New Directions in K-12 mathematics Curriculum Evaluation

Steven W. Ziebarth, Assoc. Prof.

WMU Mathematics Department

Evaluation Café Presentation: October 17, 2006

Talk Outline

1. Overview and work of the CSMC and its subgroups:
   a) Doctoral Program and Course Development
   b) Research and Research Groups
   c) Conferences, Proceedings, Monographs
2. Some recent historical benchmarks in K-12 mathematics curriculum development, research, and evaluation
3. NRC Report (2004): Evaluating the K-12 Evaluations
   a) Background and Structure of the Report
   b) Major Findings and Recommendations
   c) Effects on the direction of Math Curriculum Evaluation
4. Discussion, Q/A

Overview of CSMC

Mission

To advance the research base and leadership capacity supporting K-12 mathematics curriculum design, analysis, implementation, and evaluation.

Center Partners

- Michigan State University
- University of Missouri
- Western Michigan University

  - Columbia MO Public Schools
  - Grand Ledge MI Public Schools
  - Kalamazoo MI Public Schools

    - Horizon Research, Inc.
    - University of Chicago

Major Areas of Work

- Understand the influence and potential of mathematics curriculum materials [Research]
- Develop and study models for enabling teacher learning through curriculum material investigation and implementation. [Teacher Learning]
- Build capacity for developing, implementing, and studying the impact of mathematics curriculum materials. [Doctoral Program]

Examples of the Work of CSMC

- Doctoral Course Development
- Research: State Standards Analysis
- International Curriculum Conference
- Cross-Site Study of Factors Affecting Curriculum Use
- Monograph Series
- Online Resources

Resources & Opportunities

Curriculum Literature Database
Research Instrument Database
State Standards links
Doctoral Course Materials
Monographs
CSMC Research Associate Program

Recent Historical Benchmarks in K-12 Mathematics Curriculum

1. Prior to 1989, little systematic evaluation of K-12 mathematics curricula/textbooks
2. Early 1990s: NCTM Publication of the *Curriculum and Evaluation Standards for School Mathematics* (1989), *Professional Standards for Teaching Mathematics* (1991), *Assessment Standards for School Mathematics* (1995)
3. 1993-2003: NSF Funded K-12 Curriculum Development Projects
   a) Elementary (3), MS (5), HS (5); Some req. formal eval.
4. 2000: NCTM Publication of the *Principles and Standards for School Mathematics*
5. 1995- 200?: “Math Wars”

Recent Evaluation of K-12 Mathematics Curriculum

Evaluation related to K-12 mathematics curricula: Standards-Based School Mathematics Curricula

a) *Standards-Based School Mathematics Curricula: What are they? What do students learn?* (Senk & Thompson, 2003)

b) WWC (2003- ?): What Doesn't Work: The Challenge and Failure of the What Works Clearinghouse to Conduct Meaningful Reviews of Studies of Mathematics Curricula (Schoenfeld, 2005, *Educational Researcher*)

c) *On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations* (National Research Council, 2004, Jere Confrey, Chair)

The NRC Report

On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations

NRC Report: Defining Curriculum Evaluation

Taking curriculum in its broadest sense as a set of materials for use at each grade level complete with all ancillary materials.

The committee "considered the meaning of an evaluation of a curriculum for this study, . . . the study had to:"

a) Focus primarily on one of the curriculum programs or compare two or more curriculum programs,

b) Use a methodology recognized by the fields of mathematics education, mathematics, or evaluation; and

c) Study a major portion (at least one grade-level) of the curriculum program under investigation. (p.39)

NRC Report: The Charge

1. ***What it is:*** Committee charged to determine whether the currently available data are sufficient for evaluating the effectiveness of these materials AND, if not, to develop recommendations about the design of a project that could result in the generation of more reliable and valid data for evaluating these materials.

2. ***What it is not:***
   a) A direct evaluation of curriculum materials themselves
   b) An attempt to rate or rank specific curricular programs

NRC Report: Organization

1. 19 Curricula were studied:

   13 NSF Funded; 6 Commercially Published

2. 698 Studies found; eliminated 281 as not relevant; 225 were seen as providing background information; and remaining 192 categorized into “four major evaluation methodologies”

3. Content analysis (36) - expert model;

   Comparative analysis (95 - 32(not “at least min. meth. adeq” **) = 63 considered

   Case Studies (45 - 13 (lacking meth. rigor) = 32 considered

   Synthesis Studies (16)

   Total = 147 (small number and unevenness prevented conclusions about effectiveness of individual programs)

NRC Report: Content Analysis Characteristics

1. "Many of the evaluations were of the type known as connoisseurial assessments because they relied nearly exclusively on the expertise of the reviewer and often lacked an articulation of a general method for conducting the analysis."
2. "Evaluators reviewed a specific curriculum for accuracy and for logical sequencing of topics relative to expert knowledge."
3. " Some evaluators explicitly the curriculum being analyzed to international curricula in countries in which students showed high performance on international tests."
4. These connoisseurial assessments "involve judgment and values and hence depend on one's assessment of of the qualifications and reputation of the reviewer."

NRC Report: Comparison Studies

1. ****Comparison Studies:** "at least minimally methodologically adequate" criteria:

   a. Include Quantifiable measures of outcomes

   b. Adequate information about comparability of samples

   Plus, at least one of the following:

   c. A report of implementation fidelity or PD activity

   d. Results disaggregated by content strands or performance by student subgroups

   e. Multiple outcome measures or precise theoretical analysis of a measured construct (e.g., number sense, proof, etc.)

NRC Report: Comparative Study Considerations

"Discussion focused on seven critical issues faced by evaluators in the conduct of comparative studies":

1. Selecting study type: experimental or quasi-experimental
2. Establishing of comparability across groups
3. Selecting of a comparative unit of analysis
4. Measuring/Documenting implementation fidelity
5. Conducting and impact assessment or outcome measure
6. Selecting/Conducting statistical tests
7. Determining limitations to generalizability relative to sample selected

NRC Report: Case Studies "Ethnographic Evaluation"

1. "documented how program theories and components played out in a particular case or set of cases"
2. "used triangulation of evidence from multiple sources, including direct observations, interviews, documents, archival files, and actual artifacts"; may include pre- and post-outcome measures, open-ended testing, etc.
3. recognized "that a slow, and sometimes agonizing process of analyzing cases provides the detailed structure of argument often necessary to understand and evaluate complex phenomena"
4. Can focus on decision making processes, documenting implementation fidelity, patterns of instruction, student needs

NRC Report Findings: The Quality of Evaluations

1. “the corpus of evaluation studies as a whole across the 19 programs studied does not permit one to determine the effectiveness of individual programs with a high degree of certainty due to the restricted number of studies for any particular curriculum, limitations in the array of methods used, and the uneven quality of the studies”

Therefore:

2. “No second phase of this evaluation review should be conducted to determine the effectiveness of any particular program or set of curricular programs dependent on the current database”

NRC Report Recommendations: Content Analysis

1. "Content analyses should be recognized as a form of connoisseurial assessments, and thus should be conducted by a variety of scholars, including mathematical scientists, mathematics educators, and mathematics teachers and well-qualified individuals, who should identify their qualifications, values concerning mathematical priorities, and potential sources of bias regarding their execution of content analysis"
2. "A content analysis should clearly indicate the extent to which it addresses the following three dimensions:
   a. Clarity, comprehensiveness, accuracy, depth of mathematical inquiry and reasoning, organization, and balance (disciplinary perspectives).
   b. Engagement, timeliness and support for diversity, and assessment (learner-oriented perspectives).
   c. Pedagogy, resource, and professional development (teacher- and resource-oriented perspectives)."

NRC Report Recommendations: Comparative Studies

1. “more rigorous designs” should be employed
2. “more precise measures of content-strand outcomes, especially in relation to curricular validity measures” should be used
3. “careful sampling of representative groups and examination of outcomes by student subgroups” should be included
4. Selection of correct unit of analysis
5. Better data on implementation fidelity
6. Outcome data should include a variety of measures of high quality (e.g., by question type, type of test, relation of test to practice, etc.)
7. Careful consideration of appropriate statistical tests
8. Statements of limitations and effect sizes should be reported
9. Control groups should be clearly identified

NRC Report Recommendations: Case Studies

1. “Case studies should stipulate clearly what they are cases of, how claims are produced and backed by evidence, and what events are related or left out and why, and should identify explicit underlying mechanisms to explain a rich variety of research evidence.”

2. “It is worth noting that case studies often reveal aspects of program components, implementation components, and interactions among these two that behave differently than intended by program designers, and therefore provide essential insights into program effectiveness. The committee emphasizes that a case study should be conducted as rigorously as any other form of study.”

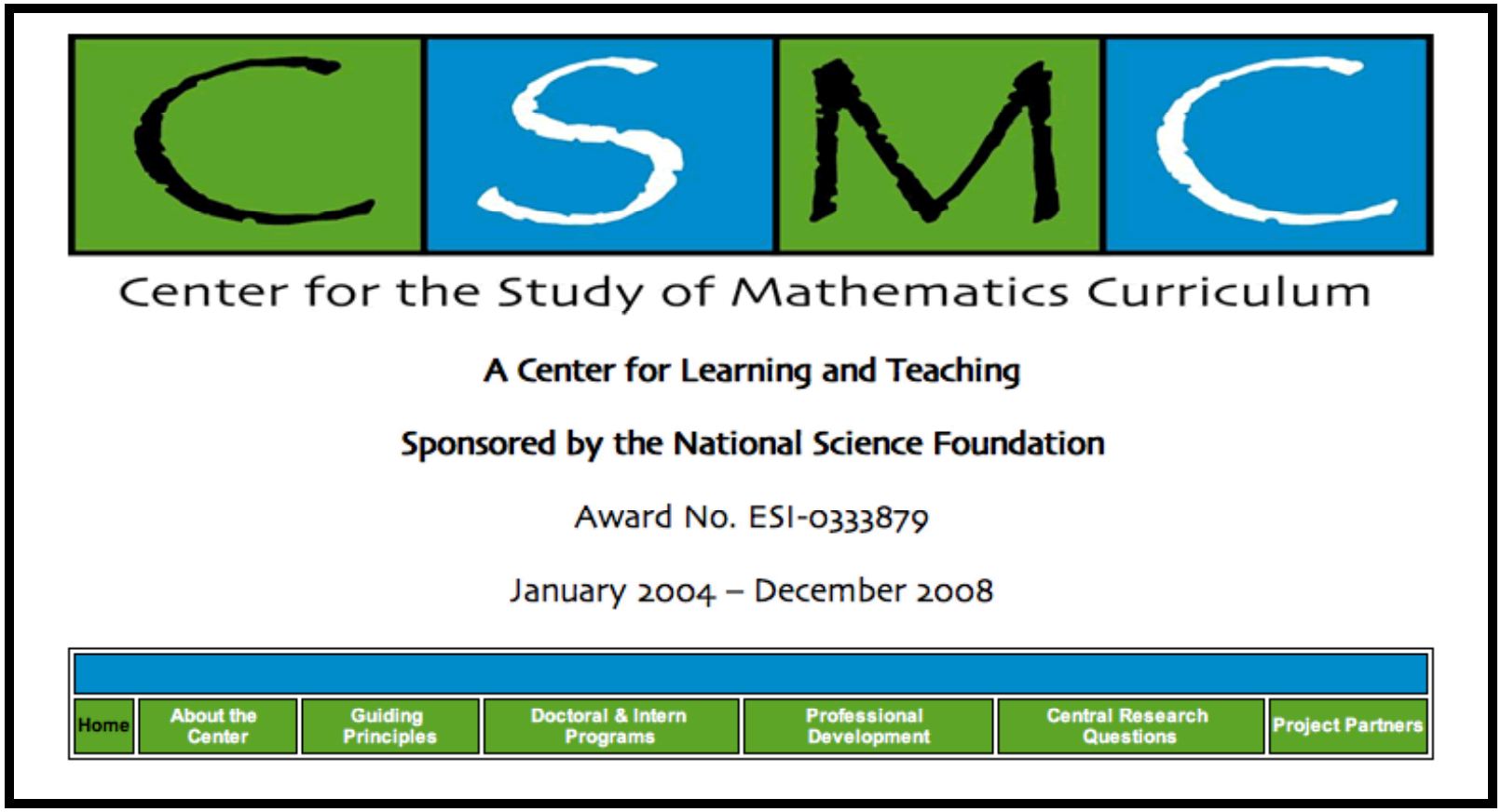
NRC Report Recommendations:

For Others

1. Federal Agencies should “improve the nation’s capacity in mathematics curriculum evaluation. Individuals and teams charged with curriculum evaluation should show evidence of understanding the interdisciplinary nature of the task.”
2. Federal Agencies should “include more explicit expectations for evaluation of curricular initiatives” in their calls for proposals.
3. A Federal Agency should “develop a program for district- and school-level data collection and maintenance” related to curriculum implementation.
4. Publishers should 1) differentiate between market research and scientifically valid evaluation studies, and 2) make such evaluation data available to potential clients.
5. The DoE and state DoEs should provide local and district decision makers with training on how to conduct and interpret valid studies of curricular effectiveness.

NRC Report : Final Thoughts

"These 19 curricular projects essentially have been experiments. We owe them a careful reading on their effectiveness. Demands for evaluation may be cast as a sign of failure, but we would rather stress that this examination is a sign of success of these programs to engage the country in scholarly debate on the question of curricular effectiveness and the essential underlying question, What is most important for our youth to learn in their studies in mathematics? To summarily blame national decline on a set of curricula whose use has a limited market share lacks credibility. At the same time to find out if a major investment in an approach is successful and worthwhile is a prime example of responsible policy. In experimentation, success and worthiness are two different measure of experimental value. An experiment can fail and yet be worthy. The experiments that probably should ***not*** be run are those in which it is either impossible to determine if the experiment has failed or it is ensured from the start, by design, that the experiment will succeed." (p.188)



http://mathcurriculumcenter.org