in student outcomes.

Given the prior research linking practitioner competence to treatment integrity (Gresham, 1989; Margison et al., 2000; Perepletchikova & Kazdin, 2005; Waltz, Addis, Koerner, & Jacobson, 1993), it was hypothesized that practitioners who are more likely to develop evidence-based BIPs are also more likely to implement the plans as written. If this is true, then the evidence-based quality of a BIP should affect student outcomes via treatment integrity. The opposite of this could actually be true, considering that evidence-based BIPs may contain strategies that require too much time and effort and, therefore, are implemented with less integrity. Across both groups of data informants, the correlation analyses revealed that BIP quality was significantly, positively, and strongly correlated with all four variables reported to assess treatment integrity. The SEM analysis corroborated the results from the correlation analyses. Together, these results revealed that evidence-based BIPs are likely to be implemented with higher levels of reported integrity than BIPs that are missing key evidence-based components.

Several studies have established the relationship between treatment integrity and student outcomes (McIntyre, Gresham, DiGennaro, & Reed, 2007; Noell & Witt, 1999). Therefore, treatment integrity has important implications for the degree of positive student outcomes that are likely to be seen as the result of a BIP. In line with this research, the results from this study indicated that the degree to which plans were implemented as written was significantly related to whether students improved; that is, both groups of informants reported that students improved more when the plans were implemented with greater integrity than with lower integrity. Thus, as predicted, treatment integrity had a significant positive relationship with improved student outcomes.

A related research inquiry guiding this study was to examine whether the substantive quality of BIPs would have an indirect effect on student outcomes via the extent to which they were implemented as planned. In this way, treatment integrity would serve as a mediator between BIP quality and student outcomes. As discussed above, this hypothesis was predicated on the research by Perepletchikova and Kazdin (2005), who contend that practitioner competency is meaningfully related to the selection and design of evidence-based services as well as the integrity with which the evidence-based services are delivered as intended. Following this logic, one would predict that teams with more competent members not only would develop better BIPs that are aligned with the

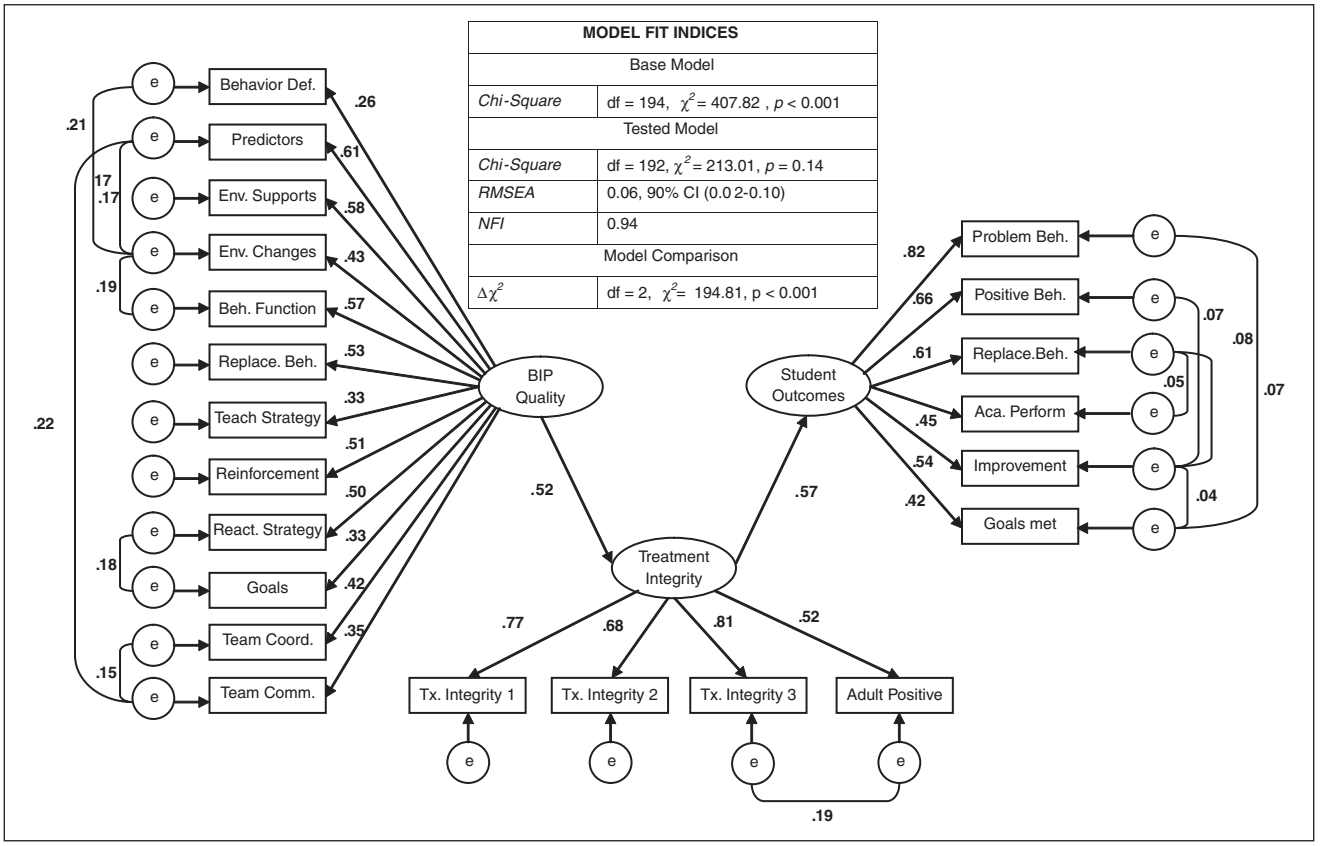

Figure 3. Cross-validation model with primary implementers.

research but also would be more likely to implement the plan as designed, with the end result being improved student outcomes.

The design of this research was deliberately longitudinal in nature to assess the evidence-based quality of BIPs at an earlier time point than student outcomes and treatment integrity because staggered data collection is a key feature of establishing a meditational relationship. The SEM analysis revealed that the model depicting treatment integrity as a full mediator of the effect of BIP quality on student outcomes provided the best fit to the observed data. Additional support for the meditational model was demonstrated by cross-validating it with the data from the primary implementers. These results provide support for the hypothesized meditational relationship among BIP quality, treatment integrity, and student outcomes and indicate that it may represent an accurate account of what actually unfolds in everyday school environments.

Implications

The findings from this research suggest that school staff should strive to develop BIPs that include critical evidence-based components. The components outlined in the literature review can serve as a useful starting point when developing a BIP. However, developing a BIP that is consistent with the research base should be viewed as a necessary but not sufficient condition to effect positive behavior change. Rather, school staff must diligently monitor and ensure the integrity of its implementation. Ensuring the integrity of implementation is paramount because it represents the mechanism by which key evidence-based concepts are translated into actual practice (Perepletchikova, Treat, & Kazdin, 2007). One must also consider the importance of treatment integrity beyond student outcomes, as it also has merit in terms of providing the context for making valid and legally defensible decisions to modify, intensify, maintain, or discontinue a BIP (Yell & Katisiyannis, 2000).

Treatment integrity is well recognized as a fundamental component in the delivery of evidence-based practices (e.g., Perepletchikova & Kazdin, 2005; Moncher & Prinz, 1991). It is one problem to not have the knowledge and skills to develop a high-quality BIP, but it is another one to develop a high-quality BIP that is likely to result in positive student outcomes but does not get implemented accurately or consistently. In the sample of plans assessed in this study, there were 13 (9%) BIPs that received total plan scores above

the mean but were implemented with less than 50% integrity. Consider the findings from Etschdeit (2006), whose analysis of 52 published court decisions indicated that the first thing hearing officers look for when making a decision is whether the BIP was implemented as planned (i.e., treatment integrity). Failure to faithfully implement the BIP may result in continued problem due process litigation and the school losing valuable resources (e.g., time, money, and credibility).

There are additional implications of this research that are noteworthy. First, this study demonstrates that school staff are capable of developing evidence-based BIPs that produce positive student outcomes without strict researcher oversight. Although some of the PENT members had received prior in-service training, many had not, and none of them developed and implemented the BIPs with assistance from researchers. Thus, with the appropriate training, knowledge, and skills, school personnel are capable of developing evidence-based BIPs that result in improved student outcomes.

Second, this research provides needed clarification on the content and substance of BIPs, which currently cannot be found in the law. Third, the findings take on increased importance when linked to the research on staff training. The literature on staff training suggests that school personnel's knowledge and skills to develop high-quality, individualized positive behavior supports can be enhanced via training (Browning-Wright et al., 2007; Gettinger & Stoiber, 2006; Kraemer et al., 2008). Merging the findings from these studies with the present research suggests that the improvements in educator competency observed after training may indeed translate into improved outcomes for students. This, however, remains unclear, as no research has rigorously assessed the impact of training on student outcomes using rigorous methods of scientific inquiry.

Limitations

The most obvious limitation of the present research is the reliance on indirect measures for the collection of student outcome and treatment integrity data. Research by Wickstrom, Jones, LaFleur, and Witt (1998) on the poor correspondence between direct (direct observation) and indirect (teacher report) methods demonstrates the limitations of relying on indirect methods to collect treatment integrity data. In light of this limitation, and in an effort to provide more robust support for the relationships under investigation, two data informants involved with the development and implementation of the BIP were recruited to provide information on the students' response to the plan and degree of plan integrity. Results revealed moderate to strong correlations between the two groups of informants for treatment integrity and student outcome.

This research was also correlational in nature, and therefore all the limitations associated with correlational research are warranted here. Interpretation of the findings should be tempered to reflect the inability to draw definitive causal conclusions regarding the data. The goal of this research, however, was not to conduct a highly rigorous evaluation of the impact of BIPs on student outcomes to arrive at internally valid conclusions. Rather, the goal was to take the important first step in exploring the connection between BIPs and student outcomes and determining whether the evidence-based, substantive quality of BIPs has any impact on student behavior.

There are also limitations in this study with regard to assessing the generalization and maintenance of student outcomes. Assessing whether BIPs are capable of producing generalized behavior and lasting effects represents an important empirical question that should be addressed in future research. This research did not attempt to assess the contextual fit of the BIPs, which has been argued to be an important aspect of the effectiveness of a BIP (Benazzi, Horner, & Good, 2006). Furthermore, this research did not assess the congruence between the FBA data and the information specified in each BIP, nor did it make an attempt to assess the accuracy of the identified function of behavior for each BIP. Despite these limitations, BIPs are capable of being assessed in their own right because it is the plan and not the FBA data that ultimately guides the actions of the team to improve student behavior.

Future Considerations

As with most studies, the results from this investigation raise additional questions that should be explored in future research. Namely, researchers will want to examine the pattern of results revealed herein using more rigorous methodology, such as direct measures of student behavior and treatment integrity. Also, the results from this study should stimulate additional research into the moderators of BIP effectiveness. That is, research should examine with whom and under which conditions BIPs are maximally effective or not effective. Research to this end will allow educators to more intelligently design and match services to student need. For instance, the degree of contextual fit of a plan may serve to differentially affect student outcomes, whereby only substantively adequate plans that also are appropriately matched to the context in which the plans are implemented produce desirable outcomes for students.

Researchers will also want to extend this research by assessing the impact of training as it relates to improvements in practitioner competency and student outcomes. This research is much needed in the area of BIPs, as prior research has indicated that the vast majority of BIPs (~90%) developed in actual practice were rated as inadequate and missing key evidence-based components (Cook et al., 2007).

Conclusion

This study represents an initial effort to address the missing translational research related to the BIP literature. Although this study is not a pure effectiveness study, it was performed in the spirit of effectiveness research to examine the relationship among BIP quality, treatment, and student outcomes under natural educational conditions. The lack of effectiveness research brings into question the IDEIA (2004) mandates that call for the development and implementation of a BIP when a student's behavior impedes his or her learning or that of others. This research, however, provides preliminary practical support for the IDEIA discipline mandates in that evidence-based BIPs were significantly related to positive student outcomes. Although the percentage of plans that would be considered evidence based may represent the minority of BIPs developed in actual practice, the findings from this research suggest that when plans are developed to include evidence-based components, improved student outcomes may follow suit, that is, *if the plan is implemented as written*.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

Arbuckle, J. L. (1997). *Amos user's guide.* Chicago, IL: Small-Waters.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Benazzi, L., Horner, R. H., & Good, R. H. (2006). Effects of behavior support team composition on the technical adequacy and contextual fit of behavior support plans. *The Journal of Special Education*, *40*, 160–170.

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

Browning-Wright, D., Mayer, G. R., Cook, C. R., Crews, S. D., Kraemer, B. R., & Gale, B. (2007). A preliminary study on the effects of training using Behavior Support Plan Quality Evaluation Guide (BSP-QE) to improve positive behavioral support plans. *Education & Treatment of Children*, *30*, 89–106.

Browning-Wright, D., Saren, D., & Mayer, G. R. (2003). *The Behavior Support Plan-Quality Evaluation Guide*. Retrieved from http://www.pent.ca.gov

Chandler, L. K., Dahlquist, C. M., Repp, A. C., & Feltz, C. (1999). The effects of team-based functional assessment on the behavior of students in classroom settings. *Exceptional Children*, *66*, 101–122.

Codding, R. S., Feinburg, A. B., Dunn, E. K., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis*, *38*, 205–219.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Conroy, M., Clark, D., Gable, R. A., & Fox, J. J. (1999). Building competence in the use of functional behavior assessment. *Preventing School Failure*, *43*, 140–144.

Cook, C. R., Crews, S. D., Wright, D. B., Mayer, G. R., Gale, B., Kraemer, B., & Gresham, F. M. (2007). Establishing and evaluating the substantive adequacy of positive behavioral support plans. *Journal of Behavioral Education*, *16*, 191–206.

Drasgow, E., Yell, M. L., Shriner, J., & Bradley, R. (1999). The IDEA amendments of 1997: A school-wide model for conducting functional behavioral assessments and developing behavior intervention plans. *Education and Treatment of Children, 22,* 244–266.

DuPaul, G. J., & Ervin, R. A. (1996). Functional assessment of behaviors related to attention-deficit/hyperactivity disorder: Linking assessment to intervention design. *Behavior Therapy*, *27*, 601–622.

Ellingson, S. A., Miltenberger, R. G., Stricker, J., Galensky, T. L., & Garlinghouse, M. A. (2000). Functional assessment and intervention for challenging behaviors in the classroom by general classroom teachers. *Journal of Positive Behavior Interventions*, *2*, 85–97.

Etschdeit, S. K. (2006). Behavioral intervention plans: Pedagogical and legal analysis of themes. *Behavior Disorders*, *21*, 223–243.

Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventative Medicine*, *15*, 451–474.

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, *25*, 186–192.

Gettinger, M., & Stoiber, K. C. (2006). Functional assessment, collaboration, and evidence-based treatment: Analysis of a team approach for addressing challenging behaviors in young children. *Journal of School Psychology*, *44*, 231–252.

Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review*, *18*, 37–50.

Gresham, F. M., Quinn, M. M., & Restori, A. (1999). Methodological issues in functional analysis: Generalizability to other disability groups. *Behavioral Disorders*, *24*, 180–182.

Horner, R., & Carr, E. (1997). Behavioral support for students with severe disabilities: Functional assessment and comprehensive intervention. *The Journal of Special Education*, *31*, 84–109.

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

Individuals With Disabilities Education Act Amendments of 1997, Pub. L. No. 105-17, 111 Stat. 37 (1997).

Individuals With Disabilities Education Improvement Act Amendments of 2004, Pub. L. No. 108-446, 118 Stat. 2647 (2004).

Ingram, K., Lewis-Palmer, T., & Sugai, G. (2005). Function-based intervention planning: Comparing the effectiveness of function-based and non-function-based intervention plans. *Journal of Positive Behavior Interventions*, *7*, 224–236.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*, 197–209.

Kenny, D. A., Bolger, N., & Korchmaros, J. D. (2003). Lower-level mediation in multilevel models. *Psychological Methods*, *8*, 115–128.

Kraemer, B. R., Cook, C. R., Browning-Wright, D., Mayer, G. R., & Wallace, M. D. (2008). Effects of training on the use of the Behavior Support Plan Quality Evaluation Guide with autism educators: A preliminary investigation examining positive behavior support plans. *Journal of Positive Behavior Interventions*, *10*, 179–189.

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*, 83–104.

Margison, F. R., Barkham, M., Evans, C., McGrath, G., Clark, J. M., Audit, K., & Connell, J. (2000). Measurement and psychotherapy: Evidence-based practice and practice-based evidence. *British Journal of Psychiatry*, *177*, 123–130.

McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in *Journal of Applied Behavior Analysis* studies from 1991–2005. *Journal of Applied Behavior Analysis*, *40*, 659–672.

Moncher, F., & Prinz, R. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, *11*, 247–266.

Newcomer, L. L., & Lewis, T. J. (2004). An investigation of assessment reliability and effectiveness of function-based interventions. *Journal of Emotional and Behavioral Disorders*, *12*, 168–181.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002). Retrieved from http://www.ed.gov/legislation/ESEA02/

Noell, G. H., & Witt, J. C. (1999). When does consultation lead to intervention implementation? Critical issues for research and practice. *The Journal of Special Education*, *33*, 29–35.

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, *12*, 365–383.

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, *75*, 829–841.

Smith, B. W., & Sugai, G. (2000). A self-management functional assessment-based behavior support plan for a middle school student with EBD. *Journal of Positive Behavior Interventions*, *2*, 208–217.

Smith, C. R. (2000). Behavioral and discipline provisions of IDEA '97: Implicit competencies yet to be confirmed. *Exceptional Children*, *66*, 403–412.

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology* (pp. 290–312). San Francisco, CA: Jossey-Bass.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, *9*, 1–14.

Van Acker, R., Boreson, L., Gable, R. A., & Potterton, T. (2005). Are we on the right course? Lessons learned about current FBA/BIP practices in schools. *Journal of Behavioral Education*, *14*, 35–56.

Vollmer, T. R., & Northup, J. (1996). Some implications of functional analysis for school psychology. *School Psychology Quarterly*, *11*, 76–91.

Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices.* Belmont, CA: Wadsworth/Thomson.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, *61*, 620–630.

Weisz, J. R. (2000). President's message: Lab-clinic differences and what we can do about them: I. The clinic-based treatment development model. *Clinical Child Psychology Newsletter*, *15*(1), 1–10.

Wickstrom, K. F., Jones, K. M., LaFleur, L. H., & Witt, J. C. (1998). An analysis of treatment integrity in school-based behavioral consultation. *School Psychology Quarterly*, *13,* 141–154.

Yell, M. L., & Katisiyannis, A. (2000). Functional behavioral assessment and IDEA '97: Legal and practice considerations. *Preventing School Failure*, *44*, 158–162.

About the Authors

Clayton R. Cook, PhD, is an assistant professor of school psychology at Louisiana State University. His current interests include the application of response to intervention principles and procedures to address students with or at risk for emotional and behavioral disorders. He has coauthored two books on response to intervention and published several articles on educational programming for students with emotional and behavioral disorders.

G. Roy Mayer, EdD, is a professor at California State University, Los Angeles. He has served as a teacher, counselor, and school psychologist and taught at Indiana University and Southern Illinois University before teaching at CSULA. His research interests are in the areas of classroom management, schoolwide discipline, and consultation and collaboration.

Diana Browning Wright, MA, is a licensed educational psychologist, behavior analyst, and teacher. She continues to consult with educators nationally and formerly served as the director of the California Department of Education's statewide initiative—the Positive Environments, Network of Trainers. Her research interests focus of the real-world application of multitiered service delivery models in behavior and academics.

Bonnie Kraemer, PhD, is an associate professor of special education at San Diego State University. She worked as an assistant professor of special education at the University of New Mexico in Albuquerque for 3 years prior to coming to San Diego State in fall 2003. Her areas of teaching and research lie in the field of severe intellectual disabilities, with a specific focus on families, instruction, transition, quality of life, and positive behavior supports

Michele D. Wallace, PhD, is an associate professor at California State University, Los Angeles and the current applied behavior analysis program coordinator. She has authored and coauthored many publications and has presented numerous times at state and national conferences. Her current research interests are related to the refinement of assessment and treatment methodologies with respect to behavior problems, parent and staff training, and the acquisition of verbal behavior.

Evan Dart, BA, is a doctoral student in the School Psychology Program at Louisiana State University. Evan's research interests involve the assessment and treatment of students with challenging behaviors.

Tai Collins, BA, is a doctoral student in the School Psychology Program at Louisiana State University. His research interests focus on African American male overrepresentation in punitive discipline, special education referral and identification, and placement in restrictive settings.

Alberto Restori, PhD, is an associate professor and co-coordinator of the School Psychology Program at California State University, Northridge. His research interests include positive behavior support for children with emotional or behavioral disorders, response to intervention, and youth involvement in gangs.