Developing and Refining a Model for Measuring Implementation Fidelity for an Instructionally Embedded Assessment System

Jennifer L. Kobrin

Accessible Teaching, Learning, & Assessment Systems
Instructionally embedded assessments are designed to lead directly to action by teachers and students.

• Theory of Action (ToA).
  • Represents an assessment’s intended effects, components and action mechanisms (e.g., Bennett, 2010; FAST SCASS, 2018).
  • Action mechanisms connect an assessment system’s components to the assessment’s intended effects.

• It is important to measure the extent to which action mechanisms take place.
IMPLEMENTATION FIDELITY

• Implementation fidelity, common in evaluation research, can guide the evaluation of action mechanisms in an assessment’s ToA.
  • “The extent to which an enacted program is consistent with the intended program model” (Century et al., 2010, p. 202).

• Comparison between the critical components of intended program model and the components that are actually enacted.
CENTURY ET AL. (2010) FRAMEWORK

Structural – Procedural
What a teacher needs to do

Structural – Educative
What a teacher needs to know

Instructional – Pedagogical
Expected teacher actions

Instructional – Student Engagement
Expected student actions

Accessible Teaching, Learning, & Assessment Systems

The University of Kansas
Achievement & Assessment Institute
• Illustrate a six-step iterative process for developing and evaluating a model of implementation fidelity for an instructionally embedded assessment system.

• Conduct exploratory analyses to collect initial evidence for the validity of the implementation fidelity model.
Dynamic Learning Maps Alternate Assessment (DLM)

• Measures alternate achievement standards in English language arts (ELA) and mathematics for students with the most significant cognitive disabilities.

• Instructionally embedded model has two 15-week administration windows (fall and spring).

• Teachers choose standards and level of assessments within blueprint requirements.
Excerpt from DLM Theory of Action

(a) Educators administer assessments with fidelity.

(b) Students interact with the system to show their knowledge, skills, and understandings.

(c) Mastery results indicate what students know and can do.

(d) Results can be used for instructional planning, monitoring, and adjustment.

(e) Educators make instructional decisions based on data.

(f) Students make progress toward higher expectations.

Accessible Teaching, Learning, & Assessment Systems
(a) Educators administer assessments with fidelity.

(b) Students interact with the system to show their knowledge, skills, and understandings.

(c) Mastery results indicate what students know and can do.

(d) Results can be used for instructional planning, monitoring, and adjustment.

(e) Educators make instructional decisions based on data.

(f) Students make progress toward higher expectations.

Accessible Teaching, Learning, & Assessment Systems
6-Step Iterative Process:

1. Develop logic model identifying critical components
2. Identify process data and indicators
3. Develop hypotheses about expected patterns and define criteria for implementation fidelity
4. Conduct analyses to test the hypotheses
5. Use results to refine indicators and criteria
6. Evaluate strength of evidence and identify gaps
STEP 1: LOGIC MODEL

This cycle is completed at least once for each blueprint requirement per window.

Plan
Teacher reviews blueprint requirements.
Teacher creates instructional plan(s): chooses standard(s) and level(s) for instruction.
Teacher adjusts level.

Instruct
Teacher provides instruction in selected standard(s).

Assess
Teacher administers test/test(s) to student.

Evaluate
Teacher accesses results.

Provide more instruction on standard
Select different standard
Teacher re-assesses standard at same or different level.

Critical component at blueprint level
Critical component at standard level
Optional component at standard level

Accessible Teaching, Learning, & Assessment Systems
<table>
<thead>
<tr>
<th>Step</th>
<th>Century et al. (2010) critical component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan</td>
<td>Structural–procedural</td>
<td>Completing blueprint requirements and creating instructional plans</td>
</tr>
<tr>
<td></td>
<td>Instructional–pedagogical</td>
<td>Adjusting levels for assessment</td>
</tr>
<tr>
<td>Instruct</td>
<td>Instructional–pedagogical</td>
<td>Providing instruction on selected standard(s)</td>
</tr>
<tr>
<td>Assess</td>
<td>Structural–procedural</td>
<td>Administering assessment(s) according to published procedures</td>
</tr>
<tr>
<td>Evaluate</td>
<td>Instructional–pedagogical</td>
<td>Viewing reports and using results to make instructional decisions</td>
</tr>
<tr>
<td>Re-Assess</td>
<td>Structural–procedural</td>
<td>Administering assessment(s) according to published procedures</td>
</tr>
<tr>
<td></td>
<td>Instructional–pedagogical</td>
<td>Choosing to re-assess students at the same level or a different level to assess mastery or progress</td>
</tr>
<tr>
<td>Outside system</td>
<td>Structural–educative</td>
<td>Completing required training to administer assessments</td>
</tr>
<tr>
<td></td>
<td>Instructional–student eng.</td>
<td></td>
</tr>
</tbody>
</table>
**STEP 2: IDENTIFY PROCESS DATA & INDICATORS**

This cycle is completed at least once for each blueprint requirement per window.

1. **Blueprint coverage**

   - **Plan**
     - Teacher reviews blueprint requirements.
     - 2. *First levels chosen*
     - Teacher chooses standard(s) and level(s) for instruction.
     - Frequency of adjustment of ILS: upward and downward
     - 3. *Subsequent levels chosen*

   - **Instruction**
     - Teacher provides instruction in selected standard(s).
     - 4. *Timing of testlet creation*
     - 5. *Timing of testlet completion*
     - 6. *Number of testlets*
     - 7. *Number of days between first and last testlet*
     - 8. *Number of days between each testlet*
     - 9. *Number of standards assessed*
     - 10. *All standards assessed*

   - **Evaluate**
     - Teacher accesses results.
     - Teacher uses results for instructional decision making.
     - 11. *Number of standards reassessed same or different level.*

**Provide more instruction.**

**Select different.**

---

**Accessible Teaching, Learning, & Assessment Systems**

**KU THE UNIVERSITY OF KANSAS**

**Achievement & Assessment Institute**
## STEP 3: IMPLEMENTATION FIDELITY CRITERIA

<table>
<thead>
<tr>
<th>Implementation Level</th>
<th>Criteria</th>
</tr>
</thead>
</table>
| 1                    | Blueprint coverage not met  
All assessments assigned and completed within one week  
All possible content standards assessed |
| 3                    | Met or exceeded blueprint coverage  
Time between first and last assessment is at least 60 days  
Median days between assessments suggests adequate time for instruction  
At least one content standard is re-assessed |

**Accessible Teaching, Learning, & Assessment Systems**

**KU The University of Kansas**

**Achievement & Assessment Institute**
• Steps 4 and 5 conducted in tandem in a few iterative cycles.
  • Conducted first round of analysis on the indicators, presented findings to our TAC and received feedback.
  • Revised indicators and criteria and ran additional analyses.
STEPS 4-5: ANALYSES & REFINEMENT

• Data from the DLM for the fall 2019 administration
  • 14,021 students in grades 3-11
  • 4,505 teachers
• Descriptive statistics for the indicators by implementation level
• Effect sizes and odds ratios for pairwise mean differences in indicators
• 31% in Level 1, 68% in Level 2, and 0.5% in Level 3.
• Larger percentage in Level 1 in math compared to ELA.
• Many of the indicators differentiate the three implementation levels according to our hypotheses.
  • Level 3 - greater number of assessments, longer testing window, more spacing between assessments and more frequent re-assessment.
  • Level 1 - most likely to complete all testing in either the first or last 20% of the assessment window and least likely to meet threshold for median days between assessments.
STEP 6: EVALUATE STRENGTH OF EVIDENCE

• Current gaps in the implementation fidelity model:
  • The actual amount of instructional time spent on standards.
  • The extent and ways in which teachers access and use assessment results (Evaluate).
  • The instructional-student engagement critical component.
CONCLUSIONS & NEXT STEPS

• Replicate analyses on future years’ data and continue to refine indicators
• Explore alternative hypotheses
• Develop profiles of instructionally embedded assessment use
• Explore relationships of implementation fidelity to student outcomes
• Collect qualitative data to examine teachers’ assumptions, motivations and rationales
Thank you!

Contact me at:

jennifer.kobrin@ku.edu