Systematizing the Measurement and Reporting of Intervention Delivery in Education Research

Catherine Darrow
Barbara Goodson
Beth Boulay

Society for Research on Educational Effectiveness
March 2014
Overview of the Measurement System

Counterfactual condition: Business as usual

Treatment condition: Program/intervention model

Specified in Logic Model
- Planned program inputs & activities
- Planned mediators:
  - Teacher behavior / knowledge
  - Experiences of students

Delivered program inputs & activities
- Actual mediators:
  - Teacher behavior / knowledge
  - Experiences of students

Common teacher / student experiences across conditions
- Teacher behavior / knowledge
- Experiences of students

Treatment/Control Contrast

System of measurement for fidelity of implementation - assesses the degree to which inputs & activities [key components] of an intervention are delivered.
Recent reviews of fidelity measurement have claimed that diverse approaches in defining and measuring fidelity abound.

However, there is general agreement in the field that it is necessary to:

- Clearly define the intervention model (Century, Rudnick, & Freeman, 2010; Nelson, et al., 2012)
- Identify the critical elements that are theoretically tied to impacts of interest (Mowbray, et al., 2003; O’Donnell, 2008)
- Develop original or use pre-existing measures to assess levels of fidelity (Century, et al., 2012; Hume et al., 2011)
- Use fidelity data to confirm delivery and interpret program impacts (Durlak & Dupre, 2008; Hulleman & Cordray, 2009)
Background and Rationale

• Several frameworks have recently been published on how to define and approach measuring intervention-specific elements (Century, et al., 2010; Nelson, et al., 2012)
• No guidance yet on the process of collecting and pooling implementation data (often large amounts from varying sources) into quantified values
• No systematic method of analysis that enables fidelity results to be compared across different interventions, across evaluations
• Fidelity measurement is tied to specific intervention - assessment of the degree to which unique elements of a preschool curriculum is implemented, for example
• Systems of measurement are often developed and presented in isolation with few connections across educational interventions
Focus and Advantages of Measuring Implementation

- Able to monitor the delivery of inputs and confirm that the foundation of the intervention has been laid
- Provide funders with confirmation that grantees have indeed delivered necessary intervention inputs for which they have received money to do so
- Collecting and analyzing this type of implementation data is less expensive than other methods and instruments used in the process are fairly simple and reliable (e.g., documentation of attendance, etc.)
- [Note: This approach measures the initial components of the hypothesized theory of change which has a weak and indirect relationship with distal outcomes]
Measurement system employs 6-phase approach to assessing fidelity of implementation

1. Develop an intervention logic model
2. Identify key components of the intervention
3. Operationalize key components into measurable, score-able indicators, specifying data sources and data collection plan
4. Calculate scoring range for each indicator, a summary range across indicators, and a threshold for each key component
5. Roll up thresholds across schools (or other unit) to a sample level
6. Produce cross-component, sample-level determination of fidelity for each year of implementation
The MA Department of Elementary and Secondary Education (ESE) launched the Expanded Learning Time (ELT) initiative, and 16 school districts received grants from the state to explore adding time to and redesigning their school day. Abt carried out a 6-year evaluation of the ELT Initiative.

The analysis presented here uses data collected during one year of the evaluation and retrospectively applies this system of measurement.

Abt, ESE, and the National Center on Time and Learning (formerly Mass 2020) developed a scoring system and thresholds at the school-level.

These thresholds were used to inform this analysis and reporting process.

The authors of this paper further developed thresholds at the sample-level (across schools) and across components without direct consultation with any of the original developers and evaluators.
Logic Model and Key Components

Phase 1

- Inputs
  - School-Wide dissemination of instructional focus and assessment analysis
  - Scheduling of instructional time and academic supports
  - Enrichment activities and specials
  - Teacher professional development
  - Principal support
  - ELT support

- Outputs

- Short-term Outcomes
  - Engagement
  - Communication and problem-solving skills
  - Positive behavior

- Long-term Outcomes
  - Academic achievement

Phase 2

- Opportunities for educators to plan, collaborate, and participate in PD

- Enrichment opportunities that engage students in
## Scoring and Thresholds for Key Components

### Phase 3

<table>
<thead>
<tr>
<th>Key Elements of Component</th>
<th>Operational Definition for Indicator</th>
<th>Data Sources</th>
<th>Data Collection Schedule</th>
<th>Fidelity Scoring System (School)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 1: School-wide dissemination of instructional focus and assessment analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator 1: Academic focus</td>
<td>School-wide Academic Focus (SWAF) is identified by teachers</td>
<td>Teacher survey</td>
<td>Spring (Year 1)</td>
<td>0 = 75% and fewer teachers identify focus; 1 = &gt;75% teachers identify</td>
</tr>
<tr>
<td>Indicator 2: Dissemination of focus</td>
<td>Academic focus is posted publically or disseminated</td>
<td>Teacher survey</td>
<td>Spring (Year 1)</td>
<td>0 = 75% or fewer teachers agree; 1 = &gt;75% teachers agree</td>
</tr>
<tr>
<td>Indicator 3: Academic-based instructional practices</td>
<td>SWAF influence of instructional practices</td>
<td>Teacher survey</td>
<td>Spring (Year 1)</td>
<td>0 = 50% or fewer teachers report influence; 1 = &gt;50% teachers report</td>
</tr>
<tr>
<td>Indicator 4: Data analysis system</td>
<td>Progress monitoring / instructional adjustment through data analysis</td>
<td>Teacher survey</td>
<td>Spring (Year 1)</td>
<td>0 = 75% or fewer teachers report data are used to monitor student progress and to adjust instructional practices; 1 = &gt;75% teachers report</td>
</tr>
</tbody>
</table>

**School level score range (0-4)**

**School level threshold: High fidelity = score of 3 or above**
## Sample Level Thresholds and Determination

### Phase 5

<table>
<thead>
<tr>
<th>Component</th>
<th>Implementation</th>
<th>Sample Level Threshold</th>
<th>Sample level Score</th>
<th>Determination of Fidelity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 1</td>
<td>Implementation of Instructional Focus</td>
<td>With fidelity = 75% or more schools score “high fidelity” (i.e., score of 3 or above)</td>
<td>14 out of 18 schools scored high (78%)</td>
<td>Yes</td>
</tr>
<tr>
<td>Component 2</td>
<td>Implementation of time and supports</td>
<td>With fidelity = 75% or more schools score “high fidelity” (i.e., score of 4 or above)</td>
<td>12 out of 18 schools scored high (67%)</td>
<td>No</td>
</tr>
<tr>
<td>Component 3</td>
<td>Implementation of Enrichment Activities</td>
<td>With fidelity = 75% or more schools score “high fidelity” (i.e., score of 5 or above)</td>
<td>13 out of 18 schools scored high (72%)</td>
<td>No</td>
</tr>
<tr>
<td>Component 4</td>
<td>Implementation of Teacher Professional Development</td>
<td>With fidelity = 75% or more schools score “high fidelity” (i.e., score of 3 or above)</td>
<td>5 out of 18 schools scored high (28%)</td>
<td>No</td>
</tr>
<tr>
<td>Component 5</td>
<td>Implementation of Principal Support</td>
<td>With fidelity = 75% or more schools score “high fidelity” (i.e., score of 3 or above)</td>
<td>8 out of 18 schools scored high (44%)</td>
<td>No</td>
</tr>
<tr>
<td>Component 6</td>
<td>Implementation of ELT Support</td>
<td>With fidelity = 75% or more schools score “high fidelity” (i.e., score of 3 or above)</td>
<td>16 out of 18 schools scored high (89%)</td>
<td>Yes</td>
</tr>
<tr>
<td>Across components</td>
<td></td>
<td>With fidelity = at least 3 of 6 components are implemented with fidelity</td>
<td>2 out of 6 (33%) components were implemented with fidelity</td>
<td>No</td>
</tr>
</tbody>
</table>
Other Considerations

- When measuring indicators
  - Indicators may be initially measured at different levels (e.g., teacher, school, district)
  - Multiple data sources could be used to assess a single indicator
- When establishing thresholds
  - Thresholds can vary across indicators
  - Thresholds set at the indicator and component level can vary from year to year
- When designing systems of data collection and reporting
  - How beneficial is it to combine across components?
  - What sample should be include (all schools? same as impact analysis? Random?)
  - Should indicators/components be weighted differently?
  - How frequent is the need to accommodate models in which key components of the intervention change from year to year?
  - Should measurement system be applied every year of implementation?
Discussion

• What does using this system tell us about Mass ELT?
  – Analysis indicates that only 2 of 6 key components were adequately implemented (using sample-level thresholds established by the authors)
  – Gives us reason to believe that supports and resources essential for ELT implementation at the intervention-, process-, mediator-level were not enough to expect widespread program success

• How does this system benefit educational evaluation?
  – Systematizes implementation fidelity data collection, analysis, and reporting across potentially divergent interventions
  – Provides guidance for evaluators as to what to measure, why, and how to use implementation data in a meaningful way
  – Allows large-scale evaluations across multiple studies (e.g., national evaluation of i3) a means to report common elements of implementation
Author Contact Information

Catherine Darrow
<cattherine_darrow@abtassoc.com>

Barbara Goodson
<barbaragoodson@gmail.com>

Beth Boulay
<beth_boulay@abtassoc.com>