Toward a Framework for Studying Research-Based Science Curricula

Janet Carlson
Joseph Taylor
Susan Kowalski
Christopher Wilson

24 March 2010
NARST International Conference
Philadelphia, PA

www.BSCS.org
Purposes

• Provide an overview of a framework for curriculum-based research
  – Few examples in the science education literature
  – Parallels an example from mathematics education
  – Offer a perspective on the role of research in curriculum development process

• Describe studies that
  – Establish feasibility of implementation (field test)
  – Establish promise of positive outcomes for all students (field test)
  – Make causal inferences of effectiveness in specified conditions and for specified student populations (efficacy)
...curriculum is a written instructional blueprint and set of materials for guiding students’ acquisition of certain culturally valued concepts, procedures, intellectual dispositions, and ways of reasoning (Clements, 2007, p. 36).
### Clements’ Basic Ideas

<table>
<thead>
<tr>
<th>Category</th>
<th>Questions Asked</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Prio <em>ri</em> Foundations</td>
<td>What is already known that can be applied to the anticipated curriculum?</td>
</tr>
<tr>
<td>Learning Models</td>
<td>How might the curriculum be constructed to be consistent with models of students’ thinking and learning?</td>
</tr>
<tr>
<td>Evaluation</td>
<td>How can market share for the curriculum be maximized?</td>
</tr>
<tr>
<td></td>
<td>Is the curriculum usable by and effective with various student groups and teachers?</td>
</tr>
<tr>
<td></td>
<td>What is the effectiveness of the curriculum, now in its complete form, as it is implemented in realistic contexts?</td>
</tr>
</tbody>
</table>
What is a Curriculum Study?

A curriculum study is an organization that determines its work by considering a philosophical perspective as well what research says about learning, teaching, and the role of curriculum and endeavors to translate that research into programs and generate new knowledge to further the field.
A Priori Foundations

- Scientific content must be represented with accuracy and integrity.
- Scientific inquiry is critical to conveying the nature of the disciplines of science.
- Our approach to learning is driven by the findings of rigorous research.
- All students are capable and have the right to learn science.
- Curriculum materials have a central role in the teaching and learning of science.
- Transformative professional development is essential for deep teacher learning.
Learning Model

• Engage – Provide an opportunity for mental engagement in a key concept.
• Explore – Provision of common experiences to “level the playing field.”
• Explain – Provide students with opportunities to construct scientific explanations.
Learning Model, continued

• Elaborate – Provides students the opportunity to transfer understandings to new contexts; students may begin to make abstractions or generalizations pertaining to the major concept of the unit.

• Evaluate – Provides teachers and students opportunities to assess understanding.
Relationship Between our Learning Model and our A Priori Foundations

• The 5Es emphasize **scientific accuracy and integrity** because of the role of explanations and evidence.
• The model stresses an active role of the learner in gaining understanding of science concepts, which supports **inquiry**-based instruction very well.
• **Research** in learning, literacy, ELL, and science informs the curriculum development process.
Relationship Between our Learning Model and our A Priori Foundations

• The model emphasizes starting with the understanding in each learner’s mind, students construct their explanations based on their own knowledge and experiences. Common classroom experiences contribute to explanations. This honors the individual while providing for common experiences that enhances the opportunities for Equity.

• Materials can enhance the coherence, rigor and focus of the content and the sequencing and strengthen the role of the learning model.

• Transformative professional development challenges current understandings and beliefs to improve teaching practice.
Curriculum Development

1. **Design**: We begin designing a unit by using the *Understanding by Design* process.
2. **Development**: Create a prototype version of the curriculum materials.
3. **Field Testing of Prototype**: Program tested in volunteers’ classroom.
4. **Revision**: Data from field tests inform revision of the unit.
5. **Second Field Test Cycle**: Ideally, phases 3 and 4 are repeated.
6. **Production and Publication**: After all field-testing and revision cycles are complete, the materials go into a final production phase.
7. **Efficacy Trial of Publicly Available Program**: This study aims to compare outcomes from the program to a comparison program for a limited population.
8. **Effectiveness Study**: compare outcomes from the program to a comparison program for a broad population.
# Types of Curriculum Research Studies

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Primary Goal</th>
<th>Potential for Causal Claims</th>
<th>Potential to Generalize</th>
</tr>
</thead>
</table>
| Design Study     | Define the learning problem that warrants a new program  
Examine the landscape through a review of the literature, surveys, interviews, and case studies to identify why the problem exists  
Propose core features of the new program based on the landscape assessment  
Define the implementation requirements and the indicators of success | NA                          | NA                      |
## Types of Curriculum Research Studies

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Primary Goal</th>
<th>Potential for Causal Claims</th>
<th>Potential to Generalize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Test</td>
<td>Establish usability of the prototype program</td>
<td>Low</td>
<td>Varies</td>
</tr>
<tr>
<td></td>
<td>Establish feasibility of implementation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Establish promise of positive outcomes for all students</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Types of Curriculum Research Studies

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Primary Goal</th>
<th>Potential for Causal Claims</th>
<th>Potential to Generalize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy</td>
<td>Make causal inferences of effectiveness in specified (often ideal) conditions and for specified student populations</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Make causal inferences of effectiveness in varied, authentic conditions and for varied student populations Establish effectiveness in the absence of extraordinary implementation support from developers</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
Curriculum Field Test Study
Example 1: Focus on Fidelity of Implementation

Joseph Taylor, BSCS
Janet Carlson, BSCS
Pamela Van Scotter, BSCS
Doug Coulson, PS International

www.BSCS.org
Need for the Study

- NSF has funded materials development for decades
- Theoretical frameworks recommend that learning be sequenced and organized using a learning cycle (Atkin & Karplus, 1962; Piaget, 1975) or an instructional model (Bybee, 1997)
- *How People Learn* (NRC, 2001) suggests that teachers address students’ prior knowledge, help students connect new knowledge to a rich framework of big ideas, and support students in monitoring and taking control of their own learning.
- Limited research on how these materials influence student learning and the role that implementation fidelity plays in the materials’ ultimate impact.
A *Priori* Foundations

- Scientific content must be represented with accuracy and integrity.
- Scientific inquiry is critical to conveying the nature of the disciplines of science.
- Our approach to learning is driven by the findings of rigorous research.
- All students are capable and have the right to learn science.
- *Curriculum materials have a central role in the teaching and learning of science.*
- Transformative professional development is essential for deep teacher learning.
Primary Research Question

What is the relationship between the teacher’s level of fidelity of implementation and student achievement?
Context for the Study

The professional development program included:

- Workshops on the goals and purpose of the instructional model (5Es). Teachers were engaged as learners of science in investigations where the facilitator modeled exemplary use of the 5E model.
- Workshops on inquiry and effective ways to create a climate of inquiry
- Sessions on *Understanding by Design* which provided key principles for the design of the program.
- Focused sessions on science content

This work was funded by National Science Foundation Grant No. ESI-9911614
Student Sample: 326 9th graders
Teacher Sample: 15 randomly selected from 30 teachers

- Colorado
- Florida
- Illinois
- Massachusetts
- Michigan
- Tennessee
- Vermont
- Washington
- West Virginia
- Wisconsin

Field-Test II Sites
2002
Measures

• Student learning: scores from identical pre-post chapter test
  – Multiple choice and constructed response items
• Fidelity of Implementation: observation protocol adapted from the Horizon, Inc. Classroom Observational Protocol (HRI, 2000)
  – 11 individual rating scales that addressed the use of specific teaching strategies consistent with the BSCS 5E Instructional Model.
  – Based on 11 ratings, holistic classifications each teacher’s fidelity of use as either “low,” “medium,” or “high.”
  – *This protocol emphasizes adherence to the curriculum as written.*
# Number of Teachers at Each Fidelity Level

<table>
<thead>
<tr>
<th>Observed use of strategies and learning sequences that are consistent with the 5Es</th>
<th>Holistic Rating of Fidelity Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Minimal</td>
</tr>
<tr>
<td></td>
<td>Basic</td>
</tr>
<tr>
<td></td>
<td>Extensive</td>
</tr>
<tr>
<td>Number of Teachers Rated at each Fidelity Level</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>
When we clustered students by their teachers’ fidelity level...

- Student **post-test** scores appeared to vary by teacher fidelity level

- Student **pre-test** scores also appeared to vary by teacher fidelity level
Adjusted Post-test Means by Level of Fidelity
## Effect Sizes (Cohen’s d)

<table>
<thead>
<tr>
<th>Teacher Group</th>
<th># of Students</th>
<th>Adjusted Mean Post-test Score (% correct)</th>
<th>Standard Deviation</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Level Fidelity</td>
<td>70</td>
<td>41</td>
<td>2.9</td>
<td>4.7 (between low and medium)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.0         (between low and high)</td>
</tr>
<tr>
<td>Medium Level Fidelity</td>
<td>168</td>
<td>54</td>
<td>2.6</td>
<td>1.3         (between medium and high)</td>
</tr>
<tr>
<td>High Level Fidelity</td>
<td>88</td>
<td>51</td>
<td>2.0</td>
<td></td>
</tr>
</tbody>
</table>
Primary Finding

Students whose teachers used the instructional materials with medium or high levels of fidelity scored considerably higher on the post-test achievement measure than students of low fidelity teachers.

Note: Complementary findings can be found in similar studies of research-based instructional models (e.g., Ates, 2005; Ebrahim, 2004; Lord, 1997).
Discussion

• Why didn’t extensive use of the 5Es yield student achievement that was significantly higher from basic use of the instructional model?
  – It is possible that the materials are fidelity-resilient (within limits) - even when a teacher implements the material with only medium fidelity, the inherent design of the instructional materials likely contributes to enhanced understanding for students.
  – Extensive use under field test conditions is extensive adherence

• Possible additional explanation
  – Some teachers were rated holistically at medium levels of fidelity because their use of the 5Es was not consistent. It is possible that a large percentage of these teachers made extensive use of the most effective aspects of the 5Es.
Implications

• Curriculum materials that reflect research on learning make a difference in student achievement, but do not replace the teacher.

• Conduct an analysis across selected individual fidelity rating scores from the observation protocol. This analysis would help us make more direct connections between student achievement and the use of specific instructional strategies within the 5E instructional model.

• Test differences in student achievement as a function of fidelity using *hierarchical linear modeling* to account for the nested data (larger teacher sample size).
Instructional Materials, Equity, and Science Achievement

Susan Kowalski
Pamela Van Scotter
Molly Stuhlsatz
Joseph Taylor

www.BSCS.org
Goals of this Presentation

• Provide an example of a field test study and how it relates to the curriculum development framework
• Present findings from the field test related to student achievement and equity
• Situate the findings within the broader context of the literature on instructional materials and equity
Framework: Alignment of this Study with *A Priori* Foundations

- **Aspects of study**
  - Study grounded in the field test of *BSCS Science: An Inquiry Approach, Level 3*, under development at the time
  - Materials designed to be inquiry- and standards-based

- **A Priori Foundations**
  - Curriculum materials have a central role in the teaching and learning of science.
  - Scientific content must be represented with accuracy and integrity.
  - Scientific inquiry is critical to conveying the nature of the disciplines of science.
Framework: Alignment of Study with *A Priori* Foundations (continued)

- Aspects of study
  - Materials designed with careful attention to research on effective teaching and learning, including:
    - Preconceptions
    - Metacognition
    - Coherence
    - Integrated use of student notebooks
    - Sense-making strategies
    - Literacy strategies
    - Collaborative learning
    - Multiple representations of ideas
    - Socially relevant contexts (as much as possible within materials intended for a national audience)
    - Educative teacher materials

- *A Priori* Foundation
  - Our approach to learning is driven by the findings of rigorous research.
Framework: This study and the Learning Model

- Materials developed within the framework of the BSCS 5E Learning Model

- Reviews involve careful consideration of whether or not writers have remained faithful to the spirit of each “E” during development.

- Attend to teacher integrity to the spirit of each “E” during classroom observations
Framework: This Study and Curriculum Development Research

• Purpose of this study
  – Assess the usability and feasibility of the materials within authentic classroom settings
  – Determine the extent to which the materials show the promise of efficacy
    • Learning gains should be larger than might be expected from commonplace instruction
    • Learning gains should, at a minimum, be similar across student demographic groups

• Purpose of field test research within our curriculum development framework
  – Establish usability of the prototype program
  – Establish feasibility of implementation
  – Establish promise of positive outcomes for students

• Alignment of purpose with *a priori* foundation: All students are capable and have the right to learn science.
Broader Need for the Study

• According to surveys conducted by the National Science Foundation (NSF), women and minorities do not participate in the science and engineering workforce at rates consistent with their populations (Rosser, 2000).

• In spite of calls for “science for all,” achievement gaps by race/ethnicity, gender, and socioeconomic status remain (Boe & Shin, 2005; Clewell & Campbell, 2002)

• Effective teaching matters in closing achievement gaps (Johnson, Kahle, & Fargo, 2007)

• When teachers lack time and/or content knowledge, they rely heavily on instructional materials (Ball & Feiman-Nemser, 1998)

Can high quality instructional materials help all teachers be effective across diverse groups of students?
About the Teachers

• 18 teachers from across the US (MI, WV, NH, OH, VT, FL, PA, OR, OK, PA, WA)
• Varied level of experience
• Many teachers were new to multidisciplinary science
• None of the teachers were well-practiced in teaching science as inquiry before implementing the curriculum.
• Taught in urban, suburban, and rural schools
• Teachers experienced a 2-day professional development session on the theory and use of the materials
About the Students

- N = 410 self-identified as
  - Hispanic or Latino/a (N = 7)
  - Asian (N = 10)
  - African American (N = 47)
  - American Indian or Alaska Native (N = 4)
  - White (N = 283)
  - Native Hawaiian (N = 3)
  - More than one race/ethnicity (N = 36)

- For statistical analyses, we grouped students as either White (N = 283) or as members of a minority group (N = 107)
- Receive FRL: Yes (N = 78); No (N = 312)
- Male (N = 217); Female (N = 173)
Assumptions

• Student pretest scores predict posttest scores (Schochet, 2005)
• Student math level predicts posttest scores (Sadler & Tai, 2007)
• Achievement gaps often expand over the course of instruction when teachers use commonplace instructional materials (Lynch et al., 2005; Wilson et al., 2010)
Research Questions

1. When teachers use the instructional materials with students, do students show promising gains from pretest to posttest?

2. When teachers use the instructional materials with students, does gender, race/ethnicity, or free or reduced-price lunch status predict variation in student post-test scores, above and beyond variation accounted for by pre-test score and math level?
Measures

• Pre/Posttests
  – Released TIMSS, NAEP items
  – Original items developed when necessary to align with materials
• Demographic survey
• Teacher surveys of materials
• Student surveys of materials
• Classroom observations
• Informal interviews with field test teachers
Preliminary Analyses

• Classical Item Analysis of Pre/Posttest items
• Elimination of items that analyses showed were particularly problematic
• Cronbach’s alphas:
  – Earth/Space Science, $\alpha = .836$
  – Life Science, $\alpha = .746$
  – Physical Science, $\alpha = .736$
  – Science & Society, $\alpha = .631$
Primary Analyses

• Matched-pairs t-test for comparison of pretests to posttests
• OLS Hierarchical Regression
• Order of predictor variables added to the model such that those expected to account for the greatest variation in posttest scores were added first:

Step 5 Model:

$$\hat{Y}_{posttest} = b_0 + b_1 X_{pretest} + b_2 X_{math} + b_3 X_{lunch} + b_4 X_{race} + b_5 X_{gender}$$

(Schochet, 2005; Sadler & Tai, 2007; Rothstein, 2004; Muller, Stage, & Kinzie, 2001)
Primary Analyses

• Calculated an F test of change for each model; $\alpha = .05$
• Model assumptions were met, including
  – normality of residuals
  – homogeneity of variances
  – the presence of a linear relationship between the covariate and $Y$
  – homogeneity of regression
• Regression diagnostics showed little evidence of outliers
• Students in classrooms were from intact groups; thus, independence of residuals is suspect.
Primary Analyses, continued

- Correlation existed between pretest score and math level, but variance inflation factors remained well within acceptable limits (VIF < 2.3).
- No significant correlation existed between pretest and any other factor in the analysis.
- Estimation of mean square pure error showed no indication of model mis-specification.
- Data were found to be missing completely at random.
Findings

• Student posttest scores were significantly higher than pretest scores ($p < .001; d = 1.03$)
• Math level significantly accounted for variation in posttest scores above and beyond pretest
• FRL status, race/ethnicity, gender, did not account for variation in posttest scores above and beyond pretest and math level
Findings

\[ R^2 = .486 \quad F(4, 395) = 9.03; \quad p < .05 \]

\[ \hat{Y}_{\text{posttest}} = b_0 + b_1 X_{\text{pretest}} + b_2 X_{\text{math}} + b_3 X_{\text{lunch}} + b_4 X_{\text{race}} + b_5 X_{\text{gender}} \]

\[ R^2 = .496 \quad F(3, 381) = 2.52; \quad p > .05 \]
Summary of Findings

• Students showed large, significant gains from pretest to posttest
• Students demonstrated similar content gains, regardless of demographic characteristics, once pretest and math level were taken into account
• Preexisting achievement gaps neither expanded nor diminished
Findings in Context of Curriculum Research Literature

• Lynch et al. (2005) compared the highly-rated Chemistry that Applies (CTA) to commonplace instruction. They found that the use of commonplace materials expanded gaps.
• Like the BSCS materials, CTA neither expanded nor closed gaps.
• Lee et al. (2005) found that the use of reform-based materials with professional development designed to incorporate students’ languages and cultures actually reduced achievement gaps for some subgroups of students.
Commonalities among Curricula: *BSCS Science: An Inquiry Approach*, CTA, and Lee et al. materials

- Inquiry-based
- Emphasis on collaborative learning
- Use of science notebooks for CTA and BSCS
- Professional development related to the theory behind and use of instructional materials
- Focused on a small number of “big ideas”
- Emphasis on describing, explaining, reporting, and drawing conclusions
- Use of literacy strategies (BSCS and Lee et al.)
- Educative teacher materials
Differences

• Lee et al.’s materials incorporated:
  – Vocabulary provided in English, Spanish, and Haitian-Creole
  – Strategies for teachers to incorporate students’ linguistic and cultural experiences in science instruction
  – Ideas for how teachers could adapt strategies for their students’ own linguistic and cultural contexts

• CTA and BSCS materials did not specifically attend to students’ languages and cultures.
Discussion

• Some studies have shown that commonplace materials tend to result in expanded achievement gaps (Lynch et al., 2005; Wilson et al., 2010).

• Inquiry- and Standards-based materials that attend to research on teaching and learning in the absence of attention to students’ languages and cultures appear to produce no increase or decrease in achievement gaps.

• Inquiry- and Standards-based materials that attend to both research on teaching and learning and students’ languages and cultures appear to have the ability to narrow gaps.

• Curriculum developers need to negotiate the tension between designing materials for a national audience and designing materials that incorporate students’ languages and cultures.
Curriculum Efficacy Studies – Comparisons to Commonplace Curriculum and Teaching

Christopher Wilson
Joseph Taylor
Susan Kowalski
Janet Carlson

www.BSCS.org
**A Priori Foundations**

- Scientific inquiry is critical to conveying the nature of the disciplines of science.
- Our approach to learning is driven by the findings of rigorous research.
- All students are capable and have the right to learn science.
- Curriculum materials have a central role in the teaching and learning of science.

BSCS 5E Instructional Model
A Priori Foundations: Scientific inquiry is critical to conveying the nature of the disciplines of science.

- The uncommonness of inquiry-based teaching.
- Conflicts in the ages of reform and accountability.
- Recent challenges to inquiry-based teaching.
- Evidence on achievement gaps.
Curriculum Study R&D Process:  
*Efficacy Study of a Program*

Compares outcomes from the program to a comparison program for a limited population.

Results inform future editions of the curriculum materials.
Curriculum Study R&D Process:
**Efficacy Study of an Instructional Model**

Compares outcomes from a program organized around an instructional model to the same program organized around a different instructional approach (for a limited population).

Results inform:
- Internal R&D
- Decision-makers
- *A priori* foundations
A Priori Foundations

- Scientific inquiry is critical to conveying the nature of the disciplines of science.
- Our approach to learning is driven by the findings of rigorous research.
- All students are capable and have the right to learn science.
- Curriculum materials have a central role in the teaching and learning of science.
Commonplace Teaching ↔ Inquiry-Based Teaching (5Es)
Primary Research Questions

1. To what extent can differences in student learning between the 5E-based and commonplace groups be attributed to randomized group assignment?

2. What differences in achievement by treatment group exist specific to the learning goals of knowledge, reasoning, and argumentation?

3. Does student race/ethnicity, gender, or socio-economic status account for variation in posttest scores above and beyond variation accounted for by pretest scores and group assignment?
Strengthened for stronger alignment with inquiry teaching.

Rebuilt around commonplace teaching strategies as defined by:

- The 2000 National Survey of Science and Mathematics Education.
The Commonplace Unit: Examples from the HRI data

<table>
<thead>
<tr>
<th>Activity</th>
<th>Never</th>
<th>A few times a year</th>
<th>Once or twice a month</th>
<th>Once or twice a week</th>
<th>All or most lessons</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Students) Design or implement their <em>own</em> investigation</td>
<td>8%</td>
<td>42%</td>
<td>41%</td>
<td>8%</td>
<td>1%</td>
</tr>
<tr>
<td>(Students) Make a formal presentation to the rest of the class</td>
<td>17%</td>
<td>49%</td>
<td>29%</td>
<td>5%</td>
<td>1%</td>
</tr>
<tr>
<td>(Teacher) Conduct a pre-assessment to determine what students already know</td>
<td>16%</td>
<td>38%</td>
<td>29%</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>(Teacher) Introduce content through formal presentations</td>
<td>0%</td>
<td>3%</td>
<td>15%</td>
<td>59%</td>
<td>22%</td>
</tr>
</tbody>
</table>
## The Commonplace Unit: Examples from the HRI data

<table>
<thead>
<tr>
<th>Activity</th>
<th>Never</th>
<th>A few times a year</th>
<th>Once or twice a month</th>
<th>Once or twice a week</th>
<th>All or most lessons</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Students) Design or implement their <em>own</em> investigation</td>
<td>8%</td>
<td>42%</td>
<td>41%</td>
<td>8%</td>
<td>1%</td>
</tr>
<tr>
<td>(Students) Make a formal presentation to the rest of the class</td>
<td>17%</td>
<td>49%</td>
<td>29%</td>
<td>5%</td>
<td>1%</td>
</tr>
<tr>
<td>(Teacher) Conduct a pre-assessment to determine what students already know</td>
<td>16%</td>
<td>38%</td>
<td>29%</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>(Teacher) Introduce content through formal presentations</td>
<td>0%</td>
<td>3%</td>
<td>15%</td>
<td>59%</td>
<td>22%</td>
</tr>
</tbody>
</table>
Sample

58 students aged 14-16 participated in the study, from 24 schools from seven districts. Five of the students attended private schools and two were home-schooled.

<table>
<thead>
<tr>
<th></th>
<th>Commonplace Unit (n=28)</th>
<th>Inquiry-Based Unit (n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>61% male, 39% female</td>
<td>47% male, 53% female</td>
</tr>
<tr>
<td>Race (% non-white)</td>
<td>21%</td>
<td>23%</td>
</tr>
<tr>
<td>Age (mean)</td>
<td>15.1</td>
<td>14.9</td>
</tr>
<tr>
<td>Free/Reduced Lunch Status</td>
<td>12%</td>
<td>10%</td>
</tr>
</tbody>
</table>
Constants between the Treatment and Comparison Groups

- Both groups of students were taught toward the learning goals in the NIH-BSCS Sleep Module.
- Both groups were taught by the same master teacher.
- Both groups received five, 2-hour lessons, provided at the BSCS headquarters in Colorado Springs.
- Both groups completed the same pretests and posttests, and we’re interviewed in a post-post test four weeks after the unit.
Learning Goals

1. Knowledge

Multiple-choice items and True/False items focused on vocabulary and simple facts about sleep.
## 2. Reasoning with Scientific Models

<table>
<thead>
<tr>
<th>Level</th>
<th>Common Errors</th>
<th>Exemplar</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>Model-based accounts connected across scales</td>
<td>Despite the new light cues in London, they would still be sleeping on the Denver time because their biological clock can't reset that quickly.</td>
</tr>
<tr>
<td>4.</td>
<td>Appropriate but superficial connections between organismal and physiological systems</td>
<td>Recognizes that an internal biological clock plays a role in sleep behavior, but cannot explain how.</td>
</tr>
<tr>
<td>3.</td>
<td>Alludes to hidden physiological mechanisms</td>
<td>Some scientific vocabulary is used to suggest cellular/internal control of sleep behavior, but no specific mechanism is described.</td>
</tr>
<tr>
<td>2.</td>
<td>Accounts restricted to the organismal level</td>
<td>Observable changes occur in direct response to the environment, with no intermediate physiological mechanism.</td>
</tr>
<tr>
<td>1.</td>
<td>Stories at the organismal level based on personal experience / cultural models</td>
<td>Sleep behavior attributed to conscious effort. Ideas about the body refueling during sleep.</td>
</tr>
<tr>
<td>0.</td>
<td>No response / unintelligible / negligible</td>
<td></td>
</tr>
</tbody>
</table>

---

### Explanation:
- **Level 5**: Model-based accounts connected across scales. Responses may contain errors such as east/west time zone mix-ups, or details of REM-NREM cycling. For example: Despite the new light cues in London, they would still be sleeping on the Denver time because their biological clock can't reset that quickly.
- **Level 4**: Appropriate but superficial connections between organismal and physiological systems. Recognizes that an internal biological clock plays a role in sleep behavior, but cannot explain how. For example: I shaded this area because after 1 day in London the person will still be on the same sleep schedule as they would in Denver, CO. This is due to their biological clock.
- **Level 3**: Alludes to hidden physiological mechanisms. Some scientific vocabulary is used to suggest cellular/internal control of sleep behavior, but no specific mechanism is described. For example: The person would probably be asleep when it is morning here because their brain wasn't used to the time in England. Jetlag!!
- **Level 2**: Accounts restricted to the organismal level. Observable changes occur in direct response to the environment, with no intermediate physiological mechanism. For example: Now your body has changed to London time.
- **Level 1**: Stories at the organismal level based on personal experience / cultural models. Sleep behavior attributed to conscious effort. Ideas about the body refueling during sleep. For example: You wouldn't be tired if you slept on the plane, so you probably wouldn't go to bed until noon.
- **Level 0**: No response / unintelligible / negligible.
3. Argumentation

<table>
<thead>
<tr>
<th>Level</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Claim:</strong></td>
<td>An assertion that answers the original question.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Evidence:</strong></td>
<td>Scientific data that supports the claim. Data need to be appropriate and sufficient.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reasoning:</strong></td>
<td>A justification that links the claim and evidence, using appropriate and sufficient scientific principles.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results
Differences in Implementation

**RTOP**

<table>
<thead>
<tr>
<th>Section</th>
<th>Commonplace Unit</th>
<th>Inquiry-Based Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesson Design &amp; Implementation</td>
<td>3.0 ± 0.5</td>
<td>3.5 ± 0.7</td>
</tr>
<tr>
<td>Content</td>
<td>2.5 ± 0.4</td>
<td>3.2 ± 0.3</td>
</tr>
<tr>
<td>Classroom Culture</td>
<td>1.5 ± 0.2</td>
<td>2.0 ± 0.3</td>
</tr>
</tbody>
</table>

**CLES**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Mean CLES Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commonplace Unit</td>
<td>60 ± 2</td>
</tr>
<tr>
<td>Inquiry-Based Unit</td>
<td>62 ± 3</td>
</tr>
</tbody>
</table>

*Denotes significant difference between Commonplace and Inquiry-Based units.
## Differences in Implementation

<table>
<thead>
<tr>
<th></th>
<th>Commonplace</th>
<th>Inquiry</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classroom Activity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture</td>
<td>265</td>
<td>145</td>
</tr>
<tr>
<td>Problem Modeling</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Lecture with Discussion</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Writing Work</td>
<td>160</td>
<td>275</td>
</tr>
<tr>
<td>Small Group Discussion</td>
<td>70</td>
<td>245</td>
</tr>
<tr>
<td>Teacher/Faculty Member Interacting with Students</td>
<td>5</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Student Attention to Lesson</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (80% or more off task)</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Medium (mixed engagement)</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>High (80% or more of the students engaged)</td>
<td>435</td>
<td>475</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cognitive activity</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Receipt of knowledge</td>
<td>435</td>
<td>205</td>
</tr>
<tr>
<td>Knowledge construction</td>
<td>25</td>
<td>235</td>
</tr>
</tbody>
</table>
1. To what extent can differences in student learning between the inquiry-based and commonplace groups be attributed to randomized group assignment?

\[ \hat{Y} = b_0 + b_1 X_{\text{pretest}} + b_2 X_{\text{group}} \]

39.7%, \( p<0.001 \)

44.3%, \( p<0.05 \)
1. To what extent can differences in student learning between the inquiry-based and commonplace groups be attributed to randomized group assignment?

**Effect Size = .47**
2. What differences in achievement by treatment group exist specific to the learning goals of knowledge, reasoning, and argumentation?

**Level 5 Understanding**

- Commonplace Unit
  - Pretest: 0.0
  - Posttest: 0.4

- Inquiry-Based Unit
  - Pretest: 0.1
  - Posttest: 0.3

\[ d = 0.68 \]

**Argumentation**

- Claim
  - Commonplace Unit: 0.58
  - Inquiry-Based Unit: 0.74

- Evidence
  - Commonplace Unit: 0.58
  - Inquiry-Based Unit: 0.59

- Reasoning
  - Commonplace Unit: 0.58
  - Inquiry-Based Unit: 0.59
3. Does student race/ethnicity, gender, or socio-economic status account for variation in posttest scores above and beyond variation accounted for by pretest scores and group assignment?

\[ \hat{Y} = b_o + b_1 X_{pretest} + b_2 X_{group} + b_3 X_{lunch} + b_4 X_{race} + b_5 X_{gender} \]

- 39.7%, \( p<0.001 \)
- 44.3%, \( p<0.05 \)
- \( p>0.05 \)
3. Does student race/ethnicity, gender, or socio-economic status account for variation in posttest scores above and beyond variation accounted for by pretest scores and group assignment?
Conclusions: Science as Inquiry

• The uncommonness of inquiry-based teaching.

• Conflicts in the ages of reform and accountability.

• Recent challenges to inquiry-based teaching.

• Evidence on achievement gaps.
Conclusions: A *Priori* Foundations

- Scientific inquiry is critical to conveying the nature of the disciplines of science.
- Our approach to learning is driven by the findings of rigorous research.
- All students are capable and have the right to learn science.
- Curriculum materials have a central role in the teaching and learning of science.
This research was supported by the Office of Science Education at the National Institutes of Health.
Implications

• Carefully developed curriculum materials may be a critical tool teachers require to offer effective instruction for all students
• Teachers and leaders should focus on analyzing curriculum materials, not creating them
• Preservice programs do not always include curriculum analysis processes, but often include curriculum development
Discussion

• Framework for Curriculum Research
• Two field-test studies and an efficacy study
• What are your thoughts about
  – The *a priori* foundations?
  – The research framework?
  – The types of studies for now and in the future?
To download this presentation, visit www.bscs.org/sessions