The Relationships Among Middle-School Inquiry-Based Science Student Achievement Measures: The Results of Structural Equation Modeling Analyses

Carlos Ayala, Sonoma State University

Paul R. Brandon, University of Hawaii at Manoa

Presentation to the annual meeting of the
American Evaluation Association, November 2006

Purpose of this session

- To report the results of using structural equation modeling (SEM) to examine student level data collected to measure the effects of an inquiry based science program
  - Nature of Science (NOS)
  - Student Self-efficacy in conducting science investigations (SE)
  - Student achievement (MC)

Structure of this session

- Context of these measures
- Development of these measures
- Study method
- Study results
- Implications

Larger Study

- Funded by NSF Grant No. REC-0228158
- Measures for use in a study comparing two versions of PD for an interdisciplinary middle-school inquiry-based science program were prepared.
  - One version is shorter than the other and incorporates long-term technology-based, support.
  - The curriculum used was the Foundational Approaches in Science Teaching (FAST).

FAST 1-3

- Three courses in FAST:
  - *The Local Environment*
  - *Matter and Energy in the Biosphere*
  - *Change Over Time*
- Our focus was physical science in FAST 1
- FAST has been shown to positively affect achievement and other student outcomes

Inquiry-Based Science

- In inquiry-based science, students learn science by doing scientific investigations. They learn how to
  - develop questions.
  - collect and explain evidence.
  - connect explanations to existing knowledge.
  - communicate and justify explanations.

FAST Activities

- Students in small groups model science as it is practiced.
- Students spend ¾ of their time in field studies.
- Teachers are “research directors;” they probe student’s thinking through the effective use of questioning (“Socratic inquiry”).

Instruments Developed in the Larger Study

- Questionnaire for measuring program implementation and program context
- Log for measuring program implementation
- Observation protocol for measuring implementation, with a focus on teacher questioning strategies
- A suite of student measures both achievement and attitudinal

PD Differences in Student Measures

Student's

- Content knowledge
- Science inquiry skills
- Views of the nature of science
- Efficacy toward science learning
- Value of science
- Science anxiety

Final Assessment Suites

- Pre-test
  - 30-item multiple-choice/ short answer content knowledge test ($\alpha = .77$)
  - 10-item IRT model ($\alpha = .77$)
  - Attitudinal Survey
- Post-test
  - 30-item multiple-choice/ short answer content knowledge test
  - Attitudinal survey
  - Rocky River performance assessment

Student measures for SEM study

- Student's content knowledge
- Student's views of the nature of science
- Student's efficacy toward science learning

Content Validity

- Reviewed curriculum and created content matrices
- Curriculum developers parsed content down
- Linked matrices to materials
- Created test linking items to matrices
- Surveys reviewed by experts
- Piloted assessments and surveys
- Talk alouds

Content Knowledge

- Measured with multiple choice test
- Same items here

Nature of Science

- As students become more engaged in the FAST curriculum, students will understand that
  - Anyone can be a scientist
  - Science knowledge is useful
  - Science knowledge builds over time
  - Science is creative
- Sample Items
  1. Scientists always get the same results.
  2. All good scientists work in the same way.
  - *Higher score more positivistic*

Lederman, Abd-El-Khalick, Bell and Schwartz (2002) *Views of Nature of Science Questionnaire; Toward Valid and Meaningful Assessments of Learner' Conceptions of the Nature of Science.*

Self Efficacy

- As students become more engaged in the FAST curriculum, the greater control they will feel towards science, science investigations and science knowledge.
- Sample Items
  1. I can make accurate measurements during a science investigation.
  2. I can make appropriate predictions about what will happen during a science investigation.

Britner, S. and Pajares F., (2001) *Self-Efficacy Beliefs, Motivation, Race and Gender in Middle School Science.*

Method

- 356 with complete achievement and aptitude data sets
- Teachers administered pre and post assessments when they were done with curriculum or at the end of the school year.
- Constructed response items scored by trained science teachers

Teacher Characteristics

Table 1
Background Characteristics of Teachers

Teacher characteristics	All surveyed teachers			Surveyed teachers who administered the test		
	N	Mean	St. dev	*N*	Mean	St. dev
N year k—12 teacher	79	12.1	8.5	10	8.0	5.3
N year k—12 science teacher	79	9.7	8.0	10	4.5	4.5
N undergrad science courses	72	13.6	17.7	9	10.1	9.6
N graduate science courses	74	5.0	8.2	9	2.2	5.0
Year born	79	1963.9	10.9	10	1968.2	7.5

Teacher Characteristics

Table 2
Additional Background Characteristics of Teachers

Teacher characteristics	All surveyed teachers		Surveyed teachers who administered the test	
	N	%	*N*	%
Male	31	39%	0	0%
Female	48	61%	10	100%
Public school	54	68%	5	50%
Private school	25	31%	5	50%
Science teacher licenses				
Yes	53	67%	4	40%
No	26	33%	0	60%
Highest degree:				
Bachelors	46	59%	7	70%
Master’s or doctorate	33	41%	3	30%

Structural Equation Modeling

- SEM is used to examine relationships among a set of variables
  - Remove measurement error
  - Identify paths
  - Often used to show causation
- Our purpose is examine validity and not show causation

Two Steps in SEM

- Develop measurement model
  - Confirmatory factory analysis
  - Items or subscales form latent variables
- Develop structural model
  - Path analysis

Our Measurement Model

- Used items and subscales to form latent or indicator variables
- Analyzed student variables using IRT (Multilog) to find items and subscales that discriminated students well
  - Motivation, anxiety, epistemic beliefs subscales did not.
  - NOS, SE, MC items discriminated best.

Our Measurement Model

- Used the highest IRT discriminating variables
  - Three NOS items
  - Four Self-Efficacy items
  - 10 MC items

SEM Subscales	Student Scores			
	N	Mean	St. dev	Scale
Nature of Science	356	2.9	2.1	1-5 high
Self Efficacy	356	74.4	20.5	1-100 high
Multiple Choice Pre	356	4.6	2.1	0-10
Multiple Choice Post	356	5.0	2.5	0-10

Validity

- SAS PROC CALIS
  - Nonsignificant chi square (p=.37)
  - Three fit indexes (all statistically significant)
    - Goodness of fit = .975
    - Non-normed index = .997
    - Comparative fit = .998
    - Composite reliability and variance extracted statistics all suggested an acceptable model

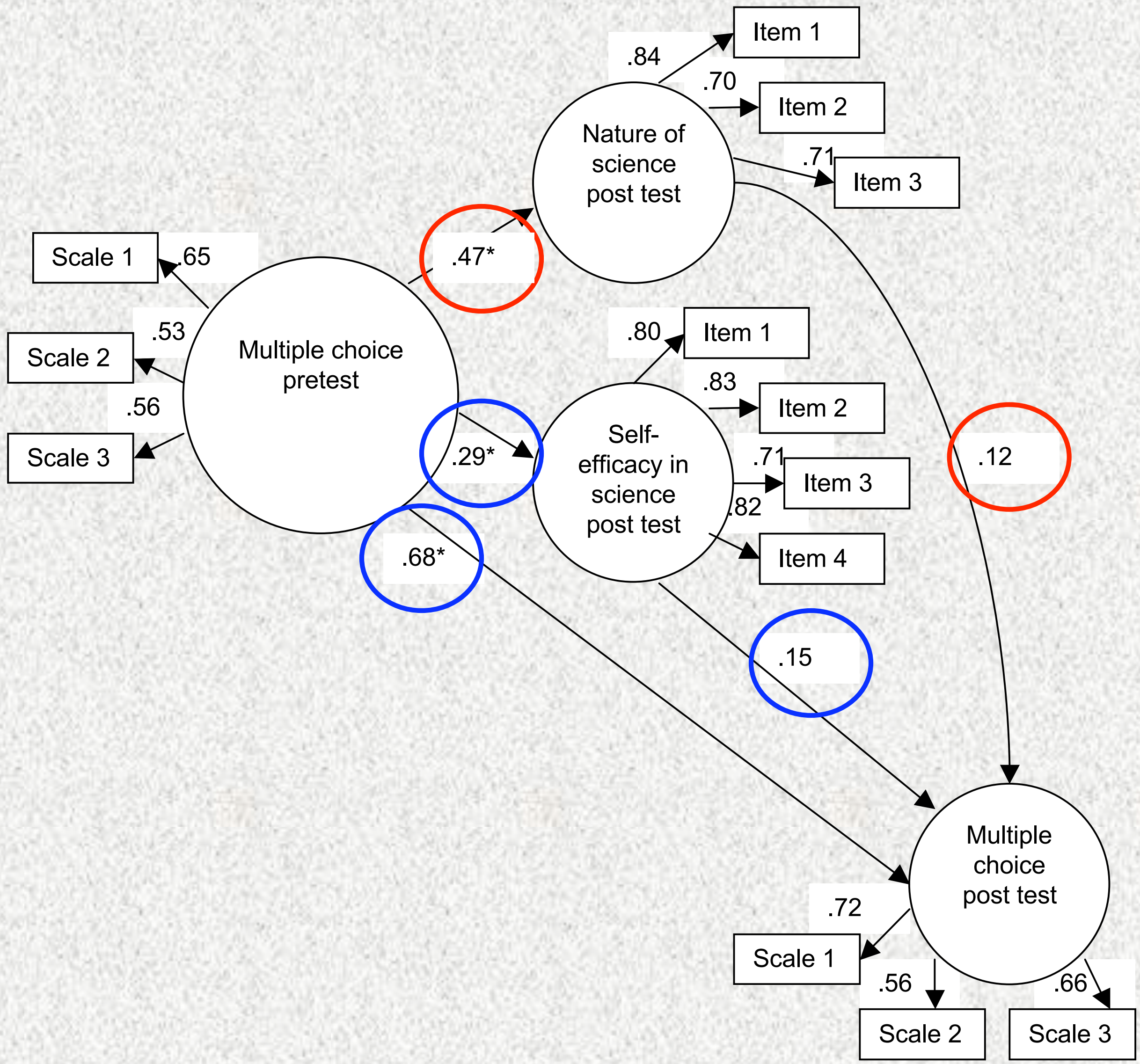

Figure 1. Structural equation model of three inquiry science student outcome measures (* indicates significant @ .05 level).

Conclusions

- Could not form a well fitting model that included NOS and SE pretests
  - Students are under informed about the nature of science and their self-efficacy towards science investigations prior to the inquiry unit.
  - Students developed a greater understanding about these measures during the unit.

Conclusions

- The multiple choice pretest showed weakness as a pretest measure.
  - If the test had been more reliable, relationships with both NOS and SE might have been stronger.
  - This probably reflects the difficulty of the measure; the test was hard for students.

Conclusions

- Prior achievement showed moderate relationships with both NOS and SE.
  - Content knowledge predicts opinions about the nature of science and beliefs about one's self-efficacy somewhat.
    - This helps lend credence to the appropriateness of the attitudinal scales as measures of inquiry science
  - However, the higher the scores on the pretest, the higher the NOS posttest scores.
    - After participating in inquiry science classes, the higher achieving students believe that science is less positivistic.
    - Science is more attainable to them.

Conclusions

- Relationships between the NOS and SE constructs and posttest scores not strong.
  - Supports that NOS and SE are independent outcome measures for an inquiry base science class.
  - Achievement alone is insufficient as a measure of outcomes of inquiry science.
  - Once prior achievement is accounted for, NOS and SE are mostly unrelated to current achievement.

Implications

- Researchers should be sure to include attitudinal measures of outcomes in their studies.
- Educators should expect that students' understanding of the overall nature of science and their beliefs about their efficacy as scientists-in-training will be minimal before they begin conducting their own scientific investigations in inquiry science classes.

Cautions About Generalizability

- This study sampled low-achieving students.
  - This might have affected the results on the instruments discussed here (e.g. the pattern of relationships and the low pretest reliability).
  - It might also account for the lack of discrimination on other scales (e.g., motivation).
- The results of the study might not be generalizable beyond the FAST program.