Fidelity to Evidence-based Practice: Our Obligation to Effective Supervision and Service Delivery

Kimberly Gentry Sperber
Talbert House

CORRECTIONAL AGENCIES FACE increasing internal and external pressure to implement evidence-based practices to reduce recidivism. Internal pressures stem from an agency's desire to excel in the industry, to maximize positive individual and public safety outcomes, and to conduct business in an ethical and fiscally responsible manner. External pressures, on the other hand, have taken the form of outcomes-driven contracting, increased regulatory audits, mandatory participation in standardized program evaluations via the Correctional Program Assessment Inventory (CPAI) and the Evidence-Based Correctional Program Checklist (CPC), and attempts to mandate recidivism reduction through legislation. To meet these demands, correctional administrators often look to the empirical literature to tell them what services or practices work to reduce recidivism.

While necessary, knowledge of "what works" alone is not sufficient to sustain long-term change in an organization. Agencies also need effective implementation planning and execution skills and a comprehensive infrastructure designed to support and sustain evidence-based practices for the long term. Key to a robust implementation infrastructure are processes to ensure that the agency and staff are continually adhering to the organization's evidence-based practices, otherwise known as fidelity. This requires that agencies develop procedures to assess, monitor, improve, and maintain fidelity to evidence-based practices. Establishing and maintaining fidelity evaluation in real-world settings can be fraught with challenges, however. Agencies need a strong understanding of their evidence-based practices, the role of fidelity in producing organizational outcomes, the benefits associated with fidelity evaluation, and the resources required to effectively implement fidelity evaluation in real-world settings.

The Foundation of Evidence-Based Practice in Corrections

In a 1990 meta-analysis, Andrews and colleagues described and tested three principles of what they termed "appropriate correctional service" (Andrews, Zinger, Hoge, Bonta, Gendreau, & Cullen, 1990). Now well-established in the field of corrections, these principles are risk, need, and responsivity and are the foundation of the Risk-Need-Responsivity (RNR) Model. In short, these principles assert that: (a) correctional practitioners should identify individuals who have a higher probability of committing future crimes and reserve more intense services and supervision for those individuals; (b) correctional services should intentionally target for change those individual attributes that have been shown to be strongly correlated to criminal behavior (i.e., criminogenic needs); and (c) correctional programs should be based on cognitive-behavioral and social learning approaches while also attending to individual and organizational attributes that impact the individual's ability to respond to correctional interventions.

To test the principles of risk, need, and responsivity, Andrews et al. (1990) compared the performance of programs that adhered to these principles to the performance of programs that did not. Study results demonstrated that appropriate programs produced recidivism reductions of 30 percent while inappropriate programs *increased* recidivism by 6 percent. Since the publication of these results, a number of subsequent meta-analyses have replicated the finding that programs that adhere to the principles of risk, need, and responsivity produce greater recidivism reductions than those programs that do not adhere to these principles. This pattern of findings has been demonstrated with an array of justice-involved populations, including adults, juveniles, females, individuals convicted of sex offenses, and individuals convicted of violent offenses (e.g., Brusman-Lovins, Lowenkamp, Latessa & Smith, 2006; Dowden & Andrews, 2000; Dowden & Andrews, 1999; Lipsey, 1999; Hanson, Bourgon, Helmus, & Hodgson, 2009; Lovins, Lowenkamp, & Latessa, 2009). Consequently, the RNR Model is now a well-established empirical framework for working with justice-involved individuals.

In the 30 years following the Andrews et al. (1990) meta-analysis, researchers and practitioners have made significant progress in operationalizing the RNR principles into concrete tools and strategies for implementation in real world settings. Examples include, but are not limited to: (1) an array of empirically validated criminogenic risk and needs assessment instruments available to help agencies identify the risk level of the individuals they serve so that they can appropriately triage supervision and treatment according to an individual's risk to re-offend (see James, 2015; Hanson & Morton-Bourgon, 2009; Hoge, 2002); (2) empirically established guidelines about the

appropriate level of treatment dosage to provide to correctional clients based on their criminogenic risk (e.g., Lipsey, Landenberger, & Wilson, 2007; Makarios, Sperber, & Latessa, 2014); (3) empirical guidelines related to the appropriate density of criminogenic needs to target for change in high-risk correctional clients (e.g., Gendreau, French, & Taylor 2002; Lowenkamp, Pealer, Smith, & Latessa, 2006); (4) a variety of cognitive-behavioral curricula to treat correctional clients across a variety of correctional populations and services settings, including cognitive-behavioral models of probation and parole supervision (e.g., Bourgon, Bonta, Rugge, Scott, & Yessine, 2010; Gehring, Van Voorhis, & Bell, 2010; Lipsey, Landenberger, & Wilson, 2007; Lowenkamp, Hubbard, Makarios, & Latessa, 2009); and (5) identification of evidence-based practices within problem-solving courts (e.g., National Association of Drug Court Professionals, 2018a; National Association of Drug Court Professionals, 2018b).

Building an Infrastructure to Support and Sustain Evidence-Based Practices

During this same time frame, researchers across disciplines in human services, social work, education, addiction science, mental health, and corrections have also made significant progress identifying the organizational practices that are required to promote the systematic uptake and integration of evidence-based practices into daily operations that are separate from the practices used by staff with correctional clients. Consequently, there is now empirical evidence to provide guidance to agencies on an array of topics related to effective, sustainable implementation of evidence-based practices. In addition to fidelity evaluation, examples of such practices include staff recruitment, staff training, staff supervision and coaching, organizational change management, quality improvement processes, and data and decision support systems, to name a few (e.g., Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Schoenwald, Sheidow, & Chapman, 2009; Durlak & DuPre, 2008; Lipsey, 2009; Landenberger & Lipsey, 2005; Lowenkamp & Latessa, 2002). While all these interdependent practices contribute to an organization's successful implementation of evidence-based practices and positive recidivism outcomes, a review of these practices is outside the scope of this paper. Rather, the focus of this paper is solely on fidelity evaluation.

The Impact of Fidelity on Correctional Client Outcomes

It is important to note that there is evidence that both organizational-level adherence and individual staff-level adherence to the strategies specified by both the RNR model and implementation science affect client outcomes post-discharge from correctional programs. Research findings from early CPAI research illustrate the impact of organizational fidelity on these practices. The CPAI is an evidence-based tool developed to assess the extent to which correctional programs follow evidence-based practices. These practices include assessment and programming characteristics used with correctional clients as well as organization characteristics within the domains of program implementation, staff quality, and program evaluation. CPAI evaluators conduct site visits to gather evidence of adherence to an established set of criteria against which programs are scored. Programs achieving higher scores on the CPAI meet a greater number of the criteria than programs receiving lower scores.

The importance of the score is twofold: (1) it serves as a baseline or gauge against which programs can compare their performance in terms of evidence-based practice while also providing guidance on areas for improvement, and (2) research has shown that the scores produced from these assessments are correlated with post-discharge recidivism outcomes for clients served within these programs. For example, Lowenkamp, Latessa, and Smith (2006) examined data from 38 residential correctional programs for adults in order to examine the relationship between CPAI scores and program effectiveness. They found that CPAI scores were significantly correlated to reincarceration post-release from the programs. Programs achieving scores ranging from 0-49 percent demonstrated a 1.7 percent reduction in reincarceration compared to the comparison group programs, while programs achieving scores ranging from 50-59 percent demonstrated recidivism reductions of 8.1 percent. Programs achieving the highest scores, in the range of 60-69 percent, demonstrated the largest recidivism reductions of 22 percent.

While CPAI studies have assessed organizational-level fidelity to evidence-based practices, other studies have investigated the relationship between individual staff adherence to various evidence-based practices and post-program outcomes of justice-involved clients. Results of these types of studies reveal the same trend—individual staff adherence, or lack thereof, is associated with client outcomes even after clients have left the program. Some clear examples of this include studies of adolescent treatment, cognitive-behaviorally based probation, and core correctional practices.

Studies of adherence to evidence-based family treatments for juveniles have shown that staff adherence to the treatment model predicts post-treatment client outcomes. For example, Schoenwald, Sheidow, Letourneau, and Liao (2003) examined the impact of staff fidelity to Multisystemic Therapy (MST) in a study involving 666 youth and families served by 217 therapists in 39 sites. Therapists in these sites were rated on the Therapist Adherence Measure, a scale comprising 26 items. The youth were assessed immediately following treatment on the Child Behavior Checklist for both internalizing and externalizing behavior problems as well as the Vanderbilt Functioning Index, which assesses such factors as antisocial behavior, problems at home, problems at school, and problems with peers. Therapist adherence to MST predicted successful completion of treatment and reductions in problem behaviors of the youth immediately following treatment. Therapist adherence was later shown to predict decreased recidivism four years post-treatment (Schoenwald, Chapman, Sheidow, & Carter, 2010). Similarly, Sexton and Turner (2010) reported results from a study of Functional Family Therapy (FFT) involving more than 900 families and 38 FFT therapists. Study results showed that high-adherent therapists demonstrated a 35 percent reduction in felonies for treated youth, a 30 percent reduction in violent crimes, and a 21 percent reduction in misdemeanors at 12 months post-treatment relative to a comparison group of juvenile probationers. Of particular importance was the finding that the highest risk families had a greater probability of successful post-treatment outcomes when assigned to high-adherent therapists.

In correctional supervision research, Latessa, Smith, Schweitzer, & Labrecque (2013) found that high-risk probationers assigned to probation officers with strong fidelity to the Effective Practices in Community Supervision (EPICS) model had incarceration rates that were 12 percent lower than high-risk probationers assigned to probation officers with low fidelity to the model. Finally, Dowden and Andrews (2004) used a meta-analytic approach to demonstrate that staff use of effective use of authority, appropriate modeling

and reinforcement, problem-solving with correctional clients, effective use of community resources on behalf of correctional clients, and quality rapport and communication with correctional clients (collectively known as "core correctional practices") was associated with lower recidivism rates; this was especially true in programs that adhered to the RNR model, meaning that use of core correctional practices had the greatest impact on recidivism reductions in these programs.

While it is intuitive that programs and staff lacking in fidelity often produce inferior results compared to programs and staff with strong fidelity, studies that find increases in recidivism relative to no-treatment conditions are often surprising to practitioners and should be of particular interest given the public safety mission of corrections. Recall that as early as 1990, Andrews et al. demonstrated that programs that did not adhere to the RNR principles increased recidivism by six percent. Seeking to determine whether fidelity to RNR was just as important for supervision programs as it was for correctional treatment programs, Lowenkamp et al. (2006) found similar impacts on recidivism when fidelity was absent. They examined 66 community-based jail and prison diversion programs to determine the impact of organizational adherence to the risk and need principles. Programs in the study included intensive supervision probation, day reporting programs, substance abuse programs, electronic monitoring, and work release.

Results showed that programs that targeted higher risk offenders produced an average decrease in recidivism of five percent. Conversely, programs that did not primarily target higher risk offenders were associated with a two percent increase in recidivism on average. In addition, programs that varied the intensity of services by offender risk reduced crime on average by four percent, while programs that did not vary intensity by risk demonstrated no significant impact on recidivism. When examining adherence to the need principle, Lowenkamp et al. (2006) found that programs that provided more referrals for high-risk offenders compared to low-risk offenders reduced recidivism by seven percent, while those who did not meet this criterion only demonstrated a reduction in recidivism of 1 percent. When 75 percent of the referrals were treatment-oriented and targeted criminogenic needs, these programs reduced recidivism by 11 percent; however, when programs did not have this 3:1 ratio of service referrals targeting criminogenic needs, they increased recidivism by 3 percent on average. Finally, when examining the cumulative impact of targeting higher risk offenders, varying services by risk, providing more referrals for high-risk offenders, and ensuring that 75 percent of referrals targeted criminogenic needs, Lowenkamp et al. (2006) found that programs that did not use any of these strategies increased recidivism by 13 percent.

A 2010 statewide study in Ohio also found relationships between organizational fidelity and client outcomes, where lack of fidelity was associated with increases in recidivism (Latessa, Brusman-Lovins, & Smith, 2010). Researchers assessed 64 adult halfway houses and Community-Based Correctional Facilities in the state and included more than 20,000 correctional clients in the study sample. The study used a matched comparison group to compare the outcomes of correctional clients receiving treatment services to individuals with similar characteristics who did not receive these services. Using a subsample of treatment completers, the evaluation also assessed each program on select evidence-based practices.

One of the practices assessed was the quality of the cognitive behavioral groups offered in each program. Evaluators went into each program, observed the groups offered, and rated each group on several key characteristics, such as the use of role-plays in group and the amount of time spent in cognitive-behavioral groups. Based on these ratings, the evaluation staff then created a Cognitive Behavioral Group Scale. Each program was assigned one negative point if cognitive behavioral groups were offered but did not contain any of the positive attributes assessed, zero points if they did not offer cognitive behavioral groups, and one point if they offered cognitive behavioral groups that were offered 4 or more hours per week or allocated at least 50 percent of group time to role-playing activities. Study results showed that programs that received a score of negative one increased recidivism by 1.4 percent compared to their matched comparison group. On the other hand, programs that did not offer cognitive behavioral groups at all demonstrated a 4.8 percent reduction in recidivism relative to their matched comparison group, meaning that programs that did not offer cognitive-behavioral groups produced better recidivism outcomes than programs that implemented cognitive-behavioral groups that likely did not provide sufficient dosage and did not attend to skill-building activities that are a key ingredient within cognitive-behavioral interventions and have been shown to be important predictors of recidivism (e.g., Lowenkamp, 2004; Lowenkamp & Latessa, 2002; Sperber & Lowenkamp, 2017). Finally, programs that offered cognitive behavioral groups that met the fidelity criteria assessed in the study produced the best recidivism outcomes, with a 6.3 percent reduction in recidivism relative to their matched comparison group.

There is also research to suggest that individual staff non-adherence to evidence-based practices is associated with increases in recidivism. In one of the earlier studies to examine the association between individual staff competence in specific evidence-based models and post-treatment recidivism, Barnoski (2004) examined three groups of juvenile offenders who had participated in Functional Family Therapy in Washington State. These three groups of juveniles were those who had participated in Functional Family Therapy with therapists who had been deemed competent in the delivery of FFT, juveniles who had participated in Functional Family Therapy with therapists who had been deemed not competent or of borderline competence in the delivery of FFT, and a control group of juveniles who had not participated in FFT at all. Recidivism data were collected at 6 months, 12 months, and 18 months post-discharge from the program. At 18 months post-treatment, Barnoski (2004) found that the juveniles who had participated in FFT with therapists deemed not competent in the model had the worst recidivism outcomes, across three separate measures of recidivism. In other words, juveniles who had participated in an evidence-based intervention had higher recidivism rates than juveniles who had received no treatment at all. For example, 54 percent of juveniles assigned to non-competent therapists had committed either a new misdemeanor or new felony at 18 months, compared to 50 percent of untreated juveniles and 44 percent of juveniles treated by competent therapists. This same pattern also was observed for new felony offenses and new violent felony offenses at 18 months. At all three time points, juveniles treated by competent therapists demonstrated the lowest recidivism rates, followed by untreated juveniles, followed by juveniles treated by non-competent therapists.

Impact of Fidelity Evaluation on Staff

In addition to the impact on client outcomes, there is also evidence in the human services

literature that fidelity monitoring can have a positive impact on staff outcomes. For example, Aarons, Sommerfeld, Hecht, Silovsky, & Chaffin (2009) evaluated the impact of fidelity monitoring on children's services staff during a statewide implementation of an evidence-based model of care. Participating agencies were randomly assigned to one of four conditions: services as usual without fidelity monitoring, services as usual with fidelity monitoring, evidence-based practice model without fidelity monitoring, and evidence-based practice model with fidelity monitoring. Their analyses showed that evidence-based practice implementation paired with fidelity monitoring predicted greater staff retention relative to the other three study groups. Also of interest was that the highest turnover rate was found among staff assigned to the evidence-based practice model without fidelity monitoring condition. The study authors provided several possible explanations for this finding. The first was that implementation of evidence-based practices without fidelity support may be perceived by staff as just another change representing new demands. The second was that staff are less likely to develop a sense of mastery of a new practice when fidelity support is not provided to staff implementing new evidence-based practices. The third was that training staff in the evidence-based practice without fidelity staff to assist them with application of the model within the context of challenging cases may lead staff to view the new model simply as a mandate with no flexibility regarding how to best apply the model to a range of scenarios while still maintaining fidelity.

During the same study, Aarons, Fettes, Flores, & Sommerfeld (2009) examined the impact of assignment to the four study conditions on emotional exhaustion of staff. Aarons et al. (2009) defined emotional exhaustion as the "extent to which an employee feels that their emotional resources have been depleted (p. 2)." Results showed that the staff experiencing the highest levels of emotional exhaustion were in the service as usual with fidelity monitoring group. Their analyses also showed that there was no direct detrimental impact of fidelity monitoring on emotional exhaustion. In other words, fidelity monitoring did not demonstrate negative effects until it was paired with the service as usual condition. This led the authors to hypothesize that the increased oversight that accompanies fidelity monitoring in the absence of a clear rationale for that oversight may have eroded staff's sense of control and autonomy, thereby increasing emotional exhaustion. Organizational attention to emotional exhaustion among staff is important, as it has been associated with both staff turnover and staff adherence to evidence-based treatment models (e.g., Schoenwald et al., 2010).

Practical Considerations for Designing Fidelity Evaluation Processes

The empirical literature outlines clear benefits of monitoring and ensuring fidelity to evidence-based practices, in terms of both client and staff outcomes. However, there are a number of considerations agencies must take into account as they plan and develop a fidelity monitoring infrastructure. While not an exhaustive primer on operationalizing fidelity evaluation within correctional organizations, this section outlines issues and decisions agencies face when strategically planning fidelity evaluation initiatives.

Fidelity Measurement

Probably the first decision an organization must make is which elements of evidence-based practice to measure and monitor. This comes into play in two important ways. First, most organizations use multiple evidence-based practices that can cut across assessment practices, decisions related to triaging and brokering services, delivery of manualized curricula or models such as Cognitive-Behavioral Interventions, as well as non-manualized evidence-based approaches to working with clients, such as Motivational Interviewing. Few agencies have the resources to conduct comprehensive fidelity monitoring of numerous evidence-based practices simultaneously, thereby requiring agencies to prioritize which of its evidence-based practices to monitor at any point in time. Second, even within a single evidence-based practice, there may be multiple components that can be assessed. Fidelity evaluations of cognitive-behavioral models, for example, may focus on adherence to delivery of materials from a specific curriculum or may focus on specific techniques such as teaching a thought-behavior chain, teaching cognitive restructuring, facilitating role-plays, staff use of behavioral reinforcers, and so forth.

Once the components of an evidence-based practice have been chosen for assessment, an agency must have a method of measuring program and/or staff fidelity to the components. Because evidence-based practices vary in the extent to which they come with prepackaged measures of fidelity, agencies may find themselves needing to create measurement tools to conduct fidelity ratings. This requires having the necessary subject matter expertise to identify the "active ingredients" of the evidence-based practice (Herschell, 2010). Such ingredients are the core components of an intervention or practice that are responsible for producing the intended outcomes. Narrowing measures down to the active or core ingredients means that agencies can develop monitoring systems that focus on fewer elements of treatment and supervision, thereby simplifying the process.

Fidelity Methods

With fidelity measurement items identified, an agency can proceed to choose the methods to use to assess fidelity. These include both direct methods of assessment and indirect methods of assessment. Indirect methods of assessment include such things as staff self-report ratings of fidelity, client surveys of fidelity, and documentation reviews. Indirect methods are typically easier to implement and require fewer resources. This makes them an attractive starting point for agencies new to fidelity measurement and monitoring. Indirect methods can suffer from serious limitations, however, such as social desirability bias among staff completing self-assessments or lack of sufficient knowledge to recognize the occurrence or quality of program components by clients (Schoenwald, Garland, Chapman, Frazier, Sheidow, & Southam-Gerow, 2011). Direct methods of measurement that involve observation of staff use of evidence-based practices, on the other hand, are considered superior to indirect methods and should be the standard of measurement toward which agencies strive. While indirect methods may serve as a starting point and can be useful supplements to direct methods of measurement, they should not replace direct methods over the course of the long term (Herschell, 2010).

Direct methods involve fidelity raters directly observing staff while interacting with clients to rate their use of evidence-based practices in real-world settings using real clients rather than simulations. While these methods of measurement are preferred, agencies often face several challenges to implementing direct observation methods. The first set of barriers concerns challenges associated with having raters physically present during staff interactions. Examples include having raters sit in on individual or group treatment sessions,

individual treatment or case management sessions, or assessment appointments. The arguments against having live raters in the room for these activities are twofold: (1) they are perceived as disruptive to the clinical process and (2) having to be present at prescribed times creates time management inefficiencies with the process, especially within agencies that use supervisory or peer staff to perform fidelity observations. One viable solution to both challenges is to simply audiotape or videotape client sessions so that raters can view and assess the sessions remotely and at times most convenient to their schedules.

Fidelity Raters

Selection of the evidence-based practice to evaluate combined with the type of fidelity measurement to be used should guide the selection of fidelity raters within agencies. There are a number of important considerations here. The first is the amount of subject matter expertise required of fidelity raters, as these individuals should be trained experts in the practice that they are evaluating on behalf of the agency. The second is the role of the fidelity raters within the organization. For example, some agencies opt to use supervisory staff and to integrate the fidelity rating function into the supervision process, while other agencies use peer raters. For both supervisory raters and peer raters, integration of this added responsibility into their existing job roles can be a challenge. A third model is to have staff or contractors who serve exclusively as fidelity monitors and/or fidelity coaches for the organization. There are several advantages to this model. First, staff assigned to work exclusively within an agency's fidelity program would not have competing tasks that interfere with their ability to perform fidelity functions as would supervisory or peer staff. Second, they would not face the conflicts of interest that supervisors and peers may face. For examples, supervisors may face pressure for their staff to look good while peer raters may worry about backlash from fellow peers over ratings perceived as negative. Third, creating specialized fidelity positions means that there are fewer raters within the organization, which serves to reduce challenges associated with interrater reliability. Fewer raters who do not have competing tasks related to other organization responsibilities also makes it easier for the organization to hold these staff accountable for conducting the required fidelity procedures. The complexity of the practice to be evaluated, the sophistication of the fidelity measurement tools, expectations of fidelity raters to coach staff, and the level of experience of the fidelity staff with both the practices to be evaluated and the evaluation methodology all dictate the level and type of training that will be required for the staff chosen to implement fidelity procedures.

Logistical Considerations

Implementing a formal system of fidelity measurement also requires agencies to have systems and procedures to code, store, analyze, and report fidelity data. This includes dedicated responsible parties for completing these tasks as well as the appropriate data storage systems and software. Agencies that opt to use technology, such as audiotapes and videotapes, will also need processes for storing, securing, and destroying tapes in a way that protects the confidentiality of staff and clients. Perhaps even more important is the need to have procedures for responding to the results of fidelity measurement. This means that agencies must have an a priori understanding of the purpose and goals of their fidelity efforts, including plans for how the agency will use the data generated from the process. Examples of decisions that agencies face include whether to include fidelity ratings in formal staff performance evaluations and whether ratings will be used only for individual staff development or will be aggregated at the program level to assess programmatic trends and improvement over time. Responses to both staff deemed proficient and demonstrating high fidelity and staff deemed not proficient and lacking in fidelity should be determined and communicated to staff prior to launching any formal fidelity initiatives.

Agencies should also incorporate workload metrics into front-end planning to determine the volume and frequency of fidelity evaluations the agency can realistically manage. Examples of such metrics include the number of staff to be evaluated, the number of fidelity evaluators available to conduct ratings, whether fidelity evaluators are expected to provide feedback and coaching to staff whom they have rated, the hours required for evaluators to completed fidelity-related tasks as well as any other competing tasks, and the frequency at which the agency expects staff to be evaluated. This information can then be analyzed to create a fidelity rating schedule for staff (e.g., monthly, quarterly, semi-annual, annual).

The final consideration is whether staff will have advance notice of fidelity observations. This is a common practice in organizations currently assessing staff adherence to evidence-based practices. The primary limitation with this methodology, however, is that infrequent observations at pre-determined times simply provide evidence of a staff person's proficiency or competency in the model or practice. It does not provide evidence of fidelity. In other words, this methodology provides evidence of whether a staff person *can* perform the techniques observed but does not provide evidence of whether the staff person *does* routinely use the techniques in everyday interactions with clients. Alternative methodology can provide clearer evidence of fidelity. For example, agencies can require staff to videotape all sessions and then allow fidelity evaluators to randomly select tapes to review according to the agency's evaluation schedule (e.g., monthly). This would provide a more accurate sense of the staff person's routine use of the techniques while not requiring an increase in frequency or volume of ratings.

Conclusions

To sum, many correctional programs are allocating a great deal of resources to implementing evidence-based practices/models in an effort to improve client outcomes. However, few programs are actively and systematically monitoring staff and organizational fidelity to these models. The result is often poor fidelity to the model and corresponding poor client outcomes. Given the implications for public safety, correctional organizations have an ethical responsibility to assess and support staff fidelity to evidence-based practices to the extent afforded by their organizational infrastructure and resources.

References

Aarons, G. A., Fettes, D. L., Flores, L. E., & Sommerfeld, D. H. (2009). Evidence-based practice implementation and staff emotional exhaustion in children's services. *Behaviour Research and Therapy,* 47, 954-960.

Aarons, G. A., Sommerfeld, D. H., Hecht, D. B., Silovsky, J. F., & Chaffin, M. J. (2009). The impact of evidence-based practice implementation and fidelity monitoring on staff turnover: Evidence for a protective effect. *Journal of Consulting and Clinical Psychology, 77*, 270-280.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28,* 369-404.

Barnoski, R. (2004). *Outcome evaluation of*

Washington State's research-based programs for juvenile offenders. Olympia, WA: Washington State Institute for Public Policy.

Bourgon, G., Bonta, J., Rugge, T., Scott, T. L., & Yessine, A. K. (2010). The role of program design, implementation, and evaluation in evidence-based 'real world' community supervision. *Federal Probation, 74*, 2-15.

Brusman-Lovins, L., Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2007). Application of the risk principle to female offenders. *Journal of Contemporary Criminal Justice, 23*, 383-398.

Dowden, C., & Andrews, D. A. (1999). What works in young offender treatment: A meta-analysis. *Forum on Corrections Research,* 11, 21-24.

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology, 42,* 449-467.

Dowden, C., & Andrews, D. A. (2004). The importance of staff practice in delivering effective correctional treatment: A meta-analytic review of core correctional practices. *International Journal of Offender Therapy and Comparative Criminology, 48*, 203-214.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* National Implementation Research Network (FMHI Publication #231). https://nirn.fpg.unc.edu/sites/nirn.fpg.unc.edu/files/resources/NIRN-Monograph-Full-01-2005.pdf

Gehring, K., Van Voorhis, P., & Bell, V. R. (2010). "What works" for female probationers? An evaluation of the *Moving On* program. *Women, Girls, and Criminal Justice, 11*, 6-10.

Gendreau, P., French, S., & Taylor, A. (2002). *What works (what doesn't) revisited 2002: The principles of effective correctional treatment.* Unpublished manuscript, University of New Brunswick, St. John, New Brunswick, Canada.

Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36,* 865-891.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*, 1-21.

Herschell, A. D. (2010). Fidelity in the field: Developing infrastructure and fine-tuning measurement. *Clinical Psychology, 17*, 253-257.

Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior*, *29*, 380-396.

James, N. (2015). *Risk and needs assessment in the criminal justice system.* Washington, DC: Congressional Research Service. https://www.everycrsreport.com/files/20150724_R44087_0c47cc191ecc982888fa182c82ef0099a86eca8d.pdf

Landenberger, N., & Lipsey, M. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*, 451-476.

Latessa, E., Brusman-Lovins, L., and Smith, P. (2010) *Follow-up evaluation of Ohio's community based correctional facility and halfway house programs—Outcome study*. Center for Criminal Justice Research. University of Cincinnati.

Latessa, E. J., Smith, P., Schweitzer, M., & Labrecque, R. M. (2013). *Evaluation of the Effective Practices in Community Supervision Model (EPICS) in Ohio.* Final report submitted to the Office of Criminal Justice Services.

Lipsey, M. (1999). Can intervention rehabilitate serious delinquents? *The ANNALS of the American Academy of Political and Social Science, 564*, 142-166.

Lipsey, M. (2009). Primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims & Offenders, 2*(4), 124-147.

Lipsey, M. W., Landenberger, N. L., & Wilson, S. J. (2007). Effects of cognitive behavioral programs for offenders. *Campbell Systematic Reviews, 6*, 1-27.

Lovins, B., Lowenkamp, C. T., & Latessa, E. J. (2009). Applying the risk principle to sex offenders: Can treatment make some sex offenders worse? *The Prison Journal, 89*, 344-357.

Lowenkamp, C. T. (2004). *Correctional program integrity and treatment effectiveness: A multi-site, program-level analysis* (Order No. 3120848). Available from ProQuest Dissertations & Theses Global. (305201397). http://search.proquest.com.proxy.libraries.uc.edu/docview/305201397?accountid=2909.

Lowenkamp, C. T., Hubbard, D., Makarios, M., & Latessa, E. J. (2009). A quasi-experimental evaluation of Thinking for a Change. *Criminal Justice and Behavior, 36*, 137-146.

Lowenkamp, C. T., & Latessa, E. J. (2002). *Halfway house and CBCF findings: Program characteristics and outcome.* Center for Criminal Justice Research. University of Cincinnati.

Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy, 5*, 575-594.

Lowenkamp, C. T., Pealer, J., Smith, P., & Latessa, E. J. (2006). Adhering to the risk and need principles: Does it matter for supervision-based programs? *Federal Probation, 70*, 3-8.

Makarios, M. D., Sperber, K. G., & Latessa, E. J. (2014). Treatment dosage and the risk principle: A refinement and extension. *Journal of Offender Rehabilitation, 53*, 334-350.

National Association of Drug Court Professionals (2018a). *Adult drug court best practice standards: Volume I.* https://www.nadcp.org/wp-content/uploads/2018/12/Adult-Drug-Court-Best-Practice-Standards-Volume-I-Text-Revision-December-2018-1.pdf

National Association of Drug Court Professionals (2018b). *Adult drug court best practice standards: Volume II.* https://www.nadcp.org/wp-content/uploads/2018/12/Adult-Drug-Court-Best-Practice-Standards-Volume-2-Text-Revision-December-2018-1.pdf

Schoenwald, S. K., Chapman, J. E., Sheidow, A. J., & Carter, R. E. (2010). Long-term youth criminal outcomes in MST transport: The impact of therapist adherence and organizational climate and structure. *Journal of Clinical Child & Adolescent Psychology, 38*, 91-105.

Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research, 38*, 32-43.

Schoenwald, S. K., Sheidow, A. J., & Chapman, J. E. (2009). Clinical supervision in treatment transport: Effects on adherence and outcomes. *Journal of Consulting Clinical Psychology, 77*, 410-421.

Schoenwald, S. K., Sheidow, A. J., Letourneau, E. J., & Liao, J. G. (2003). Transportability of multisystemic therapy: Evidence for multilevel influences. *Mental Health Services Research, 5*, 223-239.

Sexton, T., & Turner, C. W. (2010). The effectiveness of Functional Family Therapy for youth with behavioral problems in a community practice setting. *Journal of Family Psychology, 24*, 339-348.

Sperber, K. G., & Lowenkamp, C. T. (2017). Dosage is more than just counting program hours: The importance of role-playing on treatment outcomes. *Journal of Offender Rehabilitation, 56*, 433-451.