Enmeshed or Engaged?
Program Evaluation
Strategies in Cross-System Programs

The National Conference on Child Abuse and Neglect (OCAN)
April 19, 2012

Disclaimers and Caveats
- We are program evaluators:
  - Using naturalistic or observational designs
  - RPG Grants - 53 grant sites in nation
- We are engaged, but sometimes worry about being enmeshed scientific advisors in a partnership model with practitioners.
- Our job: Build evidence on strategies and/or programs that work to improve child welfare outcomes.

Our Premise
- The demands of the current Children’s Bureau service grants exceed traditional role definitions between program and evaluation professionals in:
  - complexity of service delivery;
  - complexity of selection and use of evidence-based practices; and
  - how to represent what services actually do.
- The demand for evidence to guide practice in Child Welfare is greater than ever.

Ruth A. Huebner, Ph.D.,
Child Welfare Researcher
Kentucky DCBS
rutha.huebner@ky.gov

Sharon M. Boles, Ph.D.,
Evaluation Director
Children and Family Futures
sboles@cffutures.org
Objectives

- To explore the challenges of program evaluation and make suggestions for improving the rigor of program evaluation design and implementation.
- To explore emerging roles for program evaluators in projects with inter-agency collaborations.
- To demonstrate how practitioners can take a more active role in owning the evaluative processes and outcomes.
- To explore implementation research and adherence to fidelity.

The Challenges of Program Evaluation

Research Designs Challenges

- **Randomized Control Trial**
  - Discrete research question or problem:
    - Often short term
    - Readily measured and simple, often singular
    - Focus on discrete disorders
    - Controlled conditions
  - Discrete Intervention
    - Medication
    - Comparison of intervention vs. Treatment as usual in single agency

- **Cross Agency Studies**
  - Complex Problems
    - Substance abuse
    - Multiple co-morbidities
    - Families
    - Communities
    - Multiple agencies/systems
  - Complex Intervention
    - Multiple strategies
    - Many staff and turn over
    - Many individual exceptions
  - Complex outcomes

Program Evaluation and Practitioner

- The data for program evaluation often comes from practitioners who may be unsophisticated about scientific data collection methods, and anxious or uninterested in evaluation results.
- Approaches like random assignment are generally viewed with resistance or fear of unethical behavior.
- Fidelity to a practice or model may diminish or time with a return to routine clinical habits.
The Research Challenge: Can you Control the Environment?

Courts
Different Practitioners in Each Agency
Different Funding Sources and Jurisdiction
Different Restrictions and Statutes
Different Data Systems

Child Welfare
Behavioral Health

The Same Families in all Three Systems
Different Clients from the Same family

Interventions Focusing on Families are More Complex

Community
Parents
Parental Capacity
Case or Family
Substance abuse
Co-occurring disorders
Child Welfare Outcome

Children
Family Functioning
Child Characteristics

Challenges of Random Assignment
- Seen as the 'gold standard' of program evaluation but is increasingly difficult when dealing with people in naturalistic settings. Discussion by Guo & Fraser (2010)
- Randomization may not exclude hidden selection bias and is influenced by processes such as the Hawthorne effect and potential biases in assignment to conditions.
- Average treatment effects mask individual results that may stem from multiple and important factors.

Background Literature
Positive Outcomes Are Only Part of the Story

- Even studies of drug effects when replicated over time tend to show diminished results. This is surprising, given that drugs are supposedly highly consistent. Other results also diminish.
- Negative results are not usually published.
- Studies with animals and humans are highly susceptible to many sorts of perception and selection biases.

Programs Change Over Time

- Klein (July 18, 2011). Head Start Doesn’t Work. Time Magazine
- Proven in the 1960’s to improve school performance.
- Taken to scale
- Head Start Study by DHHS found that the effects were minimal and vanished by 1st grade.
- Regression toward the mean or mediocrity.

Implementation Research (Fixsen et al., 2005)

- The science related to identifying evidence-based practices has improved, however... implementing programs with fidelity and improving results in human services has lagged far behind.
- Carefully implementing strategies and programs within a multi-year process with intensive attention may improve outcomes.
- But what strategies? And how do these differ from treatment as usual in child welfare?.

Put it in perspective

- Did you know that there have never been ANY RTCs of parachute use as a prevention of death when falling from heights? Only anecdotal evidence. Smith & Pell (2008), British Medical Journal
- Is it any wonder that if we provide coordinated and targeted services to child welfare families that they improve?
- Has RTC been subjected to RTC? Yes, see references.
Criteria for Evidence-Based Practice is at least One RTC

- The design that often focuses on a discrete disorder excludes subjects with co-morbidities. Most families and adults in child welfare have multiple co-morbidities, so EBP may have limited application to typical or 'real' clients.
- EBP generally rely on well trained providers whose skill level may exceed community practice.
- If the program is adapted to meet a new context or state or county regulations, the evidence of effectiveness may be negated.
- So, how do we advance within these challenges?

A Variety of Methods (GAO Report, 2009)

- Program Evaluation A Variety of Methods Can Help Identify Effective Interventions
- Requiring randomization as the sole proof of intervention effectiveness will exclude many potentially effective and worthwhile practices.
- Components of effective program evaluation:
  - Measure the impact of the intervention (consider subgroups, treatment in comparison group)
  - Rule out competing causal explanations (Ensure the equivalence comparison groups, measure fidelity, entertain alternative ideas).

Potential Solutions

- Improve rigor of naturalistic evaluations
- Change the role of program evaluators from expert to scientific advisor or learning partner
- Focus on implementation with measuring fidelity to well-defined strategies
- Support effective strategies and programs to improve child welfare practice
Improve the Rigor of Applied or Naturalistic Evaluations

Comparison Groups
Logic Models
Cost Benefit Analysis

Crafting the Story for Child welfare – Legislators – Funding Sources

Questions
- How many families (adults/children) are getting needed services?
- Does it Work Better than ________?
- How does it work?
- Is it worth it?
- How much would it cost to expand it to more families?

Data Needed
- Numbers served, for how long, by whom.
- Comparison to ________ with results important to the audience.
- Logic model and measures of fidelity to practices.
- Costs and value of results.
- How many other families meet selection criteria for program?

Does it Work Better Than___?
- Always strive for at least a study of Impact: Some comparison of folks who receive and do not receive the program.

Some Potential Designs
- Pre-post test (weaker design – Do they improve, but not compared to what)
- stagger start intervention

Potential Designs for Comparison Groups
- Matched control cases
- Control group based on propensity score matching
- compare to naturally occurring group
- Similar counties or teams
- Referred but not served due to random events such as case loads full
- Aggregate data from other source of typical outcomes or national norms
- Compare to results achieved on other studies.
### Terminology of Program Evaluation

- **Formative or Benchmark Based:** How is the program evolving and meeting its targets, achieving short-term outcomes?
- **Process:** How does the program work? What strategies, activities, requirements, hurdles, interaction are occurring or have occurred.
- **Summative or Outcomes:** What changes are occurring, what benefits are people getting: Short and long-term.

### A Logic Model

- Statement of Mechanism of Change
  - Proximal changes or outcomes
  - Distal changes or outcomes
- What is the change agent? How does it work? How long should it be applied for? Does it apply to everyone in the program?
- Hypotheses to test
- Why does it work or not work?

### What is a Program? A Logic Model

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Process</th>
<th>Outputs</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td>Who is served?</td>
<td># of clients served</td>
<td>Changes that occur.</td>
</tr>
<tr>
<td>Personnel</td>
<td>For how long?</td>
<td># of items produced</td>
<td>Short term: learning</td>
</tr>
<tr>
<td>Costs</td>
<td>Using what techniques?</td>
<td></td>
<td>Long-term: Improved quality of life</td>
</tr>
<tr>
<td>Clients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Child Welfare and Substance Abuse Program Serves the whole family include father and mother – the unit for the child is the family. To keep children safe, we need to address the family.

Medium Outcome: Child has less repeat maltreatment – improved child safety. Both parents contribute to child safety and treatment reduces family stress.

Proximal Outcome: Adult achieves sobriety and family achieves sobriety. Entire family working toward common goal. Treatment for family dynamics such as enabler and other problems within the family. Education for all. Mothers need partner supports so they both need treatment.

Distal Outcome: More support for child and less likely for child to abuse substances.

In Public Child Welfare

Why is it Worth It?

Cost Avoidance

<table>
<thead>
<tr>
<th>Family Preservation Cost Savings Summary</th>
<th>Cost Avoidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 children (of the 701) avoid short stays of 60 days in OOH</td>
<td>2,330,700</td>
</tr>
<tr>
<td>201 children (of the 701) avoid 15 months (457 days) of OOH</td>
<td>7,136,370</td>
</tr>
<tr>
<td>995 children with 3.0 months (91 days) shorter stays in OOH</td>
<td>7,034,441</td>
</tr>
<tr>
<td>Costs of staff, stipends, and supports to foster parents</td>
<td>&gt;1,000,000</td>
</tr>
<tr>
<td>Total Cost Avoidance: At least</td>
<td>17,501,511</td>
</tr>
</tbody>
</table>
Family Preservation (FPP) Services Cost Savings Summary

- Total Costs for FPP from July 1, 2006, to June 30, 2007 = $6,139,414.80
- Total Cost Avoidance = $17,501,511
- Average cost of serving one family = $4,584.20
- Average cost of serving one child in OOHC for nine months = $21,282
- Ratio of Costs to Savings = for every $1 for FPP, savings of OOHC costs = $2.85.

Reduction of Trauma: Priceless!

Additional Need Among Neglected or Abused Infants (to 1 year) with Risks to Safety Due to Parental Substance Abuse

The Role of the Evaluator: Learning and Improving Together

Using Empowerment Evaluation Strategies

Realities of Program Evaluation

- It must be useful, relevant, practical, and scientifically defensible.
- It shapes the program and it never ends.
- It can be done by providers with experience in research, but often needs guidance, design help, and analysis.
- There is a growing demand for evaluation.
- Practitioners must be engaged in asking: “How will we know if it is working?”
Emerging Roles

<table>
<thead>
<tr>
<th>Traditional</th>
<th>Scientific Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>External</td>
<td>Internal</td>
</tr>
<tr>
<td>Expert</td>
<td>Coach</td>
</tr>
<tr>
<td>Practitioners are Dependent</td>
<td>Practitioners are Partners</td>
</tr>
<tr>
<td>Independent Judgment</td>
<td>Collaboration</td>
</tr>
<tr>
<td>Report to Prove</td>
<td>Reports to Improve</td>
</tr>
</tbody>
</table>

Scientific Partner Roles

- What is meant by ‘scientific partner’?
  - Person or persons that are a bridge between practice and science - both the science of selecting/implementing practices and the science of examining outcomes.

- Scientific Partner Roles:
  - Literature review and guiding practitioners on the information available;
  - Suggesting instruments to measure key clinical or behavioral characteristics for both clinical and research needs;
  - Suggesting ways to translate program strategies into measurable indicators. Or at least describe service characteristics and help standardize these.

Scientific Partner Role

- Check for adherence or fidelity to program procedures and serve as a catalyst for growth and increased potency for intervention. Make adjustments to program; test the reality of program assertions.

- Improve program accountability for achieving results rather than completing process or producing outputs.

- Examine the impact of strategies on results to infuse this idea: “Does what we do make a difference in what we achieve?” Are there strategies that really don’t make a difference and could be modified or eliminated.

- Guide decisions on continuing programs and showing the value - costs/benefits.

Engaging Practitioners In Program Evaluation

Overcoming Terror Among Practitioners and Learning to Learn Together
Why Engage Practitioners

Practitioners often own the data, the data entry and the results.

“If you want providers onboard when you land, make sure they are onboard when you take off.”

Engage them in design, implementation, interpretation, adaptation etc of the program evaluation.

Empowerment Evaluation

- Align with program goals: Design the measure to match the program goals from the very beginning. The program leads, but evaluation can help articulate and improve the program.
- Involve the consumers of the evaluation in the design.
- Ask the recipients of services about their perceptions of outcomes and ways to improve.
- Plan the evaluation report to include outcomes important to the practitioners.
- Express results in pictures and graphs.
- Ask what does this mean for people (providers, customers, funding agency).

Program Evaluation and Practice: Goals and Fears

- **Articulate and then monitor clear outcome expectations** (“but we will know if it is working and it is way to hard to show it!”)

- **Improve program delivery and outcomes** (“but the program is based on solid and documented principles; we know what we are doing, and this will be so boring and take so long and distracted us from our work!”).

Program Evaluation and Practice: Goals and Fears

- **Show effects of an innovative program** (“but what if we find out it isn’t working - what then?”)

- **Demonstrate benefits in terms of a contract**: funding agency, public, customer, or government (but the results are statistically significant at the .05 level!) (Research myth)
Guidelines for Engaging Practitioners

Evaluation can be a force for change and improvement.
- Program first rather than evaluation first.

Evaluation is a process not a report.
- Explore options and ideas, not jump to conclusions.
- Work together to understand program processes and outcomes.

Guidelines for Engaging Practitioners

Practitioners can tell the story too.
- Present findings in graphs, diagrams that they understand.

Overcoming terror and learning to learn together.
- Present snippets of data at every meeting and generate reports regularly.
- Ask “What does this mean for families”
- Translate results in the numbers of children or families.

Practitioner Empowerment

- Practitioners feel ownership and the excitement that ownership elicits. They can tell the story too.
- Data tells them how effective their programs are, what works or does not work, for which groups, and why.
- Data tells them how efficient they are in their service delivery and the level of quality of their service.
- They no longer walk around wearing blinders, hoping “everything is going to be OK”
- They know---and knowledge is power.

Promote Fidelity to Strategies that Work

An Essential Role of Evaluation
Not compliance monitoring to rules – fidelity to strategies.
Fixing it while you drive.
Promoting Fidelity

- Implementing a new program or strategy takes 2-4 years for full implementation and before distal outcomes can be accurately measured. Measuring strategies is as important as measuring outcomes.

- Fidelity – the program functions match the description.
  - Implementation Fidelity - Match between the program description and the quality of service delivery.
  - Organizational Fidelity - Match between the description and the overall operations – e.g., staff selection and training, funding.
  - Drift - lack of fidelity that may pull toward the ‘old way’ of doing business. Especially after the ‘honeymoon period’.

### Example Program Strategy: Quick Access to the First Treatment

**Impact of Access in 4 days for Families (N=254)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Access in 4 days</th>
<th>Access 5 days or more</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child in State Custody</td>
<td>10.9</td>
<td>16.2</td>
</tr>
<tr>
<td>Any Member of the family</td>
<td>63</td>
<td>51.4</td>
</tr>
<tr>
<td>success in tx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reunified at most recent</td>
<td>21.7</td>
<td>14.7</td>
</tr>
</tbody>
</table>

**Fathers Access to Intake in 4 Days and Outcomes (N = 198 fathers in closed cases)**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Access in &lt;=4days</th>
<th>Access &gt;=5days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attend any support group</td>
<td>72.2</td>
<td>59.3</td>
</tr>
<tr>
<td>Attend support group at closure</td>
<td>19.4</td>
<td>15.4</td>
</tr>
<tr>
<td>Making progress at closure</td>
<td>47.2</td>
<td>34.6</td>
</tr>
</tbody>
</table>

90% Receive Intake within 8 Days of Referral
Facing Bad Adults that Flee
(closed case opened >=6months; N=251)

- Fathers
- Mothers

<table>
<thead>
<tr>
<th>Fled at any time</th>
<th>Fled first rating</th>
<th>Fled last rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>5</td>
<td>16.3</td>
</tr>
<tr>
<td>25.7</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>

10 of 89 adults that fled returned to achieve sobriety

Fled = at least one monthly rating of "Through this entire month, unable to locate this adult at any time"

The “Elephant” in the room

How Can Program Evaluation Improve Child Welfare and Substance Abuse Practices and Outcomes?

Thank you and Contact Information
Ruth A. Huebner, PhD
rutha.huebner@ky.gov

Sharon Boles, PhD
sholes@cffutures.org

See Reference List for additional resources