Joseph A. Taylor, Pamela Van Scotter, Doug Coulson

Bridging Research on Learning and Student Achievement: The Role of Instructional Materials

The authors describe results from two field-test studies conducted by BSCS, both of which indicate a strong relationship between fidelity of curriculum implementation and student learning gains.

For decades the National Science Foundation has been funding the development of instructional materials whose design is based upon the recommendations of educational research. These recommendations include the idea that learning be sequenced and organized using an experiential learning cycle (see Atkin & Karplus, 1962; Piaget, 1975) or an instructional model such as the Biological Science Curriculum Study (BSCS) 5E Instructional Model (see Bybee, 1997). More recent research studies such as those synthesized in *How People Learn* (NRC, 2001) suggest that instruction 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. Work by Pellegrino, Chudowsky, and Glaser (2001) make complementary recommendations such as the following:

- Instruction should be organized around meaningful problems and goals.
- Instruction must provide scaffolds for solving meaningful problems and supporting learning for understanding.
- Instruction must provide opportunities for practice with feedback, revision, and reflection.
- The social arrangements of instruction must promote collaboration and distributed expertise as well as independent learning.

Over time, the work of curriculum developers in response to these recommendations has resulted in an extensive portfolio of research-based instructional materials that span the sciences disciplines. However, few researchers have systematically explored how these materials influence student learning and the role that implementation fidelity plays in the materials' ultimate impact. The purpose of this article is to address this issue by exploring data collected in two BSCS research studies. In the following sections we briefly describe these studies, their respective findings, and make recommendations for teacher professional development.

BSCS Field-Test Research: A Brief Overview

The BSCS, as a routine part of its curriculum field-testing process, studies the impact of the materials on student learning. In 1995, BSCS conducted case studies of four teachers who were field-testing a new high school science program. The learning experiences in this curriculum program are structured and sequenced using the BSCS 5E Instructional Model (Engage, Explore, Explain, Elaborate, and Evaluate) which is based upon the research-based learning cycles of Atkin and Karplus (1962) and Piaget (1975).

The four case studies uncovered distinct differences in the pre/post learning gains of students whose teachers implemented the program as designed as opposed to those of students whose teachers implemented the program with considerably less fidelity. Student learning was measured using a 20-item subset of questions from the National Science Teachers Association (NSTA)/National Association of Biology Teachers (NABT) biology exam, administered at the beginning and end of the school year. Fidelity was measured through classroom observations conducted by BSCS curriculum development staff. These findings are illustrated in Figure 1 and Table 1.

The case studies suggested a relationship between fidelity and student learning, and BSCS sought to explore

Figure 1: Pre/Post Results for NABT/NSTA Biology Exam

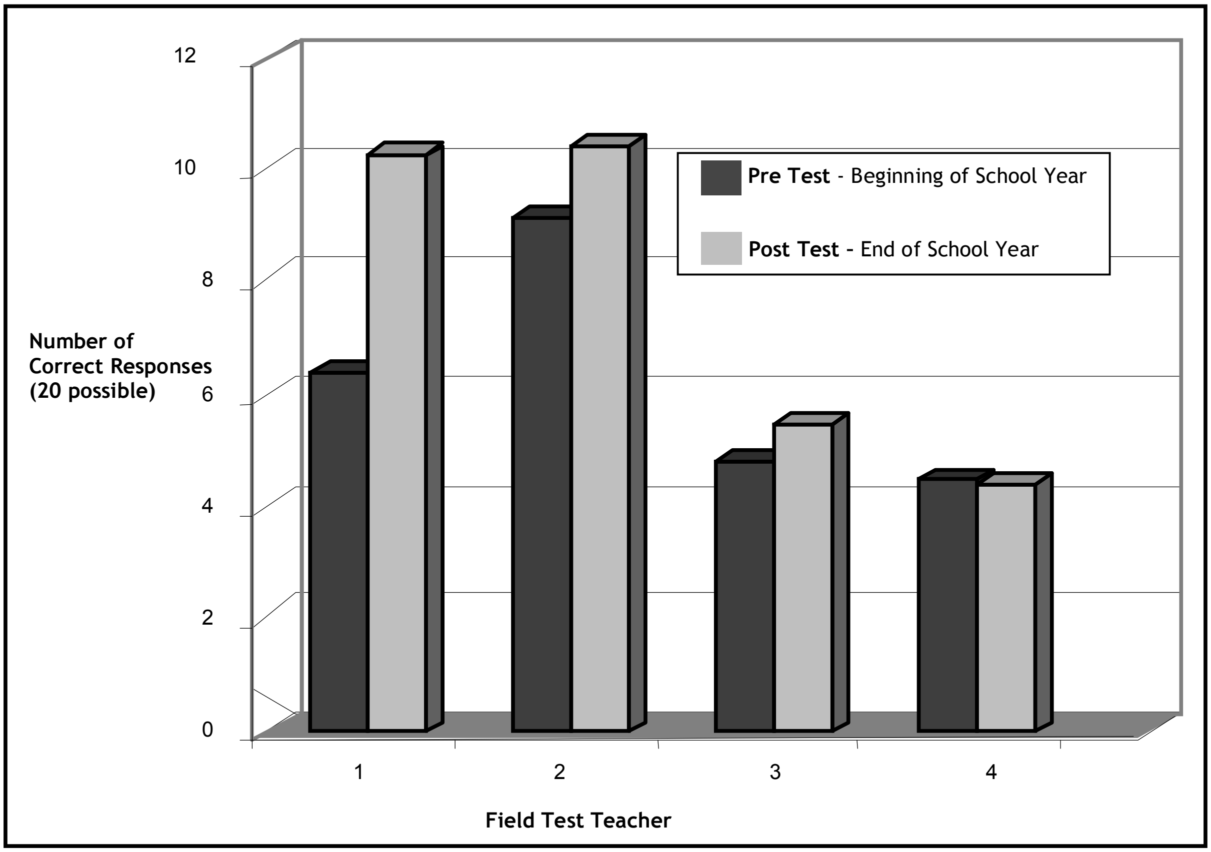


Table 1: Student Learning Gains by Teacher

Teacher	Pre-test Avg.	Post-test Avg.	Avg. Gain
1	6.4	10.3	3.9
2	9.2	10.4	1.2
3	4.8	5.5	0.7
4	4.5	4.4	0

Teacher 1 - Field tested instructional materials for two years with a high level of fidelity
Teacher 2 - Field tested instructional materials for two years with a medium level of fidelity
Teacher 3 - Field tested instructional materials for one year with a medium level of fidelity
Teacher 4 - Field tested instructional materials for one year with a low level of fidelity

this relationship with larger numbers of teachers and students in future field-test studies. This opportunity came in 2002 as BSCS began field-testing a newly-developed high school science program also structured using the BSCS 5E Instructional Model.

The 2002 Study: An Overview

One of the questions that the 2002 study sought to answer was whether the magnitude of student learning gains were in any way related to how closely teachers adhered to the intended use of the instructional materials. Specifically, the study asked: *how do the learning gains of students whose teachers are successful in implementing the program as designed (with fidelity) compare with the learning gains of students whose teachers are less successful in implementing the program as designed?*

In this study, the learning gains of 326 ninth-grade students (across 15 teachers) were determined by administering an identical chapter test before and after instruction. Implementation fidelity was measured by a research team composed of external evaluation and curriculum development staff using an observation protocol adapted from the *Horizon, Inc. Classroom Observational Protocol* (see HRI, 2000). The implementation section of this protocol included 11 individual rating scales that addressed the use of specific teaching strategies consistent with the BSCS 5E Instructional Model. These rating scales were in Likert format with scores ranging from 1 to 5; one meaning "not at all," 3 meaning "to some extent," and 5 meaning "to a great extent." For each scale, the team recorded specific notes as evidence and justification for the ratings. Specifically, the individual rating scales were intended to quantify the extent to which teachers encouraged students to: engage in metacognitive activity; communicate their understanding of concepts, and apply their understanding to new situations. These rating scales were also used to quantify the extent to which teachers used embedded assessment and student self-assessment, as well as provided students with appropriate feedback to help them construct their understanding of targeted concepts.

When student learning gains were examined in light of implementation fidelity, a strong relationship emerged.

Using these 11 ratings of teachers' use of 5E-based strategies and learning sequences, BSCS then holistically classified each teacher's fidelity of use as either "low," "medium," or "high" (see Table 2). These holistic ratings alone were used to divide teachers into low, medium, or high levels of fidelity because trichotomizing teachers on the holistic rating resulted in a clearer division than averaging the other 11 individual rating scales. The averaging process tended to cancel out the variation, making it problematic to create fidelity categories.

Findings from the 2002 Study

When student learning gains were examined in light of implementation fidelity, a strong relationship emerged. An analysis of covariance (ANCOVA) was conducted across the three fidelity groups (i.e., low, medium, and high) on post test using the pre test as a covariate. As expected, the covariate (pre-test) was significantly related to the dependent variable (post-test). There was a main effect associated with the implementation fidelity grouping ($F = 7.51$; $df = 2, 322$; $p < .01$). This analysis, summarized in Table 3, suggests that we reject the null hypothesis that the post-test scores are equal across fidelity groups.

Post hoc paired comparisons using Fisher's Least Significant Difference (LSD) method were conducted in order to identify which of the three implementation fidelity groups were statistically different. These three comparisons showed that the post-test scores (when adjusted for pre-tests) of the low fidelity group were on average statistically lower than both the medium and high level fidelity groups and that the adjusted post-test scores for the medium and high fidelity groups were not statistically different. Table 4 and Figure 2 below summarize these findings.

Discussion

The data presented in this study suggest that the students whose teachers used the instructional materials with medium or high levels of fidelity scored statistically higher on the post-test achievement measure than students of low fidelity teachers. Using the operational definitions of medium

Table 2: Fidelity Levels as a Continuum of 5E-Based Teaching

	Holistic Rating of Fidelity Level		
	Low	**Medium**	**High**
Observed use of strategies and learning sequences that are consistent with the 5Es	**None**	**Basic**	**Extensive**
Number of Teachers Rated at each Fidelity Level	**4**	**7**	**4**

Table 3: ANCOVA Output

Source	DF	Mean Square	F-Ratio	Probability
Pre Test	1	2.8368	48.5	.0000
Implementation Fidelity Group	2	.43938	7.51	.0006
Error	322	.05849		
Total Adjusted	325			

Table 4: Adjusted Means Table

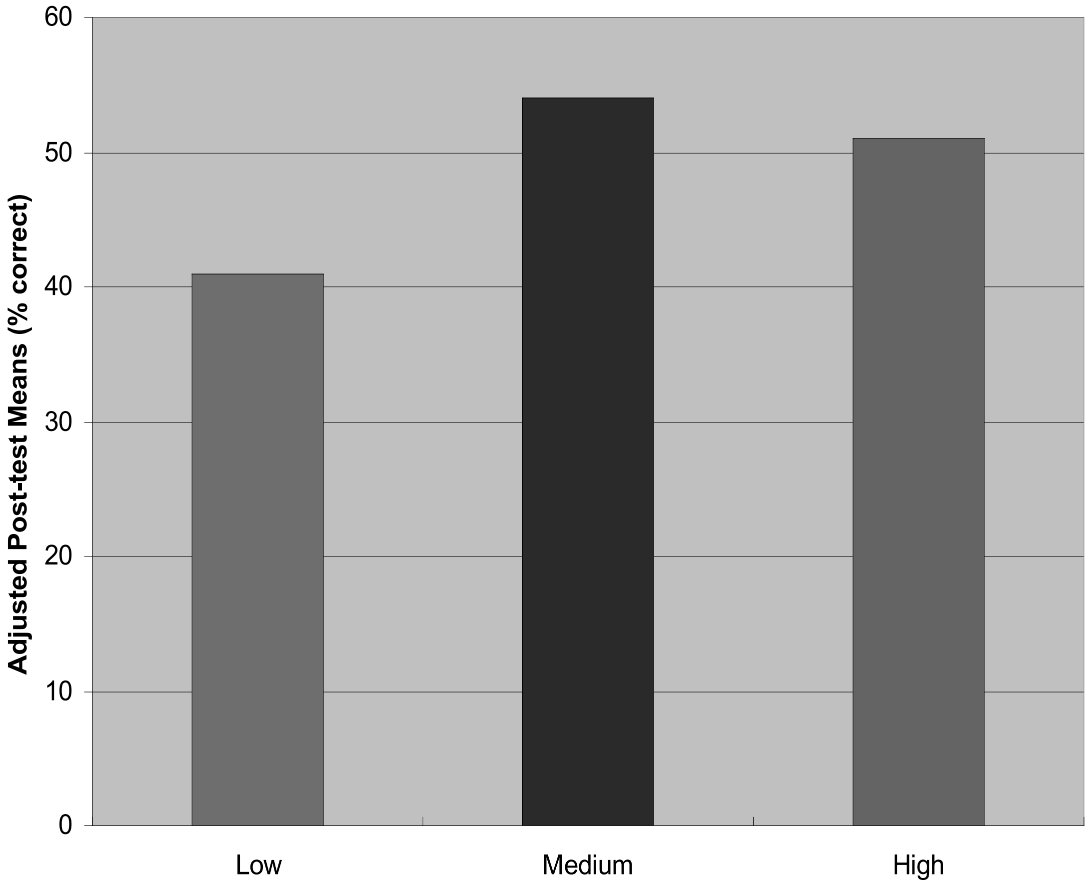
Teacher Group	#of Students	Adjusted Mean Post-test Score (% correct)	Standard Deviation	Fisher LSD (adjusted mean is significantly different from)
Low Level Fidelity	70	41	2.9	Medium and High Fidelity Group
Medium Level Fidelity	168	54	2.6	Low Fidelity Group
High Level Fidelity	88	51	2.0	Low Fidelity Group

and high levels of fidelity, we can then state that there is a statistical link between superior student achievement and ***basic*** or ***extensive*** use of strategies and learning sequences consistent with the 5Es. The observation that marked differences in achievement begin at even basic use of the 5Es makes a powerful statement about the effectiveness of the instructional model. Complimentary findings can be found in comparative studies--conducted across multiple science disciplines and grade levels--suggesting that research-based instructional models likely promote greater gains in student achievement than more didactic teaching approaches (e.g., Ates, 2005; Ebrahim, 2004; Lord, 1997).

The lower student achievement observed in low fidelity classrooms is not surprising. Too often in science education we see well-designed instructional materials, even those designed to organize everyday instruction, collecting dust on shelves or being pulled off shelves and used haphazardly as a mere resource or supplement. This often results in a patchwork approach to curriculum implementation, which leads to lack of coherence in the learning sequence for students (Rutherford, 2000; Taylor, et al., 2005). This is not to say that teacher-designed materials tend to be of poor quality. However, most classroom teachers' ability to field-test their materials with large numbers of diverse students or to have them reviewed by content experts is clearly limited. A haphazard approach to curriculum implementation does not take advantage of the thoughtful work of the science education research and curriculum development communities. The prevalent notion of teacher as curriculum developer or curriculum "hunter and gatherer" must be challenged at both the K-12 and higher education levels. The data in this study directly confront this common notion and suggest an approach to curriculum implementation where fidelity is valued.

It is interesting to speculate about why extensive use of the 5Es did not yield student achievement that was significantly different from basic use of the instructional model. It is possible that the instructional materials, which make the 5Es explicit for both teachers and students, are ***fidelity-resilient*** (within limits). That is, even when a teacher implements the material with only medium fidelity, the inherent design of the instructional materials likely contributes to enhanced un-

Figure 2: Adjusted Post-test Means by Level of Fidelity



derstanding, and thus achievement, for students.

Other hypotheses around this finding center on how the data were analyzed. For example, some of the teachers who were rated holistically at medium levels of fidelity were rated so because their use of 5E-based strategies and learning sequences was not consistent across a continuum from non-use to extensive use. It is possible that a large percentage of these teachers made extensive use of the very strategies that were most effective for students. To address this issue in future research we would like to conduct an analysis of covariance across selected individual fidelity rating scores from the observation protocol, again using pre-test scores as covariates. This analysis would help us make more direct connections between student achievement and the use of specific instructional strategies.

We also hypothesize that in some cases, teachers made departures from the design of the instructional materials that optimized student learning and were in the ***spirit*** of the instructional model but were noted by observers as showing less fidelity—resulting in lower fidelity ratings. This conclusion would suggest that it is indeed possible for teachers with a nuanced understanding of the 5Es to enhance the impact of 5E-based instructional materials with well-informed pedagogical decisions.

The question then becomes: how do we help teachers develop sophisticated understandings of instructional materials and the instructional models that drive their design? Suggested approaches to professional development are articulated in other studies that establish a link between fidelity of implementation and student learning (e.g., Cohen & Hill, 2002; Darling-Hammond, 1997). In many of these studies, high levels of fidelity are attributed in part to on-going, comprehensive professional development that is focused almost solely on helping teachers understand the design of the instructional materials. The suggestion is quite simple: if you want higher fidelity implementation, spend some professional development time and resources on helping teachers understand the instructional materials. Further, since the design of research-based instructional materials can look quite different from that of traditional instructional materials and embody instructional models of which many teachers are unfamiliar, the importance of professional development is emphasized.

A haphazard approach to curriculum implementation does not take advantage of the thoughtful work of the science education research and curriculum development communities.

In the 2002 study, the translation of this simple suggestion was to provide teachers with a professional development program that included workshops to help them understand the goals and purpose of the instructional model (5Es) that organizes and structures the learning for students. The specifics of this approach were influenced by scholars such as Loucks-Horsley, et al. (2003) who suggested that professional development mirror the instructional methods to be used with students. Therefore, teachers in this study were engaged as learners of science in investigations where the facilitator modeled exemplary use of the 5E model. To help teachers apply their science learning experience to learning about science teaching, the investigations were discussed afterward in terms of what the learner and the facilitator were doing in each stage of the 5Es.

The professional development program also engaged teachers in other foundational work to help them understand the instructional materials. For example, BSCS conducted workshops to help teachers better understand inquiry and effective ways to create a climate of inquiry in their classrooms, as well as sessions on *Understanding by Design* (Wiggins and McTighe, 2005), which provided key principles for the design of the program. The professional development program also included focused sessions on the science content since it was observed in the past that low levels of fidelity can often result from unfamiliarity or discomfort with the science content.

In summary, the data in this study suggest that research-based instructional models are most effective when they are taught with at least a basic level of fidelity. We also hypothesize that well informed modification of research-based instructional materials could optimize their impact on student achievement. However, regardless of whether the goal is to have teachers use research-based instructional materials with optimal fidelity or to modify the materials in appropriate ways, it is critical that professional development focus on helping teachers develop a vision of implementation that is consistent with the designer's intent.

References

Ates, S. (2005). The effectiveness of the learning-cycle method on teaching DC circuits to prospective female and male science teachers. *Research in Science and Technological Education* 23(2): 213-227.

Atkin, J. M., & Karplus, R (1962). Discovery or Invention, *The Science Teacher, 29 (2),* 121-143.

Bybee, R. W. (1997). *Achieving Scientific Literacy: From purposes to practices*. Portsmouth, NH: Heinemann.

Cohen, D. K., & Hill, H. C. (2002). *Learning policy: When state education reform works*. New Haven, CT: Yale University Press.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching*. New York: National Commission on Teaching and America's Future.

Ebrahim, A. (2004). The effects of traditional learning and a learning cycle inquiry learning strategy on students' science achievement and attitudes toward elementary science (Kuwait). *Dissertation Abstracts International* 65(4): 1232.

Horizon Research, Inc. (2000). *Horizon Research-2000-2001 Local systemic change classroom observation protocol*. (Found at www.horizon-research.com).

Lord, T. R. (1997). A comparison between traditional and constructivist teaching in college biology. *Innovative Higher Education* 21(3): 1127-1147.

Loucks-Horsley, S., Love, N., Stiles, K., Mundry, S., & Hewson, P. (2003). *Designing professional development for teachers of science and mathematics*, 2nd ed. Thousand Oaks, CA: Corwin Press.

National Research Council. (2001). *How People Learn*. Washington, DC: National Academies Press.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of assessment*. Washington, DC: National AcademyPress.

Piaget, J. (1975). *The Development of Thought*. New York: Viking Press.

Rutherford, F. J. (2000). Coherence in high school science. In *Making sense of integrated science: A guide for high schools*. Colorado Springs, CO: BSCS.

Taylor, J., Powell, J., Bess, K., & Lamb, T. (2005). Examining the professional growth of out-of-field physics teachers: Findings from a pilot study. *Journal of Physics Teacher Education Online, 2(4),* 16-22.

Wiggins, G., & McTighe, J. (2005). Understanding by design, 2nd ed. Alexandria, VA: Association for Supervision and Curriculum Development.

Joseph A. Taylor is director, BSCS Center for Research and Evaluation, 5415 Mark Dabling Blvd., Colorado Springs, CO 80918. Correspondence concerning this article may be sent to jtaylor@bscs.org

Pamela Van Scotter is director of the Center for Curriculum Development at BSCS in Colorado Springs, Colorado.

Doug Coulson is senior partner at PS International, an evaluation-consulting group in Annapolis, Maryland, and adjunct faculty in education statistics at the University of Maryland, College Park, Maryland.