

Transporting an Evidence-Based Classroom Management Program for Preschoolers With Disruptive Behavior Problems to a School: An Analysis of Implementation, Outcomes, and Contextual Variables

Elisa Steele Shernoff

University of Illinois at Chicago

Thomas R. Kratochwill

University of Wisconsin-Madison

The transportability of an evidence-based teacher professional development program, the Incredible Years Classroom Management Program, was evaluated. This study compared the impact of two training methods: self-administered videotape modeling (VM) and self-administered videotape modeling plus consultation (VMC) on teachers' use of classroom management strategies, reductions in disruptive behaviors, acceptability, and contextual barriers of sustaining EBIs in practice settings. Four pairs of teachers (N = 8) were randomly assigned to VM or VMC conditions in a multiple probe design. Students (1–2 per classroom; N = 13) displaying high levels of externalizing problems also participated. Statistically significant between-groups differences in teacher confidence ratings, use of positive instructional practices, and acceptability favored the VMC condition. Positive trends favoring VMC students, including clinically significant increases in social competence, were also found. Although students in both groups exhibited reductions in disruptive behaviors, significant between-groups differences related to disruptive behavior were not observed. Findings related to contextual variables (e.g., time, cost, institutional support, adaptability of the program) have implications for future adoption of this program.

Keywords: evidence-based interventions, classroom management, transportability, teacher professional development

Action Editor: Robyn S. Hess.

Elisa Steele Shernoff received her PhD in Educational Psychology from the University of Wisconsin-Madison in 2004 with a specialization in School Psychology. She is currently a Visiting Research Assistant Professor at the Institute for Juvenile Research at the University of Illinois at Chicago. Dr. Shernoff's research interests focus on the development and evaluation of school-based interventions for children with disruptive behavior problems, how evidence based practices are adopted and sustained in practice settings, and investigating the

There has never been a time in the history of education and psychology when there has been a stronger emphasis on the use of evidence-based interventions (EBIs) in mental health and educational settings (Chambers, Ringeisen, & Hickman, 2005; Kratochwill & Stoiber, 2002). EBIs are well-developed interventions in which highly regarded scientific methods have established a program as effective. In the field of education, federal, state, and local governments reportedly spend over \$330 billion per year on public education, including significant investments in educational interventions and professional development activities (Slavin & Fashola, 1998). Despite these expenditures, many school-based interventions have been inadequately researched or found to be ineffective (Kavale & Forness, 1999). Teacher training was specifically identified by the Coalition for Evidence-Based Policy (2002) as a domain in which existing practices have rarely been tested using rigorous scientific methods. Today, however, rising standards, accountability requirements, and national education legislation require that educational policies and practices are based on strong research evidence.

In the context of the national emphasis on identifying and implementing EBIs in educational settings, a high priority in research includes examining the transportability of interventions that have been successful in university-based efficacy trials (Schoenwald & Hoagwood, 2001). To date, most of the evidence amassed for EBIs has emerged in the context of efficacy trials, which are conducted in university settings or clinical research facilities using well-designed, precise methodology (Drake, Latimer, Leff, McHugo, & Burns, 2004). These studies include research participants who meet specific inclusion criteria and research assistants and principal investigators who implement prescribed interventions by adhering to research protocols and intervention manuals. Meta-analytic reviews of many efficacy studies paint a positive picture regarding the number of efficacious

role that professional development plays in educators' use of innovative practices. For her research summarized in this article, Dr. Shernoff was awarded the Outstanding Dissertation Award of Division 16 of the American Psychological Association.

Thomas R. Kratochwill received his PhD in Educational Psychology from the University of Wisconsin-Madison in 1973 with a specialization in School Psychology. In 1983 he returned to the University of Wisconsin-Madison to direct the School Psychology Program and Psychoeducational Clinic.

The research and professional recognition of Dr. Kratochwill extend into his selection as editor or associate editor of major research journals in psychology and education. He has been Associate Editor of *Behavior Therapy*, *The Journal of Applied Behavior Analysis*, *School Psychology Review*, and a guest editor of *Behavior Assessment*. He was selected as the founding editor of the APA Division 16 journal *School Psychology Quarterly* from 1984 to 1992.

Correspondence concerning this article should be addressed to Elisa Steele Shernoff, PhD, Institute for Juvenile Research, Department of Psychiatry, University of Illinois at Chicago, MC 747, 1747 West Roosevelt Road, Chicago, IL 60608. E-mail: eshernoff@psych.uic.edu

interventions to treat a variety of childhood problems, including disruptive behaviors (Brestan & Eyberg, 1998). However, few of the original studies included in these meta-analyses evaluate interventions implemented in practice settings. Effectiveness research, on the other hand, examines the degree to which intervention effects from efficacy trials hold in naturalistic settings, such as homes, schools, and community mental health agencies (Drake et al., 2004). In effectiveness research, participants are in need of immediate intervention and practitioners are not always trained in specific research protocols or target interventions.

Transportability research bridges the gap between university and real-world settings by examining the extent to which intervention effects generalize from research to practice and assessing the broader contextual variables (e.g., time, cost, training and supervision requirements, institutional support for EBIs) that mitigate the movement of EBIs into practice settings (Schoenwald & Hoagwood, 2001). These contextual variables have been found to contribute to practitioners' failure to implement EBI protocols in mental health settings (see Miller, 2001; Schmidt & Taylor, 2002; Simpson, 2002), but well-controlled studies examining the transportability of EBIs to schools has been more limited (Storch & Crisp, 2004). In the present study, we hypothesized that much more was occurring in schools to impact intervention outcomes than simply implementing a new technology (e.g., an EBI), and that these contextual variables affect how interventions are adopted by schools and delivered by teachers. In this study, transportability was evaluated by assessing whether additional training resources (i.e., teacher consultation) enhanced teacher fidelity and acceptability ratings, and behavioral outcomes for learners. We also investigated organizational factors that can impact the sustainability of EBIs.

THE INCREDIBLE YEARS CLASSROOM MANAGEMENT PROGRAM

The Incredible Years Classroom Management Program is a research-driven, teacher professional development program that draws on the effective instruction and classroom management literature related to preventing and managing disruptive behaviors (Good & Brophy, 2003). This program is based on social learning theory and the importance of adult-child socialization processes (Patterson, Reid, & Dishion, 1992). The coercion hypothesis is used to explain the development and

maintenance of disruptive behaviors, in which young children inadvertently learn and are reinforced for noncompliance at home, which perpetuates the use of coercion in dealing with teachers and peers upon school entry (Patterson et al., 1992). Furthermore, teachers with limited classroom management skills and low rates of praise often have classrooms with higher rates of aggression, which in turn can maintain behavior problems (Kellam, Ling, Merisca, Brown, & Ialongo, 1998). Thus, teacher behaviors are targeted as a vehicle for modifying children's social interactions and behaviors (Webster-Stratton, 2000).

Videotape modeling is based on modeling theory (Bandura, 1986), which posits that educators can improve their instructional skills by watching videotaped examples of teachers interacting with students in ways that promote social competence and decrease inappropriate behaviors. Efforts are made to promote modeling effects for trainees by creating positive feelings about the teachers shown on the videotapes. For example, teachers and children of different genders, ages, and cultures are depicted in unrehearsed situations (e.g., circle time, transitions, and free play) so that teachers will perceive these models as similar to themselves and their students. It is also theorized that the vignettes, which illustrate children responding to teachers positively and negatively (e.g., engagement and compliance vs. disruption and aggression) and teachers managing disruptive behaviors effectively and ineffectively, demystify the notion of "perfect teaching" and illustrate ways to learn from mistakes (Webster-Stratton, 2000).

The Incredible Years Classroom Management Program is designed to strengthen teacher classroom management strategies, promote children's prosocial behavior and school readiness skills, and reduce disruptive classroom behavior in children ages 3–10 (Webster-Stratton, Reid, & Hammond, 2001). Research in university-based settings has found this intervention to positively impact teachers' use of proactive discipline approaches, and students also exhibited greater decreases in aggression and noncompliance than children in control classrooms (Webster-Stratton et al., 2001). Although this program has been empirically supported under conditions of university-based implementation and in some school settings with extensive resources, there is limited information regarding whether this program would be equally successful when implemented under conditions of typical practice. Therefore, the purpose of the present study was to examine whether training resources (e.g., consultation) had a bearing on the transport of this program to preschools and to assess the additional contextual factors that may influence implementation. These contextual variables were examined in an effort to better match the requirements of implementing the EBI with the demands of everyday practice in schools.

METHOD

Participants

Four preschools in a medium-sized Midwestern city were recruited and selected to participate in the study based on their willingness to identify two teacher participants and agreement to random assignment to professional development conditions plus randomized start times for training. These schools shared similar national, state, and local accreditation and licensing. Two teachers from each of those four preschools ($N = 8$) expressing an interest in participating were selected. All teachers were female and European American, with the exception of one teacher who was Turkish. Teachers in the self-administered videotape modeling plus consultation group (VMC) had been employed at their schools an average of 39.3 months (Range = 18–53); teachers in the self-administered videotape modeling group without consultation (VM) were employed at their schools an average of 26.8 months (Range = 12–27). Groups were equivalent in mean years of teaching experience ($M = 9.3$). One teacher in each condition held a master's degree, while the remaining teachers possessed bachelor degrees. Thirteen students ($n = 1 - 2$ per classroom) enrolled in the eight preschool classrooms who were displaying the highest levels of challenging behaviors based on a universal screening using the *Social Competence and Behavior Evaluation—Preschool Edition* (SCBE; LaFreniere & Dumas, 1995) also participated. Students (VMC = 8; VM = 5) were included in the study if they were in the bottom 25% of the distribution of SCBE Externalizing Problems Summary Scale (T -scores ≤ 43). Demographic characteristics of student participants can be found in Table 1.

Experimental Design and Procedures

A unique variation of a multiple probe randomized design across four pairs of teachers was used. Similar to a multiple baseline design, in a multiple probe design, the independent variable is systematically and sequentially introduced to one participant at a time. However, in a multiple probe design, data are not collected as frequently, which is well suited to evaluating interventions for which continuous assessment is not feasible or when reactivity is a potential threat to validity (Tawney & Gast, 1984). Pairs of teachers from the four preschools were first randomly assigned to start times in the within-condition multiple probe design. After start times were determined, each pair of teachers was further randomly assigned to

Table 1. Demographic Characteristics of Students

Student	Gender	Age	Race/ethnicity	SCBE Externalizing T-score (Screening)
VMC				
S ₁	Male	4-3	Eur. American	41
S ₂	Male	4-2	Eur. American	42
S ₅	Female	4-6	Eur. American	43
S ₆	Male	4-6	Eur. American	43
S ₉	Male	3-10	Eur. American	38
S ₁₀	Male	3-7	Af. American	36
S ₁₃	Female	3-2	Eur. American	35
S ₁₅	Male	3-11	Eur. American	40
<i>M (SD)</i>		3-8		39.3 (2.98)
VM				
S ₃	Female	4-1	Asian	23
S ₄	Female	3-11	Latina	39
S ₈	Male	4-0	Eur. American	43
S ₁₂	Female	2-9	Asian	36
S ₁₆	Male	4-0	Eur. American	36
<i>M (SD)</i>		3-6		35.4 (7.5)

Note. SCBE = Social Competence and Behavior Evaluation-Preschool Edition; S₁ = Student 1; T-scores ≤ 43 on the Externalizing Problems Scale of the SCBE are indicative of adjustment difficulties in this sample.

one of the two training conditions. After screening and collection of baseline data, all teachers were trained in the program in a lag-order fashion. The study design included five phases (i.e., screening, baseline, training, implementation, and follow-up) over 6 months.

The Incredible Years Classroom Management Program

The professional development program covers the following five topics: (a) building positive relationships with students; (b) using praise effectively; (c) motivating students through incentives; (d) preventing behavior problems and decreasing inappropriate behavior; and (e) family outreach. Program materials included seven videotapes with over 250 brief vignettes of teachers managing challenging classroom behaviors. Video narration highlighted critical practices and strategies to address disruptive behavior. Teacher manuals, which included a recap of teacher-child interactions in each vignette, a summary of key points, and suggestions for classroom activities and homework assignments, accompanied each videotape. Chapters from the book *How to Promote Children's Social and Emotional Competence* (Webster-Stratton, 1999), which reinforced the general principles of the program, were also used.

Videotape Modeling (VM) Condition

Teachers in the VM condition were provided with a new videotape and teacher manual every 4–5 days over the course of approximately 5 weeks (seven videotapes in total). VM teachers were instructed to independently view the assigned videotapes at home or at school, review the manual, and complete the suggested classroom activities and homework assignments during that time frame. Teachers were encouraged to pace themselves and take as long as they needed to review the videotapes, manuals, and book chapters that coincided with the videotape vignettes over the 4- to 5-day time period.

Videotape Modeling Plus Consultation (VMC) Condition

VMC teachers reviewed the same videotapes, manuals, and chapters in identical sequence to VM teachers, but also participated in three 45- to 60-minute phone consultation sessions during the training phase. The consultant was a school psychology doctoral candidate with training in the program, experience with the principles and procedures of behavioral consultation and classroom management, and extensive experience working with preschoolers with disruptive behaviors. The consultation sessions followed a behavioral consultation framework (Kratochwill, Sheridan, Carlson, & Lasecki, 1999), with three standardized interviews to promote integrity to the consultation process. The Problem Identification Interview occurred during the first week of training and involved identifying strategies that were challenging to implement. The Problem Analysis Interview occurred during the third week of training and focused on identifying variables that facilitated adherence to the strategies, practicing and reinforcing newly acquired skills through implementation of strategies, and modifying instructional practices to meet the unique needs of student participants. At the end of the training phase (i.e., week 5), the Treatment Evaluation Interview was conducted to determine the extent to which plans were successful in increasing adherence to the strategies, and to modify plans to enhance maintenance and generalization.

Instrumentation

Teacher Adherence

The Teacher Strategies Questionnaire (TSQ; Webster-Stratton et al., 2001) evaluated teachers' self-reported adherence to the strategies taught in the program. The TSQ scales of interest in this study measured

teachers' confidence in managing current and future disruptive behavior (1 [*very unconfident*] to 7 [*very confident*]) and frequency in using positive and proactive instructional practices (1 [*rarely/never*] to 5 [*very often*]). In previous evaluations of this program using the TSQ, alpha coefficients ranged from .52 to .97 (Webster-Stratton et al., 2001). Teachers completed the TSQ during baseline, implementation, and follow-up phases. TSQ data along with time logs completed by teachers during the training and implementation phases were used as proxy measures of intervention integrity.

Student Behavior

The primary method of evaluating student outcomes included a behavioral observation system adapted from the Early Screening Project (ESP; Walker, Severson, & Feil, 1995; McGoey & DuPaul, 2000). Undergraduate students blind to training conditions, goals, and hypotheses of the study conducted student observations. Observers underwent extensive training by the investigator using videotapes and on-site observations of preschool children, and were trained to a criterion of 95% agreement in coding intervals during training sessions.

Behavior categories and definitions were adapted from the ESP and corresponded with the established rules of each classroom (McGoey & DuPaul, 2000). Four categories of inappropriate behaviors were coded: (a) negative social engagement (i.e., negative exchange of social signals that were verbal or physical between two or more people), (b) off-task behavior (i.e., child looks away from the activity or teacher for 3 seconds or more), (c) disobeying rules, and (d) tantrumming (i.e., yelling, kicking, and/or sulking after a social interaction). Using a 15-second partial interval recording procedure, these categories of behaviors were used to create a percentage of intervals of disruptive behavior. Each observation session was 15 minutes in duration. Due to the nature of the study design, focus students were observed three to five times during baseline, twice during training, and six times during the implementation phase (see Figure 1). Interobserver agreement was assessed on 15% of the observational sessions across participants and phases. Observers falling below 80% reliability were retrained to a criterion of 95% agreement. Interrater agreement was calculated for each behavior category using a point-by-point method, dividing the number of agreements by the sum of agreements plus disagreements, multiplied by 100. Agreement ranged from 86% to 100%, with an overall mean agreement of 99%. Mean percent agreement for each behavior category was as follows: negative social

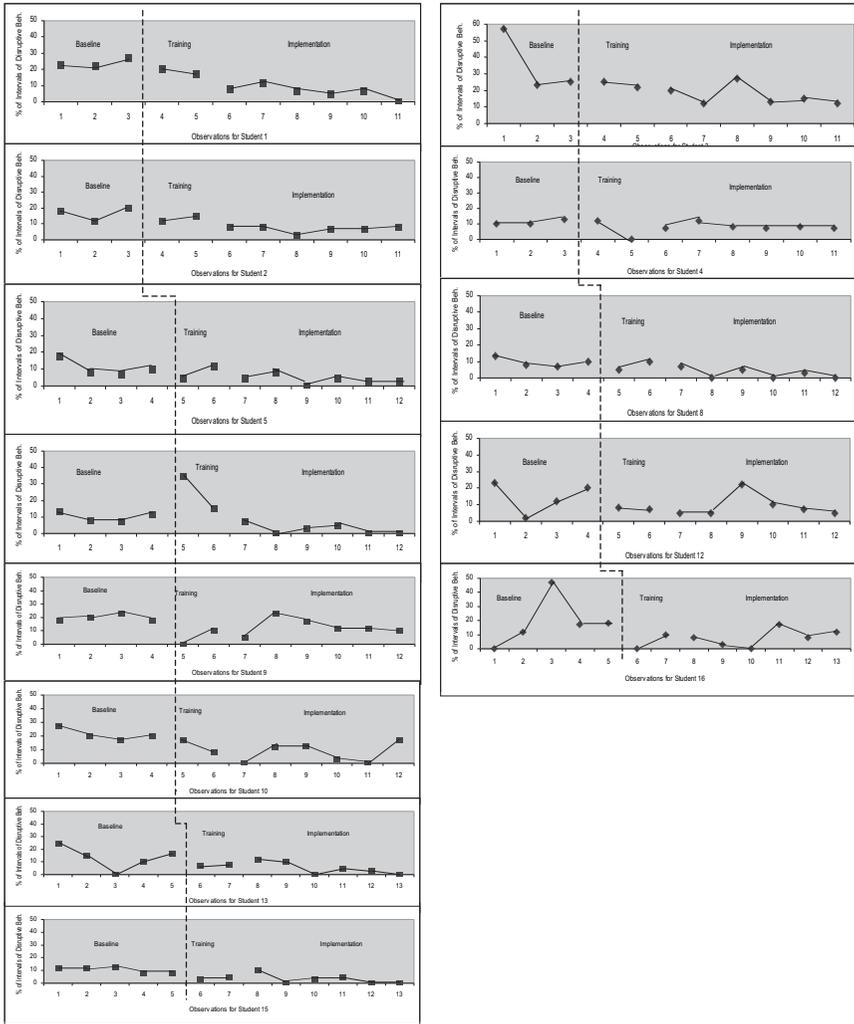


Figure 1. Percentage of intervals of disruptive behavior for VMC and VM students per observation session.

engagement = 98%; off-task behavior = 99%; disobeying rules = 99%; and tantrumming = 99%.

The *Social Competence and Behavior Evaluation–Preschool Edition* (SCBE; LaFreniere & Dumas, 1995) measures social competence and adjustment difficulties in children from preschool through age 8. Teachers rated the frequency with which students engaged in behaviors commonly seen in preschoolers on a 6-point scale, from 1 (*never*) to 6

(*always*). All scales on the SCBE are balanced for positive (competence) and negative (behavior problems) items. The Externalizing Problems Summary Scale was the primary scale of interest due to its applicability to a variety of child behavior problems. The Social Competence Summary Scale was also examined. All SCBE scores are expressed as normalized *T* scores ($M = 50$, $SD = 10$). The SCBE has good internal consistency, with Cronbach's alpha coefficients ranging from .80 to .89 (LaFreniere & Dumas, 1995). The SCBE was completed once at screening and follow-up.

Acceptability

The Treatment Evaluation Inventory (TEI; Kelley, Heffer, Gresham, & Elliott, 1989) is a nine-item questionnaire designed to assess teacher perceptions regarding the acceptability, social appropriateness, and effectiveness of strategies taught in an intervention. The nine items on the TEI are rated on a 5-point scale (1 [*strongly disagree*] to 5 [*strongly agree*]). The range of possible scores is 9–45, with higher scores representing greater acceptance of the professional development program. The TEI has a midpoint rating of 27, with scores at or above this level indicating adequate acceptability. The TEI has been shown to have acceptable internal consistency and is considered a valid measure of intervention acceptability (Kelley et al., 1989). Teachers completed the TEI once at follow-up.

Contextual Variables

Two versions of the Professional Development Evaluation Survey were developed. Both versions were identical except that VMC teachers also rated their experience with consultation on this survey. The first part of the survey included four Likert-type questions in which teachers rated the degree to which they agreed with statements regarding the effectiveness of the program components (e.g., videotapes, teacher manuals) in learning the strategies and the effectiveness of the program in reducing challenging behaviors (1 [*strongly disagree*] to 5 [*strongly agree*]). The second part of the survey assessed the adoptability of the program with similar response options. Open-ended questions elicited teachers' beliefs regarding the strengths and weaknesses of the program. VMC teachers also rated the responsiveness of the consultant, the effectiveness of the consultation sessions in helping them learn the strategies, and responded to four open-ended questions regarding the strengths and weaknesses of the consultation sessions. All teachers provided demo-

graphic information at the end of this survey and completed it after the training phase.

To better understand the process of implementing the strategies and how contextual and organizational factors affect individuals responsible for implementing EBIs, the Best Practices Interview, originally developed for use with clinicians in a mental health agency (Schmidt & Taylor, 2002), was adapted for use in the present study. This was a semistructured interview, consisting of nine open-ended questions regarding the factors teachers consider when adopting new interventions and the barriers to adopting EBIs. Due to the demand characteristics associated with the primary investigator conducting this interview, it was conducted by the project director with each teacher once at follow-up. The project director took meticulous notes during the interview and immediately added to those notes after completing the interview. The psychometric properties of these two instruments have not been evaluated.

The model used to calculate cost was empirically grounded in research on what constitutes effective professional development and was rooted in the education finance literature (Odden, Archibald, Fermanich, & Gross, 2003). This model used six core features of effective professional development to evaluate the cost of implementation: time, training and coaching, administration costs, materials/equipment, travel and transportation, and conference fees. Teachers and the consultant completed time logs during the training and implementation phases. The remaining five cost elements were based on economic data and interviews with participating preschools.

RESULTS

Confidence Ratings

On the TSQ at baseline, VMC teachers reported, on average, feeling “neutral/not sure” ($M = 4.0$; $SD = 1.15$) while VM teachers reported feeling “neutral” to “somewhat confident” ($M = 4.6$, $SD = 1.38$) in their ability to manage current and future behavior problems. The largest gain in confidence ratings occurred from baseline to implementation, with VMC teachers at implementation reporting, on average, feeling “confident” to “very confident” ($M = 6.6$; $SD = 0.48$) while VM teachers reported, on average, feeling “somewhat confident” to “confident” ($M = 5.5$; $SD = 0.58$). Mean confidence ratings remained relatively unchanged from implementation to follow up for both groups. A directional Mann–Whitney U test evaluated the hypothesis that after training, mean ranks for

VMC teachers would be higher, on average, than mean ranks for VM teachers. Results from this test were statistically significant ($z = 2.10, p = .028, \eta^2 = .64$). Using η^2 of .64 as a general measure of association (Serlin, Carr, & Marascuilo, 1982), 64% of the total variation in confidence ratings could be attributed to the training condition. Results from all nonparametric analyses should be interpreted with caution, as the z approximation gives satisfactory accuracy only with sample sizes greater than 10 per group (Kirk, 1990).

Use of Proactive Instructional Strategies

VM teachers reported, on average, more frequent use of the proactive instructional strategies taught in the program at baseline ($M = 3.3; SD = 0.21$) when compared to VMC teachers ($M = 2.7; SD = 0.49$). After completing the program, VMC teachers had greater mean increases in use of those proactive strategies ($M = 3.7; SD = 0.14$) than VM teachers ($M = 3.4; SD = 0.19$). Results from the Mann–Whitney U test were significant ($z = 2.03, p = .028, \eta^2 = .59$), suggesting statistically significant differences in mean use of positive instructional strategies at intervention in favor of the VMC group.

Student Behavior: Within-Group Analysis of Observation Data

A nonparametric randomization test provided a statistical measure of the training program's impact on students' level of challenging behavior. The Wampold and Worsham (1986) randomization test was used to analyze the dependent variable (i.e., the percentage of intervals of disruptive behavior per observation session). Due to the low occurrence of disruptive behavior for each type of inappropriate behavior, data were combined across categories to create a single percentage of intervals of disruptive behavior per observation session. The randomization test evaluated the prediction that the summed rates of disruptive behavior for students, ignoring group assignment, would be lower after teachers were exposed to the entire program (e.g., implementation) than during baseline. The proportion of 24 randomly sampled data permutations giving a combined summed rate of disruptive behavior in the predicted direction as large or larger than the experimentally obtained W was .042 for both VMC and VM students. Thus, observed differences between baseline and intervention phases were statistically significant ($p < .05$; one-tailed).

Student Behavior: Between-Group Analyses of Observation Data

Although the randomization test helps establish intervention effects by documenting when the baseline ended and the intervention was introduced, a significant result does not allow for any inferences regarding which students were most affected, or which of the two training conditions was more effective in reducing challenging classroom behavior (Todman & Dugard, 2001). Figure 1 (left column) and results in Table 2 display the behavioral observation data for VMC students at baseline, training, and implementation phases. Visual analyses allowed the comparison of students' disruptive behavior during baseline and implementation phases (after teachers were fully exposed to the program), when we expected the greatest decrease in challenging behavior. Figure 1 depicts positive intervention outcomes for students 1, 2, and 10. For example, during baseline, disruptive behavior followed an upward trend, with mean intervals of disruptive behavior equaling 24% for student 1 (Range = 22–27%), 16.7% for student 2 (Range = 12–20%), and 21% for student 10 (Range = 17–27%). When comparing disruptive behavior during the shift from baseline to intervention, there was a decrease to a mean of 6.5% for student 1 (Range = 0–12%), 6.8% for student 2 (Range = 3–8%), and 7.5% for student 10 (Range = 0–13%). Visual analysis for student 9 suggests that the strongest effects occurred at the shift from baseline (19.8%) to training (5%). Lower mean level of disruptive behavior at the start of the study for student 5 ($M\%$ = 10.8; Range = 7–18%), student 6, ($M\%$ = 10; Range = 7–13%), and student 15 ($M\%$ = 10.6; Range = 8–13%), and a downward (positive) trend detected during baseline for students 5 and 13 made it more difficult to distinguish treatment effects for these participants.

Table 2. Behavioral Observation Data at Baseline, Training, and Intervention Phases

Participant	Baseline	Training	Intervention
VMC			
S ₁	24 (2.65)	18.5 (2.12)	6.5 (3.94)
S ₂	16.7 (4.16)	13.5 (2.12)	6.8 (1.94)
S ₅	10.8 (5.0)	8.5 (4.95)	4 (2.68)
S ₆	10 (2.94)	25 (14.14)	2.5 (3.02)
S ₉	19.8 (2.36)	5 (7.07)	13.7 (6.18)
S ₁₀	21 (4.24)	12.5 (6.36)	7.5 (7.4)
S ₁₃	13.4 (9.24)	7.5 (.71)	5 (5.06)
S ₁₅	10.6 (2.41)	4 (1.41)	3 (4.0)
VM			
S ₃	35 (19.08)	23.5 (2.12)	16.5 (5.96)
S ₄	11 (1.73)	6 (8.49)	8.2 (1.94)
S ₈	9.5 (2.65)	7.5 (3.54)	2.5 (3.02)
S ₁₂	14.3 (9.4)	7.5 (0.71)	9 (6.66)
S ₁₆	18.8 (17.31)	5 (7.07)	8 (6.10)

Note. Values are Mean (*SD*) percentage of intervals of disruptive behavior.

Figure 1 (right column) and Table 2 display behavioral observation data for VM students. Fluctuations in performance during baseline for students 3 and 16 and the presence of outliers makes prediction of future performance challenging. Nonetheless, visual analysis for student 3 suggests large decreases in mean percentage intervals of disruptive behavior from baseline ($M = 35\%$; Range = 23–57%) to intervention ($M = 16.5\%$; Range = 12–27%). More modest decreases were noted for student 16 from baseline ($M = 18.8\%$; Range = 0–47%) to intervention ($M = 8\%$; Range = 0–17%). Students 4 and 8 displayed low levels of disruptive behavior at baseline ($M = 11\%$ and 9.5% , respectively) but for student 8, mean percentage of disruptive behavior decreased to 2.5% (Range = 0–7%) during intervention, while the mean rate of disruptive behavior only decreased to 8.2% (Range = 7–12%) for student 4. Student 12 had modest decreases in disruptive behavior from baseline ($M = 14.3\%$; Range = 2–23%) to intervention ($M = 9\%$; Range = 5–22%).

Effect sizes further quantified intervention outcomes for participants and allowed for comparisons of intervention effects between training conditions. Mean effect sizes were larger for the VMC condition ($M = 2.86$; $SD = 1.72$) than VM condition ($M = 1.29$; $SD = 0.87$), supporting more positive intervention outcomes for VMC students.

Student Behavior: Rating Scale Data

Clinical significance was determined by comparing the proportion of students in each condition at or below the cutoff on the SCBE at screening (T score ≤ 43) on the Externalizing Problems Scale who fell in the normal range at the end of the study, and by computing a reliability change index score (RCI). RCI scores $+1.96$ ($p < .05$) were considered statistically significant. Externalizing T scores for the VMC students at screening ranged from 35 to 43, while T scores at the end of the study ranged from 36 to 51. For VM students, Externalizing T scores at screening ranged from 23 to 43, and postintervention scores ranged from 34 to 44. Few negative RCI scores were observed overall, and for the VMC students ($n = 8$) with clinically significant Externalizing Problems T scores at screening, 50% had significant positive RCI scores and Externalizing Problems T scores that fell in the average range at the end of the study. All VM students ($n = 5$) similarly had T scores ≤ 43 at screening, and 40% had significant positive RCI scores and Externalizing T scores falling in the average range at the end of the study.

The Social Competence Summary Scale data suggest more positive outcomes for VMC than VM students, with a greater percentage of VMC

students (75%) having significant positive RCI scores regardless of clinically significant problems at screening. Thus, a greater percentage of VMC than VM students were rated as making significant positive gains in their social competence from pre to post. Furthermore, 60% of the VMC students as opposed to 0% of the VM students with clinically significant *T* scores at screening had significant RCI scores.

Acceptability

TEI ratings for the VMC condition ranged from 35 to 42 while TEI ratings for the VM condition ranged from 27 to 35. All TEI scores were ≥ 27 , indicating that overall the strategies were viewed by all teachers as acceptable and appropriate for children with behavior problems. Furthermore, the prediction that VMC teachers would have significantly higher acceptability ratings and more positive evaluations of the components of the professional development program than VM teachers was supported. The mean postintervention TEI score for the VMC condition was 38 ($SD = 3.16$), while the mean postintervention TEI score for the VM teachers was 32 ($SD = 3.83$). The between-groups comparison of post intervention acceptability ratings was statistically significant ($z = 2.07$; $p = .028$, $\eta^2 = .49$).

Teacher Perceptions of the Program

Based on responses to the Professional Development Evaluation Survey, teachers in both groups ($n = 6$) appreciated the program's positive approach and teachers ($n = 4$) also valued learning more effective methods for working with challenging students. Across both groups, teachers reported concerns regarding time demands ($n = 3$) and that some of the materials targeted children older than those in their classrooms ($n = 4$). VMC teachers ($n = 3$) reported "strongly agreeing" and one teacher reported "agreeing" that consultation increased their knowledge of how to implement the instructional practices ($M = 4.8$; $SD = 0.50$) and that the consultant addressed questions not covered or not understood from the program ($M = 4.8$; $SD = 0.50$). All VMC teachers reported "strongly agreeing" that consultation increased their ability to apply difficult strategies. Qualitative data indicated that VMC teachers valued opportunities to discuss strategies, clarify practices, and tailor the program to student's individual needs. Teachers also reported that feedback "gave me encouragement," "helped me process the information since I was working alone," and "kept me on track with the

training.” Teachers overwhelmingly perceived time as the most significant barrier to participating in consultation.

Contextual Variables

Because a primary goal of the Best Practices Interview was to capture teachers’ experience with the intervention, qualitative research methods were used to analyze interview data. Thus, the first author independently reviewed interview data and conducted a content analysis, with the goal of identifying patterns and themes across respondents. Whenever possible, these data were quantified. Table 3 summarizes results from the Best Practices Interview ordered from most to least frequently cited teacher responses collapsed across condition. Table 3 shows that time and the content of the program were the most frequently cited barriers to learning the EBI. Because assistant teachers were not trained, the most significant barrier to implementing the strategies was the added responsibility of training coteachers. Lack of congruence between the practices emphasized in the program (e.g., use of tangible rewards and time out) and those practices valued in their preschool were also frequently cited barriers to implementing the instructional strategies.

Table 3 shows that the most critical factor that teachers believed administrators should consider when adopting EBIs was whether there

Table 3. Best Practices Interview Data

Teacher response	N
Barriers to learning the EBI	
Time	6
Content of Program	5
Lack of Motivation	2
Independent format of training	1
Barriers to implementing the EBI	
Co-teachers not trained	4
Mismatch in philosophies	4
Difficulty persisting with strategies over time	3
Training occurred too late in the year	2
Factors administrators should consider when adopting EBIs	
Need for institutional resources (training/administrator support)	7
Short-term and long-term time commitment for teachers	5
Teacher motivation to be trained	4
Teacher input	2
Consistency in philosophies	1
Factors teachers consider when adopting EBIs	
Individual needs of child, family, and classroom	6
Guarantee of short and long-term effectiveness	2

Note. Data were collapsed across training condition and teachers were encouraged to provide more than one response to open-ended questions.

were ample institutional resources and supports, including training resources (e.g., long-term supervision to implement the program well) and administrative support (e.g., flexible schedule, opportunities to troubleshoot when problems emerge, and affirmation of teacher efforts to improve practice). Teachers ($n = 6$) reported focusing on the individual needs of learners when adopting new instructional practices, with one teacher stating, “Just knowing that an intervention is evidenced-based doesn’t mean it will work in my classroom with my kids.”

Time and Cost

Despite individual variation among teachers, the hours that VMC teachers logged to complete the training ($M = 12.3$; $SD = 2.05$) was comparable to VM teachers ($M = 14$; $SD = 3.46$). Similarly, VMC teachers spent an average of 18.5 hours ($SD = 2.32$), and VM teachers spent an average of 17.3 hours ($SD = 2.56$) implementing the intervention strategies. The consultant spent 11.8 hours directly consulting with teachers (2.95 hours per teacher) and 6 additional hours (1.5 per teacher) preparing for the consultation sessions. For the cost analysis, time was estimated at 25 hours based on the total time that participating preschools allocate for yearly professional development using a paid consultant. Although the total cost for an initial VMC teacher to be trained is substantial (\$2,929.23), those costs are reduced to \$508.25 for subsequent self-administered trainings with consultation. For the VM version, the cost analysis yielded a reduction of 50% of total cost (\$1,595.23) and the total cost for subsequent teachers to be trained in the VM condition is estimated at \$328.25 per teacher, all of which would be spent on teacher time.

DISCUSSION

Major Findings

Adherence

VMC teachers reported significantly higher confidence ratings and greater use of the proactive instructional practices than VM teachers. The greatest increase in self-reported confidence ratings for both groups occurred immediately following full exposure to the program (i.e., implementation), while confidence ratings remained stable from implementation to follow-up. This finding suggests that although benefits may have been sustained short-term, the extent to which teachers’ feelings of competence are maintained long-term

is unknown. These findings are consistent with research comparing the effectiveness of self-administered videotaped modeling with and without consultation using the Incredible Years Parent Training Program (Webster-Stratton, 1990). In this study, parents who participated in consultation reported less stress around managing their children's noncompliance and exhibited significantly more positive parenting strategies than parents who did not receive consultation (Webster-Stratton, 1990).

When considering the connection between consultation and adherence, Multisystemic Therapy (MST) research suggests that model-consistent supervision and consultation with change agents helps maintain fidelity to the MST approach (Henggeler, Schoenwald, Liao, Letourneau, & Edwards, 2002). Because the goals of consultation in the current study were to promote adherence and problem-solve around implementation challenges, it is hypothesized that consultation may have increased teachers' confidence and perceived self-efficacy in responding to new and challenging situations that arose when the consultant was not present.

Student Outcomes

Findings related to reductions in disruptive behavior were mixed. Results from the randomization test provided evidence that regardless of consultation, teacher exposure to the self-administered videotape modeling program was associated with reductions in challenging classroom behaviors post training. These results replicate prior parent training research using self-administered videotape modeling with and without consultation, which found that regardless of consultation, the self-administered videotape parent training was associated with significant improvements in maternal reports of child behavior problems (Webster-Stratton, 1990). In the current study, although there were positive trends favoring VMC versus VM students related to reductions in disruptive behavior (e.g., higher mean effect size data, positive visual analysis outcomes), there was not converging evidence that a greater percentage of VMC students had decreases in disruptive behavior than VM students. Because consultation was a low dose intervention (e.g., three 1-hr phone consultation sessions spanning 5 weeks) that was removed from teachers' daily practice, perhaps more pronounced reductions in challenging behavior among VMC students would have been observed had consultation reflected best practice in professional development and EBI training, including job-embedded training (Sparks & Loucks-Horsely, 1989) and high dosage intensive programming (Schoenwald & Hoagwood, 2001).

Despite the lack of differential effects related to reductions in disruptive behavior, there were important positive outcomes for VMC students

related to the development of social competence and increased adaptation to the school environment that were not observed for VM students. These results are consistent with previous studies which have found statistically significant increases in social competence for children in treatment versus control conditions based on teacher and parent report when the Incredible Years Classroom Management Program was combined with parent training (Webster-Stratton et al., 2001) and when the Incredible Years Parent Training Program was used as a stand-alone intervention (Webster-Stratton, 1998). In the current study, it is plausible that VMC teachers were using more of the positive strategies than VM teachers because VMC teachers' skills in recognizing such behaviors had improved and because there were more positive and prosocial behaviors to praise among VMC students.

Acceptability

Although the program was acceptable to teachers in both conditions, VMC teachers expressed significantly higher acceptability ratings than VM teachers. Previous research with the Incredible Years Parent Training Program documents significantly higher consumer satisfaction ratings among parents who received additional supervision when compared to parents who learned the strategies in the self-administered only format (Webster-Stratton, 1989). Although group training with supervision is not identical to individual administration with consultation, there is social support and feedback associated with these training formats that is absent from exclusive self-administration. Backer, Liberman, and Kuehnel (1986) argue that personal contact, supervision, and support are the most validated principles for knowledge transfer when organizations attempt to adopt innovative practices. Because adoption of an innovation is as much a social process as a technological one, the social component of the consultation sessions and opportunities for feedback and positive reinforcement may be critical to promoting adherence and acceptability.

Contextual Variables

Overwhelmingly, time was reportedly the most significant barrier to completing the training. This finding is consistent with previous research citing time constraints as the top-rated barrier to using EBIs in schools (Chafouleas, Elinoff, & Weinstein, 2003), internship training sites (Hays et al., 2002), and school psychology training programs (Shernoff, Kratochwill, & Stoiber, 2003). However, time constraints in the current study may have been an artifact of teachers completing this program on their own time

without compensation. This training format is unconventional given that most staff development includes training days built into teachers' schedules or flex-time for professional development (Odden et al., 2003). Nonetheless, time is an important element that schools must consider when balancing the cost-effectiveness of adopting EBIs. Implementation barriers included coteachers lack of exposure to the program and perceived lack of congruence between institutional philosophies regarding classroom management and some of the philosophies espoused in this program. To enhance sustainability and improve compatibility, full exposure to the program for all teachers would be critical, as well as balancing fidelity with adaptation of some of the program's key concepts.

Institutional resources and supports were reported as critical to adopting EBIs, which have similarly been identified as predictors of successful adoption of parent training programs (Schmidt & Taylor, 2002), substance abuse treatments (Simpson, 2002), and HIV prevention programs (Miller, 2001). Furthermore, organizational factors have been cited as a primary barrier to transporting EBIs to practice settings (Simpson, 2002). Schmidt and Taylor (2002) argue that institutional support is fundamental to the transportability of EBIs because providers need to see early in the process that the innovation is feasible to implement in that context.

Teacher interviews corroborate what has been documented as a key adoption and sustainability issue: scientific evidence is necessary but not sufficient for EBIs to be adopted by practitioners (Schmidt & Taylor, 2002). This finding reflects some of the fundamental differences cited between researchers and practitioners, the former being more invested in nomothetic issues, such as "does the program work" while the latter may be more focused on idiographic issues, such as "how does the program work" and "will it benefit my clients" (Backer et al., 1986). In the current study, teachers' experience with this EBI was fundamentally idiographic and understandably organized around the effectiveness of the program for their individual students. Evidence-based practice, thus, represents a challenging paradigm shift away from idiosyncratic practice methods that have largely characterized intervention practice up until this point (Howard, McMillen, & Pollio, 2003).

Cost

Although the initial VMC costs are substantial, costs are reduced by close to 83% after training materials are purchased. Without consultation, there was a 50% reduction in total cost for the first teacher to complete the training. When training costs for VMC are compared with what preschools in this sample budget yearly for training (e.g., average spending per teacher equals

\$250.00; $M_{\text{Range}} = \$200.00 - \300.00), the cost of subsequent VMC training is still, on average, twice the amount budgeted yearly for one teacher. VM training, in contrast, falls much closer to yearly expenditures budgeted for staff development. To assess the relative cost and benefit of each version of this program, future implementations would need to occur in these preschools, where implementation was strong and consultation was more potent.

Limitations

Teachers in this study were neither ethnically, racially, nor geographically diverse. Likewise, the teacher sample may not be an adequate representation of the larger preschool teacher population, as the area in which the sample was drawn was an educationally progressive community that may have influenced the knowledge base and instructional skills of teachers. The student sample size was also small, especially when comparing the sample size of the current study with previous evaluations of this program (Webster-Stratton et al., 2001). The fact that students were displaying lower baseline rates of disruptive behavior than in previous evaluations of the program may have created floor effects, making it more difficult to determine significant improvements in this small sample of students. All of these issues together impact the generalizability of these results.

The lack of independence among students nested in the same classrooms also impacts inferences that can be made regarding the application of the independent variable and student outcomes. Although problems with dependence cannot be eliminated entirely, dependency can be reduced by lengthening the time interval between observations and increasing the number of students. Similarly, because teachers in the same schools were not independent, there could have been “contamination” across teachers nested in the same schools. Intervention effectiveness may have been compromised due to diffusion of treatment, which poses a threat to the internal validity of the study. Finally, although the TSQ and time logs were proxy measures of integrity, direct observations of teachers would have strengthened the conclusions that could be drawn regarding the extent to which teachers adhered to the training protocol.

Future Directions

Training continues to be cited as a critical component to effectively transferring EBIs to practice settings (Drake et al., 2004). Research supporting a professional development program has limited value unless ed-

ucators in other settings can be trained in a cost-effective, efficient, and effective manner to carry out the intervention. In this study, consultation was explored as one method for enhancing outcomes. However, the transport of effective interventions to educational settings is a complex, ongoing process, rather than a one time event. Therefore, focusing on training is necessary, but not sufficient. Researchers have long struggled with this issue and have concluded that the adoption of innovative practices is anything but automatic (Davidson & Redner, 1988). Backer (2000) argues that the most common failure of past dissemination strategies is assuming that one factor (e.g., training) is enough to create change. Nevertheless, results from the current study provide some evidence that a promising professional development program that has been previously supported under conditions of university-based implementation with extensive resources can produce positive student and teacher outcomes when implemented under conditions of typical practice.

REFERENCES

- Backer, T. E. (2000). The failure of success: Challenges of disseminating effective substance abuse prevention programs. *Journal of Community Psychology, 28*, 363–373.
- Backer, T. E., Liberman, R. P., & Kuehnel, T. G. (1986). Dissemination and adoption of innovative psychosocial interventions. *Journal of Consulting and Clinical Psychology, 54*, 111–118.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology, 27*, 180–189.
- Chafouleas, S. M., Elinoff, M. J., & Weinstein, K. S. (2003, April). *Perspectives on the research-practice gap as related to school-based mental health services*. Paper presented at the annual meeting of the National Association of School Psychologists, Toronto, CA.
- Chambers, D. A., Ringeisen, H., & Hickman, A. (2005). Federal, state, and foundation initiatives around evidence-based practices for child and adolescent mental health. *Child and Adolescent Psychiatric Clinics of North America, 14*, 307–327.
- Coalition for Evidence-Based Policy (2002). *Bringing evidence-driven progress to education: A recommended strategy for the U.S. Department of Education*. Report of the Coalition for Evidence-Based Policy, November 2002.
- Davidson, W. S., & Redner, R. (1988). The prevention of juvenile delinquency: Diversion from the juvenile justice system. In R. H. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *14 ounces of prevention* (pp. 123–137). Washington, DC: American Psychological Association.
- Drake, R. E., Latimer, E. A., Leff, H. S., McHugo, G. J., & Burns, B. J. (2004). What is evidence? *Child and Adolescent Psychiatric Clinics of North America, 13*, 717–728.
- Good, T. L., & Brophy, J. E. (2003). *Looking in classrooms*. New York: Allyn & Bacon.
- Hays, K. A., Rardin, D. K., Jarvis, P. A., Taylor, N. M., Moorman, A. S., & Armstead, C. D. (2002). An exploratory survey on empirically supported treatments: Implications for internship training. *Professional Psychology: Research and Practice, 33*, 207–211.
- Henggeler, S. W., Schoenwald, S. K., Liao, J. G., Letourneau, E. J., & Edwards, D. L. (2002). Transporting efficacious treatments to field settings: The link between supervisory

- practices and therapist fidelity in MST programs. *Journal of Clinical Child Psychology*, *31*, 155–167.
- Howard, M. O., McMillen, C. J., & Pollio, D. E. (2003). Teaching evidence-based practice: Toward a new paradigm for social work education. *Research on Social Work Practice*, *13*, 234–259.
- Kavale, K., & Forness, S. (1999). Effectiveness of special education. In C. R. Reynolds & T. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 984–1024). New York: Wiley.
- Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, *10*, 165–185.
- Kelley, M. L., Heffer, R. W., Gresham, F. M., & Elliott, S. N. (1989). Development of a modified Treatment Evaluation Inventory. *Journal of Psychopathology and Behavioral Assessment*, *33*, 235–247.
- Kirk, R. E. (1990). *Statistics: An introduction* (3rd ed.). Fort Worth, TX: Holt, Reinhart and Winston.
- Kratochwill, T. R., Sheridan, S. M., Carlson, J., & Lasecki, K. L. (1999). Advances in behavioral assessment. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 350–382). New York: Wiley.
- Kratochwill, T. R., & Stoiber, K. C. (2002). Evidence-based interventions in school psychology: Conceptual foundations of the procedural and coding manual of Division 16 and the Society for the Study of School Psychology Task Force. *School Psychology Quarterly*, *17*, 341–389.
- LaFreniere, P. J., & Dumas, J. E. (1995). *Social Competence and Behavior Evaluation: Pre-school Edition*. Los Angeles: Western Psychological Services.
- McGoey, K. E., & DuPaul, G. (2000). Token reinforcement and response cost procedures: Reducing the disruptive behavior of preschool children with Attention-Deficit Hyperactivity Disorder. *School Psychology Quarterly*, *15*, 330–343.
- Miller, R. (2001). Innovation in HIV prevention: Organizational and intervention characteristics affecting program adoption. *American Journal of Community Psychology*, *29*, 621–647.
- Odden, A., Archibald, S., Fermanich, M., & Gross, B. (2003). Defining school-level expenditure structures that reflect educational strategies. *Journal of Education Finance*, *28*, 323–356.
- Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys: A social interactional approach* (Vol. 4). Eugene, OR: Castalia.
- Schmidt, F., & Taylor, T. K. (2002). Putting empirically supported treatments into practice: Lessons learned in a children's mental health center. *Professional Psychology: Research and Practice*, *33*, 483–489.
- Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services*, *52*, 1190–1197.
- Serlin, R. C., Carr, J., & Marascuilo, L. A. (1982). A measure of association for selected nonparametric procedures. *Psychological Bulletin*, *92*, 786–790.
- Shernoff, E. S., Kratochwill, T. R., & Stoiber, K. C. (2003). Training in evidence-based interventions: What are school psychology training programs teaching? *Journal of School Psychology*, *41*, 467–483.
- Simpson, D. D. (2002). A conceptual framework for transferring research to practice. *Journal of Substance Abuse Treatment*, *22*, 171–182.
- Slavin, R. E., & Fashola, O. S. (1998). *Show me the evidence! Proven and promising programs for America's schools*. Thousand Oaks, CA: Corwin Press.
- Sparks, D., & Loucks-Horsley, S. (1989). Five models of staff development for teachers. *Journal of Staff Development*, *10*, 40–57.
- Storch, E. A., & Crisp, H. L. (2004). Taking it to the schools: Transporting empirically supported treatments for childhood psychopathology in the school setting. *Clinical Child and Family Psychology Review*, *7*, 191–193.
- Tawney, J. W., & Gast, D. L. (1984). *Single subject research in special education*. Columbus, OH: Merrill Publishing Company.

- Todman, J. B., & Dugard, P. (2001). *Single-case and small n experimental designs: A practical guide to randomization tests*. Mahway, NJ: Erlbaum.
- Walker, H. M., Severson, H. H., & Feil, E. G. (1995). *Early screening project: A proven child find process*. Longmont, CO: Sopris West.
- Wampold, B. E., & Worsham, N. L. (1986). Randomization tests for multiple-baseline designs. *Behavioral Assessment, 8*, 135–143.
- Webster-Stratton, C. (1989). Systematic comparison of consumer satisfaction of three cost-effective parent training program for conduct problem children. *Behavior Therapy, 20*, 103–115.
- Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. *Journal of Abnormal Child Psychology, 18*, 479–492.
- Webster-Stratton, C. (1998). Preventing conduct problems in head start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology, 66*, 715–729.
- Webster-Stratton, C. (1999). *How to promote children's social and emotional competence*. Thousand Oaks, CA: Paul Chapman Publishing Ltd.
- Webster-Stratton, C. (2000). The Incredible Years Training Series. *Juvenile Justice Bulletin, 1–24*.
- Webster-Stratton, C., Reid, J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in Head Start. *Journal of Child Clinical Psychology, 30*, 283–302.