

Assessing the Degree of Implementation when Evaluating a State-wide Initiative

Patricia Noonan, Ph.D., University of Kansas (Chair)
Amy Gaumer Erickson, Ph.D., University of Kansas
Julie Morrison, Ph.D., University of Cincinnati
Patricia Mueller, Ed.D., Evergreen Educational Consulting, LLC
Zach McCall, Ph.D. Candidate, University of Kansas

Purpose of the Panel Session

- ▶ Explore current trends in assessing implementation of large-scale statewide initiatives.
- ▶ Discuss advances in evaluation methods and practices from a primary reliance on compliance measures to a concern for:
 - Depth of implementation
 - The role of implementation drivers
 - Web-based technology for real-time data collection

Implementation Evaluation: Historical Context and Contemporary Thinking

Julie Morrison, Ph.D.
Julie.Morrison@uc.edu

Defining Implementation

- ▶ “Implementation refers to all of the activities focused on the actual operation of a program once it moves from the drawing board and into action” (Love, 2004, pp. 63–64).

Defining Implementation

The National Implementation Research Network

When thinking about implementation,
we must be aware of ...

- ▶ Two sets of activities
 - Intervention-level activity
 - Implementation-level activity
 - ▶ Two sets of outcomes
 - Intervention outcomes
 - Implementation outcomes
- (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005)

What is Intervention Fidelity?

- ▶ Intervention fidelity is the degree to which an intervention or treatment is implemented as planned.
- ▶ Other terms used include:
 - Treatment integrity
 - Intervention adherence
 - Procedural adherence
- ▶ Traditional measures of intervention fidelity:
 - Procedural Checklists
 - Direct Observation
 - Practitioner self-report
 - Permanent Product Review

What is Implementation Fidelity?

- ▶ Implementation fidelity is the degree to which a specified set of activities designed to put into practice an activity or program of known dimensions is completed as intended.
- ▶ Other terms used include:
 - Formative evaluation
 - Process evaluation
 - Implementation analysis
 - Performance monitoring
- ▶ Traditional measures of implementation fidelity:
 - Compliance in implementing program components

Why Measure Intervention Fidelity?

- ▶ To make valid conclusions about intervention effectiveness, one must assess client outcomes and intervention fidelity (Shadish, Cook, & Campbell, 2002).
- ▶ In the absence of such data, it is impossible to determine whether poor consumer outcomes result from an ineffective intervention or an effective intervention that is poorly implemented.
- ▶ Internal Validity & External Validity

Why Measure Intervention Fidelity?

- ▶ When interventions are implemented in applied settings, they are often altered to meet the needs of the particular context. To the degree that interventions are modified in unknown ways, it makes it difficult to accurately evaluate the utility of the originally designed intervention (Lane, Bocian, MacMillan, & Gresham, 2004).

Why Measure Intervention Fidelity?

- ▶ Although it is standard practice to report with great specificity the reliability and validity of the outcome measure (dependent variable), traditionally this same level of scrutiny has not been required for the intervention (independent variable) (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000).
- ▶ Construct Validity

Why Measure Implementation Fidelity?

- ▶ Internal Validity, External Validity, Construct Validity
- ▶ In practical terms, implementation evaluation enables evaluators to be clearer about what worked and what did not to produce the intended program outcomes (Bickman & Heflinger, 1995).

Breaking Free of the Black Box

- ▶ Traditional evaluations have used a “black box” paradigm, in which evaluators assume that the program “technology” (that is, theory as applied in program activities) is controlled by the program staff and observed outcomes are caused by the program (Love, 2004).



- ▶ Black box evaluations reveal little about the process of program delivery or how to improve programs.
- ▶ They do not adequately describe the relationship of program activities, program context, and outcomes.
- ▶ Attention to treatment is limited to:
 - A description of client characteristics
 - Service “dosage” (e.g., number of contacts, duration of participation in program activities)

Embracing the Transparent Box

- ▶ Encourages the careful study of program delivery, including an assessment of how well programs are implemented and the relationship among theory, program activities, and program outputs and outcomes.
- ▶ Explicitly considers organizational and environmental factors outside the program itself and how they influence intake into the program, the acceptance and value of program activities, and the achievement of program outcomes.

(Love, 2004)

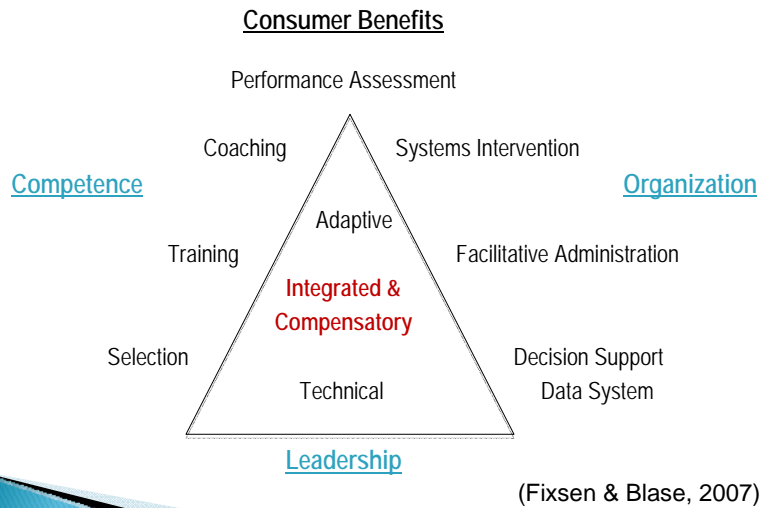
A Call to Evaluators

- ▶ The evaluation of implementation has become particularly important during the last decade.
- ▶ Spurred by major management reforms and demand by the public, organizations in all sectors (private, public, and nonprofit)
- ▶ Reformed management approaches are highly data driven.
(Love, 2004)

What Factors Influence Intervention Fidelity?

- ▶ Theoretical strength of an intervention
- ▶ Intervention complexity
- ▶ Competence of the Interventionist
- ▶ Time required
- ▶ Resources
- ▶ Number of interventionists
- ▶ Interventionists motivation
- ▶ Treatment effectiveness

What Factors Influence Implementation Fidelity?



Missouri Integrated Model: Using Evaluation Data in an Iterative Process

Amy Gaumer Erickson, Ph.D., aerickson@ku.edu

Patricia Noonan, Ph.D., pnoonan@ku.edu

Zach McCall, Ph.D. Candidate, zmccall@ku.edu

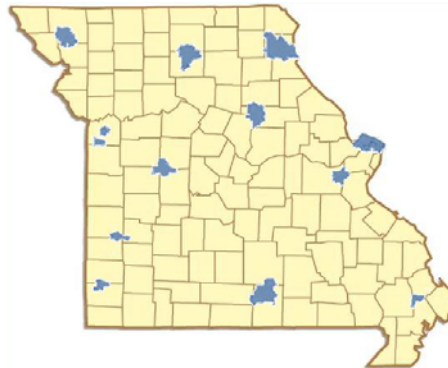
Stages of Implementation

1. Exploration and Adoption
2. Program Installation
3. Initial Implementation
4. Full Operation
5. Innovation
6. Sustainability

Implementation is a process that takes time.
It often progresses unevenly.

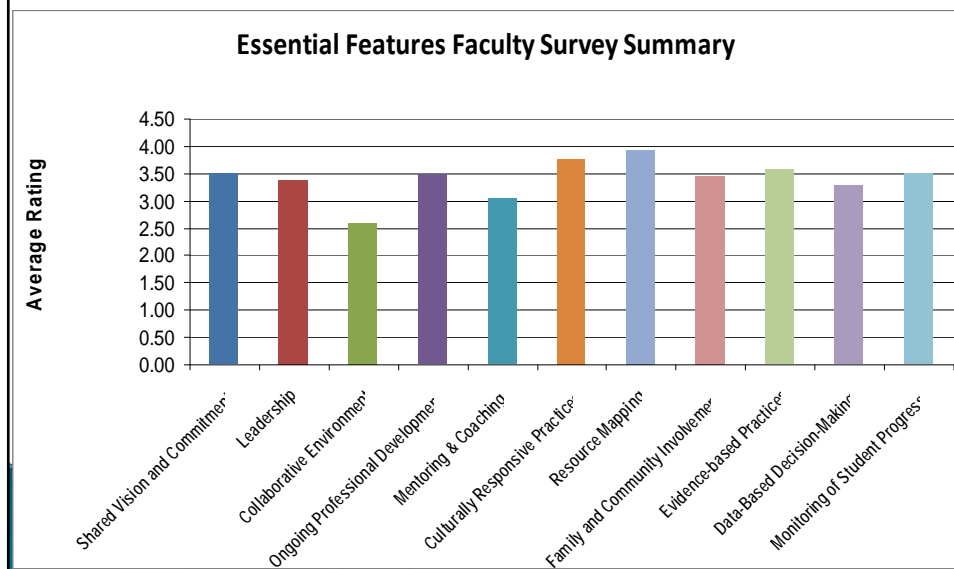
Missouri Integrated Model (MIM)

- ▶ 11 Essential Features of Systems Change
- ▶ Driven by Building and District-level Teams
- ▶ Professional Development with Coaching
- ▶ 14 Pilot Districts
- ▶ 5 Implementation Facilitators
- ▶ Management Team
- ▶ Implementation Team
- ▶ Advisory Group



www.MIMschools.org/cop

Essential Features Faculty Survey Summary Report



Self-Study Summary

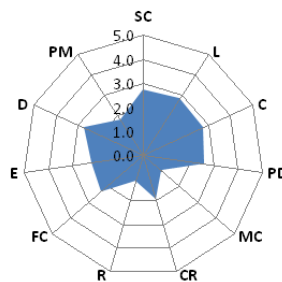


SECTION II: Summarizing the Self-Assessment

Once the self-assessment is complete, the findings should be summarized such to facilitate action on planning. This section provides a method of organizing the self-assessment findings by priority and needs for professional development, resources, and additional data.

First, calculate the average score indicating your determined current status of each MIM Essential Feature and record this value in the chart below. Second, plot the values on the spider web graph, connect the dots and shade in the area. The extent to which the web is entirely shaded indicates the degree to which the essential features are implemented fully.

	Rubric
Shared Vision & Commitment	2.7
Leadership	2.8
Collaboration	2.8
Professional Development	2.6
Mentoring & Coaching	1.0
Culturally Responsive	1.9
Resources	1.1
Family & Community Involvement	2.3
Evidence-based Practices	2.2
Data-based Decision-making	2.7
Progress Monitoring	1.7



Results of Implementation Facilitator Interview

- ▶ Long and short reports summarizing information: *Successes Observed, Observed Challenges and Responses, Anticipated Challenges, Essential Features, Supporting the IFs*
- ▶ Informs Management Team Discussion and Planning
- ▶ Included in Annual Report to Stakeholders and Pilot Schools

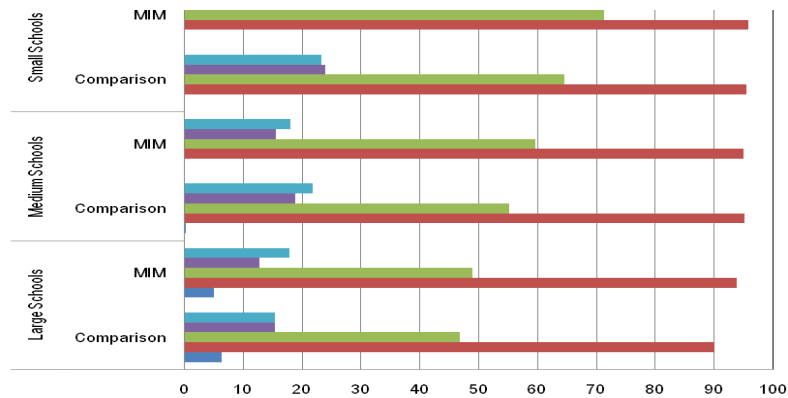
How the MIM supports IFs...

Collaboration Survey (RPDC and State) Online survey, score 1–5

- ▶ Knowledge
- ▶ Support
- ▶ Participation
- ▶ Collaboration

Results Compared and Discussed

Comparison School Data



	Large Schools		Medium Schools		Small Schools	
	Comparison	MIM	Comparison	MIM	Comparison	MIM
Percent of students with disabilities proficient on MAP* in Math	15.41	17.97	21.78	18.05	23.26	17.87
Percent of students with disabilities proficient on MAP* in Reading	15.40	12.78	18.94	15.53	23.97	16.73
Percent of students with disabilities in the regular class over 80% of the time	46.90	49.00	55.25	59.75	64.67	71.33
Average Attendance Rate	90.04	93.89	95.26	95.08	95.48	95.80
Average discipline incidence rate	6.43	5.09	0.28	0.00	0.21	0.00

Using Evaluation Data to Inform Practice in Three States

Patricia Mueller, Ed.D.
eec@gmavt.net

Selection, Training, & Coaching

- ▶ State coach team members conduct interviews for regional coaches.
- ▶ Regional coach job description includes prescriptive set of knowledge, skills & abilities.
- ▶ 6 day-training schedule in place for new coaches.
- ▶ Recognition that replicable training modules need to be developed in order for scaling-up to occur.

Performance Assessments

- ▶ Use of on-line PD/TA Log, includes some management components (insert SM log here)
- ▶ Baseline student-level data collected and trend tracking (e.g., achievement, discipline referrals).
- ▶ Annual Participating Personnel Survey to assess satisfaction with helpfulness/quantity of coaching received.
- ▶ TBD: How will the effectiveness of the coaches be assessed? How will the effectiveness of the coaching processes be assessed? Fidelity measures? (Fixen, 5/09)

Using Data to Support Decisions

- ▶ Quarterly formative evaluation data collected; annual process & outcome data reported.
- ▶ Assessment of organizational functioning (i.e., Rtl Blueprint District Self-Assessment). Baseline data collected. Discussing frequency of administration.
- ▶ Increased emphasis on buy-in at all levels of the organization. Goal is to use data to assess readiness & commitment.
- ▶ TBD: How to tie quarterly & annual reports to *organizational* processes & outcomes? (Fixen, 5/09)

Facilitative Administration Supports

- ▶ Learning curve from Year 1 Cohort to Year 2 Cohort. Lessons learned...without full administrative support, initiatives will fail.
- ▶ TBD: Need to determine types of supports administrators need to ensure success of the innovation and develop assessment tools.
- ▶ How will the effectiveness of the administrative processes be assessed? (Fixen, 5/09)

Systems Interventions

- ▶ All projects identify partnerships with IHE's and family support organizations.
- ▶ Involvement of the partners is limited. It's a challenge to bring partners to the table when much of the effort focuses on field work, but without them, the innovations won't be sustainable.
- ▶ TBD: Assessment of the partnerships with external systems to ensure success of the initiatives. What strategies are in place or will need to be created for the innovation to work with external partners? (Fixen, 5/09)

Practical Guidelines for Advancing the Assessment of Implementation

Julie Morrison, Ph.D.
Patricia Mueller, Ed.D.
Patricia Noonan, Ph.D.
Amy Gaumer Erickson, Ph.D.
Zach McCall, Ph.D.

Advancing Traditional Approaches

- ▶ Component Analysis
 - Involves a careful description of the program operations at each phase (component) of service delivery
 - The program, as it was intended, is compared to the evaluation data regarding the actual program operations.

(Love, 2004)

Contemporary Applications of Component Analysis for State-wide Initiatives

- ▶ Emphasis on collaborative approaches
- ▶ Need for direct and frequent communication (e.g., e-mail, video/phone conferencing)
- ▶ Recognition that the program itself is often evolving in real time
- ▶ Need for real-time data collection with quick feedback loop (e.g., web-based data entry portals for self-assessments and satisfaction surveys)

Balancing Replication and Mutual Adaption

- ▶ **Replication:** Implementing the program components with fidelity.
- ▶ **Mutual Adaption:** A process whereby clients are encouraged to adapt the program components to the needs of their local context (Datnow, Hubbard, & Mehan, 2002; Hubbard & Mehan, 1999; Stringfield & Datnow, 1998)

Balancing Replication and Mutual Adaption: Recommendation

- ▶ Conduct Component Analysis in light of ...
 - The **non-negotiables**: The defining components of the program
 - The **negotiables**: Aspects that can be contextualized to better fit the local application
 - The **deal breakers**: Competing factors that detract from the program's effectiveness if present

(Hawkins, Morrison, Musti-Rao, & Hawkins, 2008)

Measuring Depth of Implementation

- ▶ It is possible to come to very different conclusions about the degree of implementation depending upon whether an evaluator focuses on activity structures and materials or changes in underlying norms, beliefs, and principles (Coburn, 2003)
- ▶ Assessing depth of implementation may require in-depth interviewing, direct observation, and the review of permanent products and practitioner logs.