Research Summary: Impact of the Understanding Science Professional Development Model on Science Achievement of English Language Learner Students

Regional need and study purpose

Teachers with strong knowledge in science content and in teaching science are more likely to ask their students higher level questions, encourage them to discuss the content and think about its applications, and foresee the difficulties they may encounter (Carlsen 1991, 1993; Druva and Anderson 1983; Hashweh 1987). In contrast, teachers with little science knowledge or experience teaching science content tend to depend more on direct instruction (Heller et al. 2004), to more often use inaccurate examples in their instruction, and to restrict classroom discourse to factual and simple recall questions (Carlsen 1991, 1993). It thus seems logical that teachers with greater knowledge in science content and in teaching science might be more effective at improving student science achievement than those without such knowledge.

This study is designed to examine whether one teacher professional development program in middle school physical science, WestEd's Understanding Science Program, improves teachers' content knowledge in the topic "force and motion," their instructional strategies and confidence in teaching it, and in turn grade 8 students' content knowledge in force and motion and performance on California's standardized science test. Of particular interest in the study is the impact of the program on the performance of English language learner students—24.7 percent of the K–12 student population in California (California Department of Education 2009).

The program emphasizes inquiry-based instruction practices that depend less on English proficiency, textbook knowledge, and direct instruction. Its model emphasizes activities that provide opportunities for developing English and academic language proficiency. For example, collaborative, small-group work allows students to develop English proficiency through authentic communication about science knowledge (Lee 2002; Lee and Fradd 2001; Rosebery, Warren, and Conant 1992). Because the courses are designed to help teachers support discussion in science—to make sense of the content and help students develop academic language proficiency—the program might especially benefit the science performance of English language learner students.

Understanding Science has been evaluated using only quasi-experimental designs (Heller, Daehler, and Shinohara 2003; Heller and Kaskowitz 2004; Heller 2006). In earlier research, scores on tests of science content knowledge increased for elementary and middle school teachers who participated in the courses. Among elementary school students, students whose teachers participated in Understanding Science courses had statistically significant differences in content knowledge when compared with their control counterparts—and such differences were more pronounced for English language learner students. But because this previous nonexperimental research could not control for confounding factors when estimating the program's effects on teacher and student knowledge, it is difficult to determine whether the program alone caused these changes. This study tests similar outcomes, but using a rigorous experimental design.

The study examines the following research questions:

- What is the impact of the Understanding Science professional development sessions on teachers' knowledge of force and motion and their confidence in teaching it.
- And what is its impact on teachers' strategies for developing students' understanding of science, revealing student ideas about science, and supporting English language learner students' science learning?
- What is the impact of the Understanding Science professional development sessions on students'—especially English language learner students'—content knowledge in force and motion and their performance on California's standardized science test?

Intervention description

Incorporating physical science content, student work and thinking, and classroom instruction, the Understanding Science professional development sessions develop physical science concepts through structured investigations and narratives based on actual classroom experience with those concepts.

Composed of 8 three-hour sessions, the professional development consists of a course on force and motion for teachers held over five days in August 2009. The eight sessions are sequenced so that the science topics (speed, velocity, acceleration, balanced and unbalanced forces, for instance) build on each other—uncovering the corresponding language issues and strategies for supporting student learning and language development incrementally throughout the sessions.

Each three-hour session includes two components—science investigations and case discussions (analysis of detailed instructional scenarios). Written by classroom teachers, case discussion materials contain student work, student-teacher dialogue, context information, and descriptions of teacher thinking and behavior. Science investigations engage teachers in the particular case's core dilemma (for example, students incorrectly think that if something is moving, there must be a force acting on it). Through common standards-based curricula, teachers' experiences in these hands-on exercises parallel those of the students in the cases. Following the science investigation, teachers examine student thinking and analyze instruction from the case discussion. Such discussions are designed to lead teachers to explore alternative perspectives and solutions—possibly spurring them to rethink their instruction. Language and literacy course components teach how to more effectively support discussions of science, both to make sense of the science and to help students, particularly English language learner students, develop academic language proficiency.

This approach presumes that situating professional development in an environment of collaborative inquiry—conceptually exploring science along with focusing on student thinking and teacher instruction—increases teachers' knowledge of science content and instruction while helping them develop targeted strategies for eliciting student ideas and strengthening their science language abilities. The study posits that these outcomes will result in changes in classroom practices, such as increased accuracy of science representations and explanations, a focus on helping students develop conceptual understanding and on assessing student ideas during instruction, and a greater opportunity for students to develop the language of science. Ultimately, such changes may improve student achievement, help students develop their science language skills, and reduce achievement gaps between students, especially English language learner students and English-proficient students.

Study design

This study, an experimental trial, runs from spring 2009 to spring 2010, serving six regions in California and one urban district in Arizona. This study's unit of assignment is the teacher. At each of the six research sites, around 20 volunteer teachers participate in the study. Treatment group teachers take an Understanding Science course in August 2009, incorporating what they learn there into their instruction of force and motion in the subsequent fall semester. Control group teachers participate in their regular professional development activities and continue using their usual instructional practices in the fall. Approximately 120 volunteer teachers are randomly assigned to treatment and control groups. Where two or more teachers in a school agree to participate, teachers are randomly assigned to treatment and control groups within the school. Where only one teacher in a school is participating, likely the more common case, researchers will randomly assign schools to treatment and control groups.

With 120 teachers and an average of 40 students served by each teacher, the study is powered to detect program impacts of 0.17–0.20 standard deviation units on academic outcomes for general and English language learner students, which is equivalent to about two-thirds of a year of growth by middle school students on norm-referenced standardized tests in reading (Hill et al. 2008). And for teacher outcomes, the projected sample size is sufficient to detect impacts greater than 0.46 standard deviation units. Impacts of this magnitude would be expected at the more proximal teacher level to produce smaller subsequent impacts at the more distal student level.

One limitation of the study is the short period of implementation. Impact estimates will be based on a sample of treatment teachers who are using what they have learned from the professional development for the first time. These teachers will not have the benefit of having practiced the instructional approach in prior semesters. Another limitation is the inadequate sample size for examining differential impacts on student subgroups. Differences in program impacts for student subgroups, particularly for English language learner and English-proficient students, can only be examined in an exploratory manner. But the main effects of the intervention on the performance of all students and for English language learner students can still be examined.

Key outcomes and measures

Box 1 shows the key student and teacher outcomes with their measures for the study. These outcomes will be assessed with science content tests for students and teachers, standardized achievement tests for students, and multiple teacher surveys.

Science content tests. The researchers measure student and teacher science content knowledge using tests developed and validated as part of the ATLAST project in collaboration with the American Association for the Advancement of Science's Project 2061 (Smith and Banilower 2006a, 2006b). Funded by the National Science Foundation, ATLAST provides rigorous and well validated measurement instruments to science education studies evaluating education programs. This study uses ATLAST's Test of Force and Motion for Students and Test of Force and Motion for Teachers. These tests measure science content using the *National Science Education Standards* (National Committee on Science Education Standards and Assessment, National Research Council 1996) and reflect the research literature documenting misconceptions related to science concepts in the outcome domains. Comprising 25 multiple-choice items, the student test is administered in one 45-minute period. Although situated in instructional practices, the teacher test also comprises 25 multiple-choice items; these are of three types: knowledge of science content, use of science content to analyze student thinking, and use of science content to make decisions on instruction (Smith and Banilower 2006b).

Science achievement tests. The scores of grade 8 students in participating teachers' classes on the California Standards Test in science are collected at the end of the 2009/10 school year.

Teacher surveys. Developed for prior evaluations of Understanding Science (Heller 2006), the teacher surveys include questions eliciting teachers' classroom instructional practices on science, attitudes about teaching science, and beliefs about science and children's learning. These are all factors that could mediate the relationship between knowledge gained from the professional development and actual classroom practice.

Box 1. Study outcomes and their measures

Student outcome measures
Student content knowledge in force and motion

- ATLAST Test of Force and Motion for Students
  Student science achievement
- California Standards Test in Science
- Physical science cluster of the California Standards Test

Teacher outcome measures
Teacher content knowledge in force and motion

- ATLAST Test of Force and Motion for Teachers
  Teacher instructional strategies
- Strategies for developing students' conceptual understanding of science ideas (teacher survey)
- Strategies for revealing student ideas about science (teacher survey)
- Strategies for supporting English language learner students' science learning (teacher survey)
  Confidence in teaching abilities
- Confidence in ability to teach force and motion (teacher survey)

Data collection approach

Outcomes are measured for treatment- and control-group teachers, students, and classrooms through data collected during both the 2008/09 and 2009/10 school years. Teacher pre- and post-course outcome measures—teacher surveys and science content tests—are administered in the spring before and the winter after completing the professional development course. All participating teachers complete a precourse survey in the spring before the intervention and a postcourse survey in the winter following it.

Outcomes are also measured for students in two physical science classes per teacher, about a third of which the researchers predict are English language learner students. All students receive a science content pretest on force and motion in fall 2009 and a posttest within two weeks of receiving classroom instruction on force and motion sometime in the fall 2009 semester. Students' scores on standardized science achievement tests are obtained at the end of the academic year in which the experiment is conducted (spring 2010). Both the total scores and the subscores on the physical science cluster will be analyzed. These test score data are available only for the sites in California, or for an estimated 100 teachers. Assuming 50 teachers per condition, statistical power should be adequate to estimate program impacts on student outcomes using the California subsample.

Analysis plan

To estimate program impacts, the outcomes for students and teachers in treatment classrooms are compared with those for their counterparts in control classrooms. In addition, multilevel regression models, which account for school and classroom clustering, are used to analyze Understanding Science's effect (Goldstein 1987; Raudenbush and Bryk 2002; Murray 1998). Each impact analysis will control for baseline measures (pretests) of outcome variables and other student- and teacher-level covariates. These models will be used to estimate impacts on student performance for all participating students and for English language learner students.

In addition to examining the main effects of the program on student performance, exploratory analyses will be used to examine program impacts for certain student subgroups: English language learner and English-proficient students, low- and high-performing students, students in different racial/ethnic groups, and students from low-income households.

Procedures described by Schochet (2008) are used to account for testing multiple hypotheses involving primary outcome variables. Five outcome domains were delineated for this study: student content knowledge, student science achievement, teacher content knowledge, teacher instructional strategies, and teacher confidence in ability to teach force and motion. Multiple comparison procedures are used for each outcome domain to reduce the probability of finding statistically significant program impacts when they are due to chance alone.

Principal investigator

Joan Heller, PhD, Heller Research Associates

Contact information

Dr. Thomas Hanson
Regional Educational Laboratory West
4665 Lampson Avenue
Los Alamitos, CA 90720-5139
Voice: (562) 799-5170
Fax: (562) 799-5151
Email: thanson@wested.org

Region: West

This summary is also available at:
http://ies.ed.gov/ncee/edlabs/projects/rct_87.asp?section=ALL

References

California Department of Education. (2009). *Education Data Partnership: Fiscal, Demographic, and Performance Data on California's K-12 Schools.* Retrieved April 17, 2009, from http://www.ed-data.k12.ca.us/.

Carlsen, W.S. (1991). Subject-matter knowledge and science teaching: A pragmatic perspective. In J. Brophy (Ed.), *Advances in research on teaching,* Vol. 2. Greenwich, CT: JAI Press.

Carlsen, W.S. (1993). Teacher knowledge and discourse control: Quantitative evidence from novice biology teachers' classrooms. *Journal of Research in Science Teaching, 30* (5), 471–81.

Druva, C.A., and Anderson, R.D. (1983). Science teacher characteristics by teacher behavior and by student outcome: A meta-analysis of research. *Journal of Research in Science Teaching, 20* (5), 467–79.

Goldstein, H. (1987). *Multilevel models in educational and social research.* London: Oxford University Press. Hashweh, M. (1987). Effects of subject matter knowledge in the teaching of biology and physics. *Research and Teacher Education, 3* (2), 109–20.

Heller, J.I. (2006). *Final evaluation report for Science Cases for Teacher Learning: Impact on teachers, classrooms, and students, project years 2000–2003.* WestEd and Stuart Foundation.

Heller, J.I., Daehler, K., and Shinohara, M. (2003). Connecting all the pieces: Using an evaluation mosaic to answer an impossible question. *Journal of Staff Development, 24*, 36–41.

Heller, J.I., Daehler, K.R., Shinohara, M., and Kaskowitz, S.R. (2004, April 2). *Fostering pedagogical content knowledge about electric circuits through case-based professional development.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Vancouver, Canada.

Heller, J.I., and Kaskowitz, S.R. (2004). *Final evaluation report for Science Cases for Teacher Learning: Impact on teachers, classrooms, and students, Project years 2000–2003.* Technical report submitted to WestEd and Stuart Foundation.

Hill, C.J., Bloom, H.S., Black, A.R., and Lipsey, M.W. (2008) Empirical benchmarks for interpreting effect sizes in research. Child Development Perspectives, 2(3),172–77.

Lee, O. (2002). Science inquiry for elementary students from diverse backgrounds. In W. G. Secada (Ed.), *Review of Research in Education,* Vol. 26. Washington, DC: American Educational Research Association.

Lee, O., and Fradd, S.H. (2001). Instructional congruence to promote science learning and literacy development for linguistically diverse students. In D.R. Lavoie and W-M. Roth (Eds.), *Models for science teacher preparation: Bridging the gap between research and practice.* Dordrecht, the Netherlands: Kluwer Academic Publishers.

Murray, D.M. (1998). *Design and analysis of group randomized trials.* New York: Oxford University Press.

National Committee on Science Education Standards and Assessment, National Research Council. (1996). *National Science Education Standards.* Washington, DC: National Academy Press.

Raudenbush, S.W., and Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Sage: Newbury Park, CA.

Rosebery, A., B. Warren and F. Conant (1992). Appropriating scientific discourse: Findings from language minority classrooms. *The Journal of the Learning Sciences, 2*, 61–94.

Schochet, P.Z. (2008). *Guidelines for multiple testing in impact evaluations of educational interventions* (MPR Reference No. 6300-080). Princeton, NJ: Mathematical Policy Research, Inc.

Smith, S.P. and Banilower, E.R. (2006a, April 3–6). *Measuring middle grades students' understanding of force and motion concepts: Insights into the structure of student ideas.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Smith, S.P. and Banilower, E.R. (2006b, April 3–6). *Measuring teachers' knowledge for teaching force and motion concepts.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.