Published in final edited form as:

J Consult Clin Psychol. 2009 June ; 77(3): 410–421. doi:10.1037/a0013788.

Clinical Supervision in Treatment Transport: Effects on Adherence and Outcomes

Sonja K. Schoenwald, Ashli J. Sheidow, and Jason E. Chapman

Medical University of South Carolina

Abstract

This non-experimental study used Mixed-Effects Regression Models (MRMs) to examine relations among supervisor adherence to a clinical supervision protocol, therapist adherence, and changes in the behavior and functioning of youth with serious antisocial behavior treated with an empirically supported treatment (i.e., Multisystemic Therapy), one-year post treatment. Participants were 1979 youth and families treated by 429 clinicians across 45 provider organizations in North America. Four dimensions of clinical supervision were examined. MRM results showed one dimension, supervisor focus on adherence to treatment principles, predicted greater therapist adherence. Two supervision dimensions, adherence to the structure and process of supervision, and focus on clinician development, predicted changes in youth behavior. Conditions required to test hypothesized mediation by therapist adherence of supervisor adherence effects on youth outcomes were not met, and direct effects of each were observed in models including both supervisor and therapist adherence.

There is growing recognition across all stakeholders in effective mental health services that practice context variables may affect the implementation and outcomes of empirically supported treatments; and, that proactive dissemination efforts may backfire if implementation problems attenuate the promise of treatment effects observed in clinical trials (Schoenwald & Hoagwood, 2001). Accordingly, conceptual models have proliferated that depict various levels of the practice context -- including the client, clinician, provider organization and service system -- and variables within each level thought likely, on the basis of theory and research on technology transfer, implementation, and dissemination, to influence treatment implementation and outcomes in usual care (see, e.g., Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005; Schoenwald, Kelleher, Weisz, & The Research Network on Youth Mental Health, 2008; Southam-Gerow, Marder, & Austin, 2008). Critically, federal funding agencies have recently called for research on the implementation of empirically supported treatments (ESTs) in practice settings and strategies to support such implementation (National Institute on Drug Abuse, 2004; National Institute of Mental Health, 2006).

This report focuses on a feature of the practice context that has typified efficacy trials demonstrating positive child treatment effects but not child treatment as implemented in usual care, namely clinical supervision to support the implementation of a specific treatment (Weisz, Donenberg, Weiss, & Han, 1995; Weisz, Weiss, & Donenberg, 1992). As utilized in psychotherapy training, clinical supervision appears to vary widely. The rationale, objectives,

The first author is a Board Member and stockholder in MST Services, LLC, which has the exclusive licensing agreement through MUSC for the dissemination of MST technology.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/journals/ccp.

and desirable ingredients of supervision engender considerable published discussion and debate, largely absent an evidence base to inform the discussion (Davy, 2002). Reviews of research on clinical supervision in training note studies examining the impact of supervision on client outcomes is scant, with nine studies published from 1981 - 1993 (Ellis & Ladany, 1997) and three more published from 1993 – 1997 (Frieitas, 2002). The majority of these studies are characterized by a lack of theory and serious methodological problems (Freitas, 2002; Ellis & Ladany, 1997). Chief among these problems are the lack of validated measures of supervision, aspects of therapy it is meant to affect, and poor definitions and measures of client outcomes (client satisfaction is the most often cited “outcome”); confounds in study design; extremely small sample sizes of supervisors and trainees; and, the conduct of numerous tests of statistical significance without adjusting significance levels accordingly.

A more recent review (Wheeler & Richards, 2007) expanded the scope of studies evaluated to include psychotherapy supervision in professional practice contexts, although 10 of the 18 studies in this review pertained to trainees, and 4 of these had been examined in previously cited reviews. Only 2 of the 18 studies examined the effects of supervision on clients: No effects were found in one; in another, client change in sessions was observed over time in a single case design of the same therapist-client dyad (Milne, Pilkington, Gracie, & James, 2003). A recent review of supervision in medicine suggested direct supervision of medical residents (i.e., presence of supervising physician) is associated with better patient physical health outcomes (Kilminster & Jolly, 2000). The method of in vivo supervision common in residency training, however, is a method that differs considerably from most psychotherapy supervision.

Currently, clinical supervision in mental health services is among several common approaches to ensure service quality (others include licensure of individual practitioners and accreditation of organizations) that prevails largely without the benefit of evaluation or despite mixed evidence of their impact on clinicians or consumers (Bickman, in press). Whether supervision methods used in clinical trials should be replicated in usual care settings is an empirical question. To the extent such supervision methods characterize conditions in which positive treatment effects emerge (i.e., implementation in efficacy trials that generated positive results), it is logical to consider replication of the methods a necessary, though not sufficient, ingredient in transport of the treatment to usual care settings. If, however, the “acid test” of good supervision is client outcomes, as suggested by Ellis and Ladany (1997, p. 485), then empirical evidence of associations between supervision practices and client outcomes are needed.

Theory and research pertinent to the proposition that specialized supervision in the workplace enhances performance has often been generated outside of psychotherapy and mental health services research. Descriptive, quasi-experimental, and experimental studies of management practices in several industries suggest the importance of proactive strategies to sustain on-the-job performance of new skills and complex tasks (see, e.g., Burke & Baldwin, 1999). Research on increasing physician use of evidence-based medical procedures in the United States indicates training with follow-up in the workplace is more effective than training alone (for reviews, see Davis et al., 1999; Grimshaw et al, 2001; Grol & Grimshaw, 1999). In the United Kingdom, owing in part to government policies designed to support the implementation of evidence-based health care, training and evaluation of mental health clinical supervision in the workplace has been ongoing for some time. Kolb's multi-component theory of experiential learning has informed one program of research evaluating the effects of clinical supervisors on implementation of evidence-based mental health practices in the workplace, using observational and quasi-experimental studies (see, e.g., Milne, Dudley, Repper, & Milne, 2001; Milne & James, 2002; Milne, Westerman, & Hanner, 2002). Kolb's model proposes the acquisition of skills and understanding is optimized when reflection, conceptualization, planning, and practical experience occur within structured learning environments, and that a

facilitator such as a supervisor or consultant is needed to help the learner through the experiential learning cycle (Kolb, 1984).

Results of a recent study of the effects of workplace-based supervision on mental health nurses are consistent with this proposition. In this quasi-experimental study, nurses were trained to implement a new psychosocial intervention for adults and assigned either to receive workplace supervision in the intervention or not. Among the nurses with workplace supervision, there were significantly greater increases in knowledge and attitudes towards consumers, and significantly greater reductions in psychiatric symptoms of consumers receiving services, compared to nurses without supervision (Bradshaw, Butterworth, & Mairs, 2007). Results of a recent national survey in the U.S. suggest there may be some compatibility between the methods of clinical supervision found in community-based mental health clinics and those found in treatment efficacy trials (Schoenwald, Chapman et al., 2008). The extent to which supervision as practiced in such trials is replicable and effective in usual care, however, remains largely unknown.

The current report is based on data from the first NIMH-funded multi-site study of practice context factors affecting the implementation and outcomes of an evidence-based treatment for youth, Multisystemic Therapy (MST; Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998) in usual care settings. Specifically, the study, hereafter referred to as the MST Transportability Study, used a non-experimental short-term prospective longitudinal design to examine relations among select aspects of the organizational context (climate and structure), clinical supervision, therapist adherence, and youth outcomes for MST programs in 45 sites across 12 states and Canada. Prior publications reported relations among organizational context, therapist adherence, and youth discharge and short-term behavior and functioning (Schoenwald, Sheidow, Letourneau, & Liao, 2003), predictors of therapist adherence (Schoenwald, Letourneau, & Halliday-Boykins, 2005), and linkages among expert consultation, therapist adherence, and short-term youth outcomes (Schoenwald, Sheidow, & Letourneau, 2004).

The focus of this report is on relations among clinical supervision, therapist adherence, and changes in youth behavior and functioning problems one-year post-treatment using polynomial change over time. Mediation by therapist adherence of supervisor adherence effects on youth long-term post-treatment outcomes was hypothesized on the basis of a cross-sectional study linking supervisor and therapist adherence (Henggeler, Schoenwald, Liao, Letourneau, & Edwards, 2002); and of the demonstrated linkages in the MST Transportability Study among expert consultation, therapist adherence and short-term youth outcomes (Schoenwald et al., 2004).

Method

Participants

Youth and caregivers—A total of 1,979 youth and caregivers participated. The mean age for youth was 14.0 ($SD = 2.35$), and most were male (65.0%) and Caucasian (59.5%), with 19.3% of youth identified as African American, 6.4% Asian or Pacific Islander, and 14.8% other. The majority of the youth were not of Hispanic ethnicity (92.7%). Almost half of youth resided with their mother or mother and a significant other (48.6%), with remaining youth residing with both parents (15.5%), their father or father and a significant other (7.1%) or alternating between parents' households (.3%), in special living arrangements (16.9%), with a foster family (3.3%), or in other non-institutional settings (8%). The primary referral sources to treatment for youth were juvenile justice or corrections agencies (44.2%), social services (23.0%), mental health agencies (17.6%), or other agencies (15.1%). The most frequent referral reasons (multiple reasons could be endorsed for a given youth) included status offenses

(47.4%), criminal offenses (46.7%), substance use problems (31.3%), and school suspensions or expulsions (29.8%).

Mean caregiver age was 40.8 years ($SD = 8.48$), and most were female (87.8%). Most caregivers were Caucasian (65.0%). Remaining caregivers were African American (18.9%), Asian or Pacific Islander (6.6%), or other (9.6%). Most caregivers were not of Hispanic ethnicity (94.3%). Over half completed high school (66.2%), and one-third completed some college (33.9%). Half of caregivers (50.0%) reported annual incomes under \$20k. Thus, the community-based sample resembled samples in randomized trials of MST for serious antisocial behavior.

Clinicians (therapists and supervisors)—Primary therapists ($n = 429$) were identifiable for 1,888 (95.4%) of families. “Primary therapist” signifies the therapist treating the family for the entire treatment episode or, for families treated by more than one therapist, the therapist providing treatment for the majority of the family's treatment episode. The 91 families without a primary therapist were treated by more than one therapist for approximately equal lengths of time. Differences between families with and without a primary therapist have not been found (Schoenwald, Chapman, & Sheidow, 2006). The majority of therapists were female (74%) and held master's degrees (61%; 32% held bachelor's degrees; 3% held doctoral degrees; 3% held unspecified degrees). The primary clinical supervisor ($n = 122$) for each family's primary therapist *during the family's treatment episode* was identifiable for 1,736 of the 1,888 families (92%) with an identifiable primary therapist. The majority of supervisors were female (78%) and most held master's degrees (73%; 19% held bachelors degrees; 7% held doctoral degrees). Of the 122 clinical supervisors, 59 (12%) were also primary therapists for at least one family.

Procedures

Study procedures have been detailed previously (see Schoenwald et al., 2003; Schoenwald et al., 2004) and are briefly described here. All youth referred for MST treatment at the study sites were eligible except youth with autism or severe mental retardation. Families were recruited for study participation by clinical supervisors or therapists at the provider organizations upon referral, prior to treatment. The family consent rate was 82%, and only two therapists declined to participate. Informed consent from therapists and supervisors was obtained by investigators during site visits and via telephone for therapists employed after the study began. Research assistants administered pre, post, and follow-up assessment measures to caregivers and therapists by telephone; as well as the therapist adherence measure to caregivers. The research assistants also obtained youth and family demographic information during the pre-treatment assessment. Therapists reported on referral information (agency referring the youth, reasons for referral) and discharge circumstances (who made the discharge decision, reasons for discharge) on standardized forms used by MST programs. Caregivers were reimbursed for completed assessments. Longer-term follow-up data (2 – 4 years post-treatment) on youth criminal activity were obtained from court, juvenile justice, and criminal justice archives.

Family, therapist, and supervisor participation in the study was voluntary and the Institutional Review Board of the university approved all procedures.

Clinical intervention—Because details of the clinical intervention have been described elsewhere (Henggeler et al., 1998) the brief description here recaps information most central to understanding MST clinical supervision. MST is an intensive, family-based treatment originally developed for delinquent youths at imminent risk of incarceration or other out-of-home placements and their families that specifically targets those factors in each youth's social ecology (family, peers, school, neighborhood, and community) contributing to his or her

antisocial behavior. MST treatment is informed by the social ecological theory of human behavior articulated by Bronfenbrenner (1979) and by prospective research identifying the multiple predictors of serious antisocial and related behavior in adolescents. Given the youths' imminent risk of placement, overarching treatment goals often relate to keeping the youth in the home and reducing criminal behavior. Specific goals and the interventions to achieve them are designed and implemented collaboratively with the youth's caregivers.

The combination of intervention techniques applied and the expected impact of intervention procedures vary in accordance with the circumstances of each youth and family. To balance adequate specification of the model with responsiveness to the needs and strengths of each youth and family, nine principles are used to guide the MST assessment and intervention process. Ongoing assessment and intervention proceeds in accordance with an analytic process that encourages clinicians to generate specific hypotheses about the combination of factors that sustain a particular problem behavior, provide evidence to support the hypotheses, test the hypotheses by intervening, assess the impact of the intervention, and begin the assessment process again. Interventions typically include improving specific caregiver discipline practices, enhancing family affective relations, decreasing youth association with deviant peers, increasing youth association with prosocial peers and activities, improving youth school or vocational performance, and developing an indigenous support network of extended family, neighbors, and friends to help caregivers achieve and maintain such changes. Specific treatment techniques used to facilitate these gains are integrated from those therapies that have the most empirical support, including cognitive behavioral, behavioral, and pragmatic family therapies. A home-based model of service delivery provides comprehensive and intensive clinical interventions when and where they are needed (i.e., clinicians are available 24 hours/day, 7 days/week to respond to crises). Duration and frequency of treatment sessions vary in accordance with changing circumstances, needs, and treatment progress. Teams of 3-4 MST therapists have supervisors who devote at least half their time to this role and receive training in the MST supervisory protocol from expert consultants. Each therapist carries a caseload of 4 – 6 families, and treatment length averages 4-6 months. Average length of treatment for the current sample was 22.2 weeks ($SD = 10.4$).

Implementation protocol—As detailed elsewhere (Henggeler & Schoenwald, 1999; Schoenwald, 2008), a comprehensive quality assurance system designed to replicate procedures and resources provided therapists in randomized trials is used to support MST transport. Early work with communities seeking to import MST indicated systematic implementation support at several levels of the practice context would be essential to achieving the outcomes obtained in MST research; as workforce, organizational, and fiscal exigencies challenged adherence to MST clinical protocols. These experiences were consistent with theory and research on technology transfer (e.g., Backer, David, & Soucy, 1995), organizational implementation of innovations (e.g., Klein & Knight, 2005; Klein & Sorra, 1996), and Kolb's (1984) experiential learning theory, which converge on the notion that differentiated but coordinated strategies are needed to enable individuals and organizations to effectively and consistently implement a new technology. The resulting quality assurance system of six elements: (a) Site assessment; (b) 5-day orientation for therapists and clinical supervisors; (c) on-site clinical supervision guided by a the supervision protocol (Henggeler & Schoenwald, 1998); (d) weekly consultation with an MST expert trained in a consultation protocol (Schoenwald, 1998); (e) quarterly booster training; and (f) feedback on measures of therapist and supervisor adherence to MST protocols. In addition, ongoing organizational consultation and semi-annual formal program reviews are provided to assess and address organizational, service system, and purveyor (i.e. MST expert-related) barriers to achieving integrity of treatment implementation and favorable youth outcomes. This system is deployed through MST Services, LLC, a university-licensed technology transfer organization.

Clinical supervision—Weekly on-site group supervision is designed to help therapists apply the MST analytic process and treatment principles to the ongoing assessment and treatment process with each family, and to select, adapt, and integrate the appropriate treatment techniques accordingly. The MST model of supervision evolved via efficacy trials of MST in which doctoral student therapists received supervision from model developers and subsequent effectiveness trials in which master's level, community-based practitioners deployed MST with model developer support. The supervision protocol is specified in a manual that describes: underlying assumptions about the value of clinical supervision for professional practitioners undertaking a complex, empirically-supported treatment such as MST; objectives of supervision; recommended structure and process of supervision; strategies to support therapist adherence to MST principles; use of the analytic process to conceptualize case progress and setbacks; and guidelines for supporting clinician development (Henggeler & Schoenwald, 1998).

MST supervision is conducted in a group format (as a team) weekly, for 1-2 hours. Before meeting, therapists provide a weekly case summary for each family. The supervisor reviews these and identifies clinical priorities in advance of the meeting. Summaries outline, in a format mirroring the MST analytic process, progress and barriers to the attainment of specific treatment goals that week, plans to overcome barriers to previous goals, and goals for the coming week. Frequency and duration of MST group supervision occasionally varies to address exigencies, and individual supervision may occur. As in MST treatment sessions, the objectives of additional supervisory meetings and means to meet those objectives are clearly identified.

Measures

Child Behavior Checklist (CBC; Achenbach, 1991)—Youth behavior problems were assessed by the caregiver-reported CBC collected at pretreatment (T1), immediately post-treatment (T2), 6 months post-treatment (T3), and 12 months post-treatment (T4). The CBC is one of the best-validated measures of child behavioral functioning and has been normed with various age and ethnic groups (Achenbach, 1991; Drotar, Stein, & Perrin, 1995). The measure describes 113 behavior problem items applicable to children aged 2 to 18 years. Caregivers are asked to rate the extent to which the description is true of their child during the previous 6 months on a scale that ranges from 0, “not true,” to 2, “very often or often true.” T-scores for the broadband Externalizing and Internalizing scales were analyzed. For each of these scales, a T-score of 60 is the borderline clinical cutoff and a T-score of 64 is the clinical cutoff.

Vanderbilt Functioning Inventory (VFI; Bickman, Lambert, Karver, & Andrade, 1998)—Psychosocial functioning was assessed using the VFI. Content areas indexed by the 24-item VFI are antisocial behavior, problems at home, problems at school, problems with peers, and self-harm. Analyses of the reliability and validity of the VFI indicate adequate internal consistency (.71), concurrent validity (e.g., significant correlations with established measures in the expected directions), predictive validity (e.g., VFI scores predicted cost of treatment and use of residential care), and incremental validity (e.g., VFI scores accounted for a significant portion of variance of treatment cost and residential care after accounting for the variance accounted for by other measures) (Bickman et al., 1998). VFI probability scores are computed by summing raw item scores (0 or 1) and dividing by the number of completed items. Thus, scores can range from .00 to 1.00, and we observed a baseline mean of .42 ($SD = .20$, $Mode = .45$).

MST Therapist Adherence Measure –Revised (TAM-R Henggeler, Borduin, Schoenwald, Huey, & Chapman, 2006)—Therapist adherence was assessed monthly during treatment using caregiver reports on the TAM-R. The TAM-R is a 28-item scale developed by expert consensus to assess therapist adherence to the nine principles of MST.

The 28-item scale retains 19 of the 26 items of the original MST Therapist Adherence Measure (TAM; Henggeler & Borduin, 1992), validated in randomized clinical trials of MST. TAM items were rated on a 5-point Likert-type scale, with response options ranging from “Not at all” to “Very much.” Although caregiver, therapist, and youth reports on the measure were obtained in past MST trials, caregiver reports were the better predictors of youth outcomes (Schoenwald, Henggeler, Brondino, & Rowland, 2000). Multi-factor structures characterized by some instability across samples emerged in the first trials to use the measure (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997; Henggeler, Pickrel, & Brondino, 1999); and reliability and confirmatory factor analyses from the much larger and more diverse sample of caregivers and therapists in the MST Transportability Study supported a single-factor solution (Schoenwald, et al., 2003; Schoenwald et al., 2005). The Transportability Study also included 12 new items that indexed whether treatment sessions focused on important aspects of the youths' school, peer, and neighborhood/social support systems, consistent with the MST model. A comprehensive evaluation of the original 26 items and the 12 new items conducted using a Rasch-based approach to scale development retained 19 of the original TAM items and 9 of the new items onto a single factor and a two- rather than five-point rating scale. Consistent with psychometric evaluation of the single-factor TAM, TAM-R ratings were stable within a family's treatment episode.

Consistent with psychometric evaluation of the single-factor TAM, TAM-R ratings were stable within a family's treatment episode. TAM-R scores per administration range from 0 to 1, representing the percentage of items on which the caregiver rated the therapist as adherent. The scores for each administration were averaged by family to produce a mean level of therapist adherence experienced by a family during the treatment episode. In the Transportability sample, the mean TAM-R score was .64 ($SD = .26$), with observed scores ranging from 0 to 1.

Supervisor Adherence Measure (SAM; Schoenwald, Henggeler, & Edwards, 1998)—The 43-item, Likert format SAM was developed by expert consensus and is based on the rational constructs of supervision described in the MST Supervisory Manual (Henggeler & Schoenwald, 1998). Therapists rate their MST supervisor on the SAM at 2-month intervals. Exploratory and confirmatory factor analyses of SAM data collected from 74 MST therapists reporting on supervisors in 12 MST programs supported a three-factor solution for the SAM (Henggeler et al., 2002.). These factors were labeled: *Focus on analytic process and MST principles*, *Develop clinicians' MST competencies*, and *Expertise in MST and empirically supported treatments*. Reliabilities for the three factors ranged from .89 to .98. Some factors were associated with therapist adherence in expected directions, others not. Because the data were cross-sectional, it was not possible to determine whether supervisory focus in a specific area occurred before, or in response to, poor therapist adherence.

Repeated assessments of supervisor adherence over time and the significantly larger sample of therapists and supervisors participating in the MST Transportability Study made possible further refinement of the SAM scales and scoring procedures using the Many-Facet Rasch Rating Scale Model to evaluate separately the original four theoretical SAM scales. Results revealed there was not a substantial departure from unidimensionality for any of the scales, and thirty-seven of the original 43 items were retained (Schoenwald, Chapman, & Sheidow, 2006). These scales were: *Structure and Process of supervision (SP)*, *supervisor promotes Adherence to the MST treatment Principles (AP)*, *supervisor promotes use of the MST Analytic Process (ANP)*, and *supervisor promotes Clinician Development of the competencies needed to implement MST (CD)*. Sample items include: from the SP scale, “Case summaries were used during discussion of the cases;” from the AP scale, “Interventions discussed targeted sequences of interaction between family members;” from the ANP scale, “When interventions were not successful, discussion focused on identifying the barriers to success and actions clinicians

should take to overcome them;” and from the CD scale, “Within the past two months, the supervisor and I have set goals for my development of specific competencies in MST.”

Additionally, mixed-effects regression models were used to partition the variance in SAM ratings among the items, the therapists providing ratings, and the supervisors referenced by the ratings. According to the formulas provided by Raudenbush and Bryk (2002), item reliability within therapists for CD, SP, ANP, and AP was .58, .42, .88, and .88, respectively; and reliability of therapist average scores within supervisors was .58, .54, .57, and .55, respectively. Across SAM scales, 8%-36% of variance was attributable to the therapist providing the rating, and 6%-14% of variance was attributable to the supervisor referenced by the ratings. The implications for data analysis are described in the *Data structure and statistical models* section below.

Data Analysis Strategy

Missing data—Of the 1,736 families with one of 122 primary supervisors, 101 (6%) families did not have an available TAM-R score, 239 (14%) families did not have available SAM data, and 323 (19%) and 314 (18%) families did not complete any CBC and VFI administrations, respectively. Across the missing data categories (i.e., TAM-R, SAM, CBC, and VFI), a total of 323 (19%) or 314 (18%) of the 1,736 families were missing at least one type of data required for analysis of the CBC outcomes or the VFI outcome, respectively. Across the four measurement occasions, 15%, 7%, and 2% of families with useable data were missing 1, 2, or 3 CBC administrations, respectively; and 12%, 7%, and 2% were missing 1, 2, or 3 VFI administrations, respectively. At each measurement occasion, 5%, 2%, 10%, and 27% of the CBC administrations were missing, and 0%, 3%, 13%, and 19% of the VFI administrations were missing. Following the approach described by Hedeker and Gibbons (1997), models were computed to estimate the degree to which attrition cases differed from non-attrition cases on key outcomes and covariates, and no differences were found. The statistical models detailed subsequently made use of all available CBC and VFI data.

Data structure and statistical models—The present outcome data were analyzed according to a three-level Mixed-Effects Regression Model (MRM; Raudenbush & Bryk, 2002) with four repeated measurements of caregiver-reported youth Externalizing behavior problems, Internalizing behavior problems, and functioning problems (level-1) nested within ~1,422 youth/caregivers (level-2) who were nested within ~119 supervisors of families' primary therapists (level-3). Change over time was modeled using linear and quadratic polynomial terms. Because of variability in the spacing of assessments, the linear term was computed as months between the first assessment and subsequent assessments, and the quadratic term was computed as the square of the linear term (Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004). When combined in the same model, the linear term captures the instantaneous rate of change and the quadratic term captures the acceleration of change over time (Singer & Willett, 2003). The magnitude and direction of each term determines the shape of the trajectory, with, for example, combined negative linear and positive quadratic terms indicating a decelerating negative slope where initial change occurs more rapidly and then slows over time (Hedeker & Gibbons, 2006). Models with therapist adherence as the outcome were analyzed according to a two-level MRM with youth/caregiver-level TAM-R scores (level-1) nested within supervisors (level-2).

One additional feature of the data has implications for the data analysis strategy: SAM scores can be conceptualized and modeled in two key ways. First, there is a level of supervisor adherence that each youth's primary therapist experienced during the youth's treatment episode. This is referred to as the youth-level SAM score. Second, there is a level of supervision adherence that is typical or characteristic of each supervisor, and this is referred to as the

supervisor-level SAM score. A model that incorporates and individually evaluates the effect of SAM scores at both youth and supervisor levels is known as a frog-pond model (Enders & Tofighi, 2007). This strategy is common in the organizational literature where individual reports are provided with reference to a common entity (Raudenbush & Bryk, 2002, pp. 139-141; Klein & Kozlowski, 2000). This type of model is a flexible approach that permits simultaneous evaluation of the effect of the supervisor's characteristic level of adherence as well as the effect of the level of supervision received by the youth's primary therapist during the youth's treatment episode. Accordingly, the youth-level SAM scores were centered around the group (i.e., supervisor) mean SAM score, and the supervisor-level SAM scores were centered around the grand mean SAM score. The youth-level scores were entered in the level-2 equations for the intercept and the linear and quadratic terms, and the supervisor-level scores were entered in the level-3 equations for each of the level-2 intercepts (i.e., initial status, linear, and quadratic).

For each analysis, the model building approach detailed by Singer and Willett (2003) was used for specifying fixed and random effects, and all MRMs were performed using HLM software (version 6.04; Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004). Outcomes were evaluated using Restricted Maximum Likelihood estimation, and robust standard errors were used for the computation of the Wald (i.e., *T*-ratio) test statistic for the fixed effects (Maas & Hox, 2005). As described by Raudenbush and Bryk (2002), Hox (2002), and others, robust standard errors lead to accurate significance tests in the presence of non-normality, outliers, and model misspecification. Further, the upper-level sample size exceeded the minimum required for interpretation of test statistics based on robust standard errors. In each longitudinal model, initial status (i.e., the score on the outcome at the start of treatment) was allowed to vary randomly across youth/caregivers and across supervisors. A random effect was modeled for each of the level-1 fixed effects (i.e., the effect was allowed to vary across youth/caregivers and/or supervisors) and level-2 fixed effects (i.e., the effect was allowed to vary across supervisors) when it improved model fit according to the likelihood ratio test (Singer & Willett, 2003).

Test of indirect effects—The hypotheses underlying the present investigation imply tests of mediated, indirect, or intervening variable effects. Statistical methods exist for formal tests of mediation in single-level (e.g., Baron & Kenny, 1986; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002) and two-level (i.e., nested cross-sectional) designs (e.g., Bauer, Preacher, & Gil, 2006; Kenny, Korchmaros, & Bolger, 2003; Krull & MacKinnon, 2001). Less work, however, has been completed on tests of indirect effects for three-level longitudinal models. Thus, the method for testing multilevel mediation detailed by Krull and MacKinnon (2001) was extended to the three-level case. The product of coefficients test with asymmetric confidence limits was selected as the method for testing mediation on the basis of greater statistical power relative to the traditional “causal steps” approach (Krull & MacKinnon; MacKinnon et al.)

Three models were estimated for each SAM subscale: supervisor adherence predicting therapist adherence (Path A); supervisor adherence predicting Externalizing, Internalizing, and functioning outcomes (Path C); and supervisor adherence *and* therapist adherence predicting Externalizing, Internalizing, and functioning outcomes (Path C'). Although not a requirement of the product of coefficients approach, the effect of therapist adherence on each of the outcomes was evaluated first to establish the nature of relations among these variables.

Results

Therapist Adherence Predicting Youth Behavior and Functioning

As expected on the basis of previous analyses (Schoenwald, Carter, Chapman, & Sheidow, in press), higher levels of therapist adherence were associated with significantly greater instantaneous (i.e., linear) reductions in youth Externalizing behavior problems, $\gamma = -0.43$, $SE = .22$, $T(1,420) = -1.94$, $p = .05$, and more rapid (i.e., quadratic) early reductions in youth Externalizing behavior problems, $\gamma = 0.02$, $SE = .01$, $T(1,420) = 2.01$, $p = .04$. As illustrated in Figure 1, this translates into a reduction in T-score from 69.8 to 62.3 (7.5 points) between the start of treatment and 12-months post-treatment for the lowest observed TAM-R score and a reduction from 67.9 to 59.6 (8.3 points) for the highest observed TAM-R score. Thus, clinical levels of Externalizing behavior were reduced below clinical levels when adherence was highest, and remained above the borderline clinical level when adherence was lowest.

For youth Internalizing behavior problems, higher levels of therapist adherence were associated with significantly greater instantaneous (i.e., linear) behavior problem reductions, $\gamma = -0.49$, $SE = .23$, $T(1,420) = -2.20$, $p = .03$. This translates into a reduction in T-score from 61.6 to 55.5 (6.1 points) between the start of treatment and 12-months post-treatment for the lowest observed TAM-R score and a reduction from 61.4 to 53.1 (8.3 points) for the highest observed TAM-R score. The effect of therapist adherence on the quadratic term for Internalizing behavior problems did not reach significance. Thus, Internalizing problems among these youth referred for serious antisocial behavior were above the borderline clinical cutoff at the start of treatment and fell below the borderline clinical cutoff under conditions of higher and lower adherence, although the scores were a bit lower when adherence was better.

Therapist adherence was not significantly associated with the linear or quadratic change terms for youth functioning problems.

Supervisor Adherence Predicting Therapist Adherence (Path A)

Results of the two-level MRM for youth- and supervisor-level supervisor adherence scores predicting therapist adherence revealed that supervisor average focus on Adherence to Principles (AP) was significantly positively associated with therapist adherence, $\gamma = 0.18$, $SE = .07$, $T(117) = 2.70$, $p = .01$. This translates into a 12% higher TAM score for supervisors with the highest versus lowest observed AP score. The remaining SAM subscales (supervisor average or youth deviation) were not significantly associated with therapist adherence.

Supervisor Adherence Predicting Youth Behavior and Functioning (Path C)

Externalizing problems—Greater supervisor adherence to the Structure and Process (SP) of supervision during a youth's treatment episode (i.e., a higher youth deviation from the supervisor mean SAM score) predicted significantly greater instantaneous (i.e., linear) reductions, $\gamma = -1.45$, $SE = .66$, $T(1,420) = -2.19$, $p = .03$, and more rapid (i.e., quadratic) early reductions, $\gamma = 0.06$, $SE = .03$, $T(1,420) = 1.94$, $p = .05$, in youth Externalizing behavior problems. This translates into a reduction in T-score from 66.4 to 60.6 (5.8 points) between the start of treatment and 12-months post-treatment when supervisor adherence during the youth's treatment episode was lowest relative to the supervisor's usual level of adherence, and a reduction from 70.2 to 60.7 (9.5 points) when supervisor adherence during a treatment episode was highest relative to the supervisor's usual level of adherence. These T-score changes represent reductions in the clinical severity of Externalizing problems from just above the clinical cutoff to the borderline clinical cutoff for youth when supervisor adherence during a treatment episode was lowest, and from well beyond the clinical cutoff to the borderline clinical cutoff for youth when the supervisor level of adherence during treatment was highest.

The remaining SAM scores (youth- and supervisor-level) were not significantly associated with the linear or quadratic terms for Externalizing behavior problems.

Internalizing problems—The magnitude and direction of effects observed between SP and Internalizing problems mirrored that found between SP and Externalizing problems, but fell short of the threshold for significance ($p = .06$).

Functioning problems—Greater supervisor adherence to SP during a youth's treatment episode (i.e., a higher youth deviation from the supervisor mean SAM score) significantly predicted greater instantaneous (i.e., linear) reductions, $\gamma = -0.02$, $SE = .01$, $T(1,420) = -2.11$, $p = .04$, in functioning problems. This translates into a reduction from .35 to .23 (.12 points) between the start of treatment and 12-months post-treatment for a youth when supervisor adherence during the treatment episode was lowest relative to the supervisor's usual level of adherence, and a reduction from .41 to .20 (.21 points) when a supervisor's level of adherence during a treatment episode was highest relative to the supervisor's usual level of adherence. Thus, the proportion of youth functioning problems decreased by nearly twice as much when supervisor adherence to the structure and process of MST supervision was highest rather than lowest relative to the supervisor's average level of adherence.

For supervisors with a higher overall average focus on Clinician Development (CD), there were significantly weaker instantaneous (i.e., linear) decreases, $\gamma = 0.02$, $SE = .01$, $T(117) = 2.06$, $p = .04$, in proportion of youth functioning problems. This translates into a decrease in functioning problems from .47 to .22 (.25 points) between start of treatment and 12-months post-treatment for a supervisor with the lowest level of focus on CD adherence and only half that decrease, from .33 to .21 (.12 points), for a supervisor with the highest level of focus on CD.

Supervisor Adherence & Therapist Adherence Predicting Youth Behavior and Functioning (Path C')

As described in the "Test of indirect effects" subsection of the "Data Analysis Strategy" section, we evaluated, for each supervision subscale, the effects of both supervisor adherence and therapist adherence on each youth outcome.

Externalizing problems—Figure 2 illustrates the effect of supervisor adherence to Structure and Process during a treatment episode on reductions in youth Externalizing problems in the presence of therapist adherence. Because supervisor and therapist adherence are continuous variables, we selected the highest and lowest observed deviations on the supervisor SP, and the highest and lowest TAM-R scores to illustrate the nature of the relationships. As seen in Figure 2, when supervisor adherence to SP during a youth's treatment episode was highest (i.e., a higher youth deviation from the supervisor mean SAM score) youth experienced significantly greater instantaneous (i.e., linear) reductions, $\gamma = -1.51$, $SE = .65$, $T(1,419) = -2.31$, $p = .02$, and more rapid (i.e., quadratic) early reductions, $\gamma = 0.07$, $SE = .03$, $T(1,419) = 2.07$, $p = .04$, in Externalizing behavior problems.

For each model testing the effects of a SAM subscale and therapist adherence, higher levels of therapist adherence were associated with significantly greater instantaneous (i.e., linear) reductions, γ s ranging from -0.43 to -0.46 and p s from .04 to .05, and more rapid (i.e., quadratic) early reductions, all γ s = 0.02 and p s from .04 to .05, in Externalizing behavior problems.

Internalizing problems—Following the pattern observed in Figure 2 for Externalizing behavior problems, when supervisor adherence to SP during a treatment episode was higher (i.e., a higher youth deviation from the supervisor mean SAM score), youth experienced

significantly greater instantaneous (i.e., linear) reductions, $\gamma = -1.39$, $SE = .70$, $T(1,419) = -1.98$, $p = .05$, and more rapid (i.e., quadratic) early reductions, $\gamma = 0.07$, $SE = .04$, $T(1,419) = 2.01$, $p = .04$, in Internalizing behavior problems.

For each model testing the effects of a SAM subscale and therapist adherence, higher levels of therapist adherence were associated with significantly greater instantaneous (i.e., linear) reductions, γ s ranging from -0.49 to -0.52 and p s from $.02$ to $.03$, in Internalizing behavior problems. For the model testing the effects of SP, and the model testing the effects of AP, therapist adherence was also associated with more rapid (i.e., quadratic) early reductions, $\gamma_{210s} = .02$ and $p = .05$, in Internalizing behavior problems.

Functioning problems—When supervisor adherence to SP during a youth's treatment episode was higher (i.e., a higher youth deviation from the supervisor mean SAM score), youth experienced significantly greater instantaneous (i.e., linear) reductions, $\gamma = -0.02$, $SE = .01$, $T(1,419) = -2.11$, $p = .04$, in functioning problems. When therapists participated in supervision with a supervisor with a higher average focus on CD (i.e., the supervisor focused more on CD across all therapists supervised), the youth they treated experienced significantly weaker instantaneous (i.e., linear) decreases, $\gamma = 0.02$, $SE = .01$, $T(117) = 2.04$, $p = .04$, in functioning problems. Therapist adherence was not significantly associated with the instantaneous (i.e., linear) reductions or the acceleration of change for youth functioning problems.

In sum, when the effects of both supervisor adherence and therapist adherence on youth one-year post-treatment were examined in the same model, each exerted direct effects on reductions in youth behavior problems, but only supervisor adherence exerted direct effects on youth functioning problems. Specifically, higher supervisor adherence to the structure and process of MST supervision during a youth's treatment episode predicted greater and faster improvements in youth problem behavior and functioning. In addition, the higher average focus of a supervisor on Clinician Development (i.e., the supervisor's focus on Clinician Development across all therapists supervised) was associated with weaker improvements in youth functioning.

Discussion

This report is the first to our knowledge to examine prospectively relations among supervisor adherence to a treatment-model specific clinical supervision protocol, therapist adherence, and the long-term post-treatment outcomes of an EST for youth transported to diverse usual care settings. Two of four dimensions of MST clinical supervision, Structure and Process (SP) and focus on Clinician Development (CD), predicted long-term changes in youth behavior and functioning; and a third dimension, focus on Adherence to Principles (AP), predicted therapist adherence. Consistent with previous findings for short-term outcomes in the MST Transportability Study, caregiver-reported therapist adherence predicted long-term reductions in behavioral problems in youth. These results suggest the promise of implementing workplace-based, treatment-model specific, clinical supervision with mental health professionals for the larger-scale transport of ESTs to usual care settings. The findings also suggest the MST clinical supervision approach is on the way to meeting the “acid test” of clinical supervision – demonstrated impact on clients (Ellis & Ladany, 1997).

Supervisor focus on AP predicted caregiver-reported therapist adherence to MST principles during a treatment episode, as hypothesized. The lack of associations between therapist adherence and the other supervision subscales – Structure and Process (SP), Analytic Process (ANP), and Clinician Development (CD) -- countered expectations. Theory and experience with MST transport and implementation suggest a possible explanation for the null findings lies in the extent to which these three scales -- relative to the Adherence to Principles (AP)

scale-- embody the reflective, conceptualization, and planning functions of the experiential learning cycle proposed by Kolb (1984). Specifically, the focus of supervision on the consistency of assessment and intervention strategies with the principles of MST (i.e., focus on AP) includes discussion and practice of therapist actions that occur in treatment sessions, and which can be observed and rated by caregivers on the therapist adherence measure. In contrast, promotion of the Analytic Process (ANP) in supervision is designed to support therapist use of a scientific method of case conceptualization and planning, activities that should occur before and after each treatment session but may not be apparent in the session. The conceptualization, planning, and intervention discussion and practice (i.e., performance) functions of MST supervision are reminiscent of Kolb's learning cycle, but interleaved within the same supervision session. For this reason, and because the SAM items reference supervision and supervisors (not therapist perceptions of their own cognition and performance), it is not possible to examine relations between the supervision functions and the sequencing of therapist learning and performance.

Similarly, supervisor focus on CD is designed to shore up the specific skills a particular therapist needs to implement techniques used in MST. The CD subscale indexes supervisory efforts to address the therapist's goals, skills, and competencies; as well as therapist perceptions of the supervisor's skills and expertise. Supervisory focus on therapist development may be needed to engage therapists in ongoing workplace supervision while not translating directly into caregiver-reported adherence during sessions. Finally, the SP items index supervisor management of the weekly group supervision, in which discussion must be sufficiently focused and efficient to address the clinical priorities of 12–15 cases. Although not linked with therapist adherence, greater supervisor adherence in the SP and CD domains predicted youth outcomes.

The finding that greater average supervisor focus on Clinician Development (CD) predicted a weaker decrease in youth functioning problems could be interpreted as an indication that supervisors of therapist teams characterized by relatively lower levels of competence or skill focused more on clinician development relative to supervisors of other teams. This greater average focus of supervisors on clinician development, however, yielded limited gains with respect to youth functioning problems.

That different aspects of supervision may affect client outcomes absent their effect on therapist adherence is consistent with findings in adult psychotherapy research linking distinctive aspects of therapy (e.g., alliance, adherence, competence, therapist effects) with client outcomes. For example, adherence and competence have each been found to predict outcomes of brief dynamic therapy for adult depression (Barber, Crits-Christoph, & Luborsky, 1996) whereas competence was the better predictor of outcomes of cognitive behavioral therapy for depression (Shaw et al., 1999). Reviews indicate, however, relations between clinical competence and client outcomes are not consistent (Barber, Sharpless, Klosterman, & McCarthy, 2007). And, findings from a recent adult substance abuse treatment trial found no association between competence and outcomes and curvilinear relations between early treatment alliance and adherence (Barber et al., 2006). Finally, even absent demonstrated relations between client outcomes and specific attributes of therapists, such as experience and education or competence, therapist effects on such outcomes can be observed (Lutz, Leon, Martinovich, Lyons, & Stiles, 1997; Wampold & Brown, 2005; Kim, Wampold, & Bolt, 2006).

In contrast with adult psychotherapy research, comparatively little research has been conducted in child and family mental health treatment on either model specific processes such as adherence and competence or nonspecific processes such as alliance (Liddle & Rowe, 2006; Weisz, Doss, & Hawley, 2005). Adherence-outcomes links have been documented for one model of cognitive-behavioral treatment of adolescent depression (Kolko, Brent, Baugher,

Bridge, & Birmaher, 2000) and one model of family therapy other than MST (Hogue et al., 2008).

Finally, the mediation of the effects of supervisor adherence on youth outcomes by therapist adherence was not formally tested, given the lack of evidence linking any dimension of supervisor adherence with *both* therapist adherence *and* youth outcomes. Plausible explanations for the lack of mediation include the possibility that effects of supervision on youth outcomes are carried by other specific and nonspecific therapy factors not measured in the study (e.g., competence, alliance). Alternately, aspects of MST supervision may affect the conceptualization, planning, and skill development of therapists in ways that are not captured by caregiver reports of therapist adherence during sessions, but affect therapist effectiveness during sessions.

Study limitations

This prospective, short-term longitudinal study of the effects of clinical supervision on therapist adherence and changes in youth behavior and functioning one-year post-treatment with MST in community practice settings found modest differences in youth behavior and functioning problem reduction associated with greater and lesser supervisor adherence. For example, the average one-year post-treatment Externalizing scores for all youth had fallen from above the clinically significant cutoff at pre-treatment to below that cutoff, even when supervisor adherence was one standard deviation below the average of that supervisor's adherence during a treatment episode. On the one hand, given all youth in the sample – largely juvenile offenders at risk of out-of-home placement -- had undergone a course of MST, significant reductions in their problem behavior were to be expected on the basis of prior research. On the other hand, given the delinquent behavior and associated risk of out-of-home placement characterizing the youth in this sample, reducing clinically significant levels of Externalizing problems rapidly is a treatment priority. The current findings indicate supervisors effectively address this priority when adhering to the structure and process of MST supervision during a youth's treatment episode. We thus remain encouraged by the demonstrated capacity to prospectively link supervision with valid and reliable measures of client outcomes, a rare occurrence in the literature on clinical supervision and in the enterprise of transporting evidence-based treatments for youth to usual practice settings. And, preliminary analyses of relations among supervisor adherence, therapist adherence, and longer-term youth criminal outcomes are promising (Schoenwald, 2008).

A second limitation of the study, acknowledged in previously published reports (Schoenwald et al., 2003; 2004; 2005) is that shared method variance (i.e., caregiver report) may account for the observed associations between therapist adherence and youth behavior change. Concern about this limitation is attenuated by the fact that the caregiver reports of youth behavior and functioning at baseline preceded any adherence rating, and with the exception of the post-treatment (T2) assessment, were temporally distant (6 and 12 months following) from their reports on adherence. Recent findings linking therapist adherence and youth criminal outcomes support the predictive validity of caregiver-reported adherence absent shared method variance (Schoenwald, Chapman, Sheidow, & Carter, in press).

Finally, as noted previously, the study findings do not allow for examination of the sequencing of therapist learning and performance (other than adherence) as a function of MST supervision, and of relations among therapist learning and performance with clinical outcomes.

Implications for future research

To support the effective transport and implementation of ESTs in practice contexts, research is needed to evaluate the viability, implementation, and effects on client outcomes in usual care

settings of the training and supervision protocols used in treatment trials. Rigorous tests of the effects in practice contexts of clinician training protocols for adult substance abuse treatment are promising in this regard (e.g., Sholomskas et al., 2005); but similar studies have not yet focused either on ESTs for youth, or on clinical supervision. To the extent empirical evidence links specific aspects of the therapy process (e.g., therapeutic alliance, therapist adherence to treatment, therapist competence, client adherence to treatment), with positive treatment effects, examination of supervision effects on these processes is warranted. Evaluation of mechanisms of action by which supervision effects on client outcomes is also needed. In addition, valid, reliable, and – if they are to be used in usual care practice contexts – inexpensive and minimally burdensome measures of clinical supervision are needed, and the development and validation of such measures should take into account the data structures likely to emerge in such contexts (e.g., clients nested within therapists, therapists within supervisors, supervisors within provider organizations; repeated measures of supervision).

Implications for clinical practice

The results reported here indicate the training and clinical supervision provided to therapists in successful randomized treatment trials can be provided to mental health professionals in the workplace as part of a multi-component treatment transport strategy. The results show the use of workplace-based clinical supervision to support the implementation of a specific treatment model can affect select aspects of the therapy process (in this case, adherence) and some client outcomes. One implication of these findings for clinical practice is that, at least for complex ESTs, workplace-based, model-specific clinical supervision may be among the differentiated, but coordinated, set of practices needed to support effective implementation. Accordingly, planning for EST implementation in usual care settings should take into account not only the time and resources needed to train and support therapists, but also clinical supervisors.

Acknowledgments

Preparation of this manuscript was supported by grants DA018107 and K23DA015658 from the National Institute on Drug Abuse, grant MH59138 from the National Institute of Mental Health, The Annie E. Casey Foundation, and The John D. and Catherine T. MacArthur Foundation. The views presented in this manuscript are those of the authors alone and do not necessarily reflect the opinions of the Anne E. Casey or John D. and Catherine T. MacArthur Foundation.

References

- Achenbach, TM. Integrative guide for the 1991 CBC/4-18, YSR, and TRF profiles. Burlington: University of Vermont, Department of Psychiatry; 1991.
- Backer, TE.; David, SL.; Soucy, G., editors. Reviewing the behavioral science knowledge base on technology transfer. Rockville, MD: National Institute on Drug Abuse; 1995. NIDA Research Monograph 155, NIH Publication No. 95-4035
- Barber JP, Crits-Christoph P, Luborsky L. Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology* 1996;64:619–622. [PubMed: 8698958]
- Barber JP, Gallop R, Crits-Christoph P, Frank A, Thase ME, Weiss RD, Gibbons MB Connolly. The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research* 2006;16:229–240.
- Barber JP, Sharpless BA, Klostermann S, McCarthy KS. Assessing intervention competence and its relation to therapy outcome: a selected review derived from the outcome literature. *Professional Psychology Research and Practice* 2007;38:493–500.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 1986;51:1173–1182. [PubMed: 3806354]

- Bauer DJ, Preacher KJ, Gil KM. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods* 2006;11:142–163. [PubMed: 16784335]
- Biesanz JC, Deeb-Sossa N, Papadakis AA, Bollen KA, Curran PJ. The role of coding time in estimating and interpreting growth curve models. *Psychological Methods* 2004;9:30–52. [PubMed: 15053718]
- Bickman L. Improving mental health outcomes is not rocket science – it's harder! *American Academy of Child and Adolescent Psychiatry*. in press.
- Bickman L, Lambert EW, Karver M, Andrade AR. Two low-cost measures of child and adolescent functioning for services research. *Evaluation and Programming Planning* 1998;21:263–275.
- Bradshaw T, Butterworth A, Mairs H. Does structured clinical supervision during psychosocial intervention education enhance outcome for mental health nurses and the service users they work with? *Journal of Psychiatric & Mental Health Nursing* 2007;14:4–12. [PubMed: 17244000]
- Bronfenbrenner, U. *The ecology of human development*. Cambridge, MA: Harvard University Press; 1979.
- Burke L, Baldwin TT. Workforce training transfer: A study of the effect of relapse prevention training and transfer climate. *Human Resource Management* 1999;38:227–242.
- Davis D, O'Brien MA, Fremantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. The impact of formal continuing education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *Journal of the American Medical Association* 1999;282:867–874. [PubMed: 10478694]
- Davy J. Discursive reflections on a research agenda for clinical supervision. *Psychology & Psychotherapy: Theory, Research, & Practice* 2002;75:221–238.
- Drotar D, Stein REK, Perrin EC. Methodological issues in using the Child Behavior Checklist and its related instruments in clinical child psychology research. *Journal of Clinical Child Psychology* 1995;24:184–192.
- Ellis, MV.; Ladany, N. Inferences concerning supervisees and clients in clinical supervision: An integrative review. In: Watkins, CE., editor. *Handbook of psychotherapy supervision*. New York: Wiley; 1997. p. 447-507.
- Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods* 2007;12:121–138. [PubMed: 17563168]
- Fixsen, DL.; Naoom, SF.; Blasé, KA.; Friedman, RM.; Wallace, F. *Implementation Research: A Synthesis of the Literature*. Tampa, FL: University of South Florida, Louis De la Parte Florida Mental Health Institute, The National Implementation Research Network; 2005. FMHI Publication #231
- Freitas GJ. The impact of psychotherapy supervision on client outcome: A critical examination of 2 decades of research. *Psychotherapy: Theory, Research, Practice, Training* 2002;39:354–367.
- Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, et al. Changing provider behavior: An overview of systematic reviews of interventions. *Medical Care* 2001;39(8):II-2–II-45. [PubMed: 11583120]
- Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. *Journal on Quality Improvement* 1999;25:503–513.
- Hedeker, D.; Gibbons, RD. *Longitudinal data analysis*. Hoboken, NJ: John Wiley & Sons, Inc.; 2006.
- Henggeler, SW.; Borduin, CM. *Multisystemic Therapy Adherence Scales*. Charleston, SC: Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina; 1992. Unpublished instrument
- Henggeler, SW.; Borduin, CM.; Schoenwald, SK.; Huey, SJ.; Chapman, JE. *Multisystemic Therapy Adherence Scale - Revised (TAM-R)*. Charleston, SC: Department of Psychiatry & Behavioral Sciences, Medical University of South Carolina; 2006. Unpublished instrument
- Henggeler SW, Melton GB, Brondino MJ, Scherer DG, Hanley JH. Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology* 1997;65:821–833. [PubMed: 9337501]
- Henggeler SW, Pickrel SG, Brondino MJ. Multisystemic treatment of substance abusing and dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research* 1999;1:171–184. [PubMed: 11258740]

- Henggeler, SW.; Schoenwald, SK. The MST supervisory manual: Promoting quality assurance at the clinical level. Charleston, SC: The MST Institute; 1998.
- Henggeler SW, Schoenwald SK. The role of quality assurance in achieving outcomes in MST programs. *Journal of Juvenile Justice and Detention Services* 1999;14:1–17.
- Henggeler, SW.; Schoenwald, SK.; Borduin, CM.; Rowland, MD.; Cunningham, PB. Multisystemic treatment of antisocial behavior in children and adolescents. New York and London: Guilford Press; 1998.
- Henggeler SW, Schoenwald SK, Liao JG, Letourneau EJ, Edwards DL. Transporting efficacious treatments to field settings: The link between supervisory practices and therapist fidelity in MST programs. *Journal of Clinical Child and Adolescent Psychology* 2002;31:155–167. [PubMed: 12056100]
- Hogue A, Henderson CE, Dauber S, Barajas PC, Fried A, Liddle HA. Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology* 2008;76:544–555. [PubMed: 18665684]
- Hox, J. Multilevel analysis techniques and applications. Mahwah, NJ: Lawrence Erlbaum; 2002.
- Kazdin AE. Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist* 2008;63:146–159. [PubMed: 18377105]
- Kenny DA, Korchmaros JD, Bolger N. Lower level mediation in multilevel models. *Psychological Methods* 2003;8:115–128. [PubMed: 12924810]
- Kilminster SM, Jolly BC. Effective supervision in clinical practice settings: a literature review. *Medical Education* 2000;34:827–840. [PubMed: 11012933]
- Kim DM, Wampold BE, Bolt DM. Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research* 2006;16:161–172.
- Klein KJ, Knight AP. Innovation Implementation: Overcoming the Challenge. *Current Directions in Psychological Science* 2005;14(5):243–246.
- Klein KJ, Kozlowski SW. From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods* 2000;3:211–236.
- Klein KJ, Sorra JS. The challenge of innovation implementation. *Academy of Management Review* 1996;21:1055–1080.
- Kolb, DA. *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall; 1984.
- Kolko DJ, Brent DA, Baugher M, Bridge J, Birmaher B. Cognitive and family therapies for adolescent depression: Treatment specificity, mediation, and moderation. *Journal of Consulting and Clinical Psychology* 2000;68:603–614. [PubMed: 10965636]
- Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research* 2001;36:249–277.
- Liddle, HA.; Rowe, CL. Adolescent substance abuse: Research and clinical advances. New York: Cambridge University Press; 2006.
- Lutz W, Leon SC, Martinovich Z, Lyons JS, Stiles WB. Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology* 2007;54:32–39.
- Maas CJM, Hox JJ. Sufficient sample sizes for multilevel modeling. *Methodology* 2005;1:86–92.
- MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods* 2002;7:83–104. [PubMed: 11928892]
- Milne D, Dudley M, Repper D, Milne J. Managers' perceived contribution to the transfer of psychosocial interventions in training. *Psychiatric Rehabilitation Skills* 2001;5:387–403.
- Milne DL, James IA. The observed impact of training on competence in clinical supervision. *British Journal of Clinical Psychology* 2002;41:55–72. [PubMed: 11931678]
- Milne D, Pilkington J, Gracie J, James I. Transferring skills from supervision to therapy: A qualitative and quantitative N=1 analysis. *Behavioral and Cognitive Psychotherapy* 2003;31:193–202.
- Milne D, Westerman C, Hanner S. Behavioural and Cognitive Psychotherapy 2002;30:361–364.

- National Institute of Mental Health. The Road Ahead: Research Partnerships to Transform Services. A Report by the National Advisory Mental Health Council's Workgroup on Services and Clinical Epidemiology Research. Rockville, MD: National Institutes of Health, Department of Health and Human Services; 2006.
- Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2nd. Thousand Oaks, CA: Sage Publications; 2002.
- Raudenbush, SW.; Bryk, AS.; Cheong, YF.; Congdon, R.; du Toit, M. HLM 6: Hierarchical linear & nonlinear modeling (Version 6.02) (Computer software and manual). Lincolnwood, IL: Scientific Software International; 2004.
- Schoenwald SK. Toward evidence-based transport of evidence-based treatments: MST as an example. *Journal of Child and Adolescent Substance Abuse* 2008;17(3):69–91.
- Schoenwald, SK. Multisystemic therapy consultation manual. Charleston, SC: The MST Institute; 1998.
- Schoenwald SK, Carter RE, Chapman JE, Sheidow AJ. Therapist adherence and organizational effects on change in youth behavior problems one year after Multisystemic Therapy. *Administration and Policy in Mental Health and Mental Health Services Research*. in press.
- Schoenwald SK, Chapman JE, Kelleher K, Hoagwood KE, Landsverk J, Stevens J, et al. A Survey of the infrastructure for children's mental health services: Implications for the implementation of empirically supported treatments (ESTs). *Administration and Policy in Mental Health and Mental Health Services Research* 2008;35:84–97. [PubMed: 18000750]
- Schoenwald, SK.; Chapman, JE.; Sheidow, AJ. Implementation fidelity in MST. In: Schoenwald, SK.; Reid, J., editors. Community based model programs panel: Implementing with fidelity; Blueprints Conference, 2006: Evidence-based programs, research-to-practice conference; Denver, CO. 2006 Mar. Co-Chairs
- Schoenwald SK, Chapman JE, Sheidow AJ, Carter RE. Long-term youth criminal outcomes in MST transport: The impact of therapist adherence and organizational climate and structure. *Journal of Clinical Child and Adolescent Psychology*. in press.
- Schoenwald, SK.; Henggeler, SW.; Edwards, D. MST Supervisor Adherence Measure. Charleston, SC: MST Institute; 1998.
- Schoenwald SK, Henggeler SW, Brondino MJ, Rowland MD. Multisystemic therapy: Monitoring treatment fidelity. *Family Process* 2000;39:83–103. [PubMed: 10742933]
- Schoenwald SK, Hoagwood K. Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services* 2001;52:1179–1189. [PubMed: 11533391]
- Schoenwald SK, Kelleher K, Weisz J. Research Network on Youth Mental Health. Building bridges to evidence-based practice: The MacArthur Foundation Child System and Treatment Enhancement Projects (Child STEPs). *Administration and Policy in Mental Health and Mental Health Services Research* 2008;35:66–72. [PubMed: 18085433]
- Schoenwald SK, Letourneau EJ, Halliday-Boykins CA. Predicting adherence to a transported family-based treatment for youth. *Journal of Clinical Child and Adolescent Psychology* 2005;34:658–670. [PubMed: 16232063]
- Schoenwald SK, Sheidow AS, Letourneau EJ. Toward effective quality assurance in evidence-based practice: Links between expert consultation, therapist fidelity, and child outcomes. *Journal of Child and Adolescent Clinical Psychology* 2004;33:94–104.
- Schoenwald SK, Sheidow AJ, Letourneau EJ, Liao JG. Transportability of Multisystemic Therapy: Evidence for multi-level influences. *Mental Health Services Research* 2003;5:223–239. [PubMed: 14672501]
- Shaw BF, Olmsted M, Dobson KS, Sotsky SM, Elkin I, Yamaguchi J, et al. Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology* 1999;67:837–846. [PubMed: 10596506]
- Sholomskas DE, Syracuse-Siewert G, Rounsaville BJ, Ball SA, Nuro KF, Carroll KM. We don't train in vain: A dissemination trial of three strategies of training clinicians in cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology* 2005;73:106–115. [PubMed: 15709837]
- Singer, JD.; Willett, JB. Applied longitudinal data analysis: Modeling change and event occurrence. New York: Oxford University Press; 2003.

- Southam-Gerow, MA.; Marder, AM.; Austin, AA. Transportability and dissemination of evidence-based treatments in clinical settings. In: Steele, RG.; Elkin, TD.; Roberts, MC., editors. Handbook of evidence-based therapies for children and adolescents. New York: Springer; 2008. p. 447-469.
- Wampold BE, Brown GS. Estimating therapist variability: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology* 2005;73:914–923. [PubMed: 16287391]
- Weisz JR, Donenberg GR, Weiss B, Han SS. Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology* 1995;63:688–701. [PubMed: 7593861]
- Weisz JR, Doss AJ, Hawley KM. Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology* 2005;56:337–363.
- Weisz JR, Weiss B, Donenberg GR. The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist* 1992;47:1578–1585. [PubMed: 1476328]
- Wheeler S, Richards K. The impact of clinical supervision on counselors and therapists, their practice, and their clients. A systematic review of the literature. *Counseling and Psychotherapy Research* 2007;7:54–65.

Figure 1.
The effect of therapist adherence on youth Externalizing behavior problems one- year post-treatment

Figure 2.

The effect of supervisor adherence to Structure and Process and therapist adherence on youth Externalizing behavior problems one-year post-treatment.