Implementing *Incredible Years*: An Evaluation Perspective on an Evidence-Based Practice

Morrison Child and Family Services' implementation of the *Incredible Years* (*IY*) programs has been supported by the agency's internal evaluation department, which conducts ongoing evaluations for all of Morrison's programs. When Morrison began implementing *IY* in 2004, we evaluators expected our roles to shift, with more emphasis on fidelity monitoring and less emphasis on outcomes reporting. In fact, as we gain experience with *IY*, we are finding that is not necessarily the case.

We are fortunate to have had federal and local grants that provided us with enough resources to evaluate the program thoroughly, by measuring both fidelity and client outcomes. Having these resources gave us the flexibility to experiment with how to distribute our time between fidelity monitoring and measuring outcomes. It also supported our evaluators' efforts to be more involved in the day-to-day tracking and collecting of data.

What is Fidelity for *Incredible Years*?

Fidelity for *IY* is assessed in the following ways:

Group Leader Checklist. Group leaders review and complete a checklist customized to track the content and process of each unique weekly session. These checklists track the number of video vignettes shown during sessions as well as what "Did I's" were completed (i.e., "Did I review parents' goals?" or "Did I role play the play skills?"). For the current Morrison evaluation, these checklists first are submitted to the program's clinical supervisor for review and then to support supervision, and finally submitted to the evaluator for analysis. Currently we have Group Leader Checklist data analyzed for ten *IY* basic groups, two *IY* advanced groups and three *Dina Dinosaur* groups. (*Dina Dinosaur* is a problem-solving and so-

cial skills group curriculum for young children experiencing behavior problems.). Overall, program staff are in compliance with 92% of the required checklist items.

Cargiver evaluations. Caregiver participants complete evaluations of every group session. This provides an opportunity for group leaders to receive ongoing feedback from participants about their experiences with the group process. Weekly and final evaluations are collected and monitored by both the clinical supervisor and program evaluator. Since *IY* groups were started in the spring of 2004, 98% of participants have reported that they are satisfied or very satisfied with the *IY* programs at Morrison.

Group leader certification. Staff can pursue certification from the program's developer. To become certified, staff must submit at least one video-taped session per group series, peer and self evaluations, and all of

the fidelity paperwork listed above. Morrison currently has six staff certified in the *IY* basic parent series and two staff certified in the *Dina Dinosaur* child series. One staff member has become a certified mentor, which means she may offer authorized trainings to Morrison staff and provide mentoring for their groups.

What We Thought We'd Be Doing

As evaluators, we initially expected to monitor the collection of fidelity measures, but in fact, that role was taken on by the clinical supervisor. The clinical supervisor uses case-specific information from the fidelity measures to provide real-time course corrections to the therapists as they learn to implement a complicated package of protocols. Morrison is fortunate to have a program supervisor who understands how important the data collection and feedback loop are for the successful implementation of the model.

We also expected that we would be asked to calculate how well each group leader was performing on the fidelity measurement tools. Instead, the developer tracks the fidelity measure compliance percentage as part of each group leader's certification process. If there is a period of time when no group leader is working towards certification, this may require more fidelity compliance tracking by the evaluation team. In any case, the clinical supervisor will still be using fidelity forms as part of supervision and will continue to monitor fidelity so as to guard against drifting from the model.

What We Actually Did

This left the internal evaluation team time to evaluate outcomes, which is increasingly a requirement of funders. One of the advantages of implementing a program that had been thoroughly researched was that we could adopt a simple evaluation model using similar evaluation protocols. We knew what change to expect and what instruments to use to measure that change. For *IY* we specifi-

TABLE 1: PARENT/CAREGIVER ECBI RATINGS			
ECBI (N=147)	Pre-Intervention Mean	Post-Intervention Mean	Effect Size**
Intensity*	58.27	53.78	.61
Problem*	57.67	51.34	.64

*$p<.001$

** Effect size is a standardized measure of the magnitude of the difference between two samples, allowing comparison across different variables and sample sizes. Though interpretation of effect size requires consideration of several factors, general qualitative guidelines suggest that an effect size of .20 is small, .50 is medium, and .80 is large.

cally selected the Intensity Scale of the Eyberg Child Behavior Inventory (ECBI) to measure children's problematic behaviors and the ECBI Problem Scale to assess parents' ability to manage those behaviors, because the ECBI has been consistently used in evaluations and replication studies of the *IY* model. Data have been collected from group participants since the spring of 2004. Table 1 shows the mean pre- and post- scores for each scale as reported by parents and primary caregivers who completed an *IY* basic or advanced group series. Parents and caregivers reported significant improvements in both their parenting efficacy and their children's behavior.

Conclusion

The expected drastic changes to our roles as evaluators did not in fact occur. While we anticipated that we would have to take on the responsibility of monitoring fidelity, we learned that fidelity monitoring was an essential program component that was mostly carried out by the clinical supervisor and the developer.

For the agency as a whole, there are increased costs associated with collecting fidelity data and using the data for supervision and certification. This leaves fewer resources for examining outcomes. However, we were able to take advantage of the previous research on *IY* to design a simple and cost-effective pre/post evaluation study. It is difficult to know exactly how we will adjust our evaluation strategies when grant funds are no longer available, but more and more of our contracts are requiring both outcomes measurement and fidelity monitoring.

Authors

Michele Limpens is a program evaluator at Morrison Child and Family Services.

Linda Magnuson is Director of Program Evaluation for Morrison Child and Family Services.