THE EFFECTS OF TWO PROFESSIONAL DEVELOPMENT MODELS ON TEACHERS' KNOWLEDGE ABOUT ASSESSMENT AND CONFIDENCE, SELF-EFFICACY AND INTEREST TOWARD TECHNOLOGY

Melfried Olson	Hannah Slovin	Judith Olson	Paul Brandon	Yue Yin
University of Hawai'i	University of Hawai'i	University of Hawai'i	University of Hawai'i	University of Illinois-Chicago

This paper reports on two different models of professional development that were created to investigate the use of formative assessment in a networked classroom. Participants were divided into two groups with one group receiving formative assessment without networked technology in the first year while the second group received formative assessment along with technology. Data was gathered on participants' knowledge of formative assessment and their attitudes toward the use of technology.

INTRODUCTION

Formative assessment is well accepted as an effective way to improve students' achievement, but it has been challenging for teachers to use it efficiently and effectively. Project FANC (Formative Assessment in a Networked Classroom) is a three-year research project funded by the National Science Foundation[1]to investigate the use of formative assessment as it affects middle grades student learning of algebra concepts. In particular, we are studying the effects of two particular professional development models for using formative assessment in a connected classroom. The research is comparing these models by considering growth in student achievement, teacher and student attitudes, and effective formative assessment practices. In this paper, we report on how participation in the two professional development models in the first year affected teachers' knowledge about assessment and attitudes toward technology, specifically their confidence, self-efficacy, value and interest.

[1] The research reported in this paper was generated by the grant, *The Effects of Formative Assessment in a Networked Classroom on Student Learning of Algebraic Concepts (DRL 0723953)* funded by National Science Research and Evaluation on Education in Science and Engineering (REESE) program. The views expressed in this article are the views of the authors and do not necessarily represent the views of the National Science Foundation.

THEORETICAL FRAMEWORK

The theoretical framework for this research project was informed from both research in formative assessment and the use of technology to facilitate uses of formative assessment. Fundamental ideas from each of these are described below.

Formative Assessment

Black & Wiliam (1998a) defined “formative assessment” as “all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (p. 7). In other words, in addition to traditional assessments, *formative assessment for learning* includes instructional activities such as questioning, discussion, seatwork, and student self-assessment. Evidence has shown that formative assessments, if appropriately implemented in teaching, can produce substantial learning gains for students at different ages and across different subjects (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998a; 1998b; Wiliam, Lee, Harrison, & Black, 2004).

Shavelson, Yin, Furtak, Ruiz-Primo, Ayala, & Young (2006) classified formative assessment techniques into three categories on a continuum based on the amount of planning involved and the formality of technique used: (a) on-the-fly formative assessment, which occurs when teachable moments unexpectedly arise in the classroom; (b) planned-for-interaction formative assessment, which is used during instruction but prepared deliberately before class to align closely with instructional goals; and (c) formal-and-embedded-in-curriculum formative assessment, which is designed to be implemented after a curriculum unit to ensure that students have reached important goals before moving to the next unit.

Despite their variety, when formative student assessments are used, common steps are explicitly or implicitly involved: (a) determining achievement goals that students are expected to reach—the expected level; (b) collecting information about what students know and can do—the actual level; (c) identifying the gap between the actual level and expected level; and (d) taking action to close the gap. Sadler (1989) also addressed similar steps to (a), (b), and (d).

Rather than merely receiving a grade, formative assessment provides specific feedback for each student. Moreover, self and peer formative assessments can help students to develop the habit of reflection and to become more aware of the learning goal, learning gap, and how to close the gap (Black et al., 2004).

Formative Assessment and Technology

As straightforward as it may sound, in practice, formative student assessment has proven difficult to implement (Ruiz-Primo & Furtak, 2006; Shavelson, R. J., Yin, Y., Furtak, E. M., Ruiz-Primo, M. A., Ayala, C. C., Young, D. B., et al., 2006; Yin, 2005). One of the challenges is the amount of time many formative assessment strategies take, making them impractical (Black & Wiliam, 1998b). For example, it is time-consuming for teachers to count students' votes, and it is almost impossible for teachers to provide specific feedback on each student's work in a timely enough manner to make it useful for the lesson in progress. Other challenges involve a concern about making students' misunderstandings public.

In *How People Learn* (NRC, 1999), classroom networks are considered one of the most promising technology-based education innovations for transforming the classroom environment. Some early findings demonstrate its potential to overcome one of the greatest hurdles to improving formative classroom assessment: the collection, management and analysis of data (Roschelle, Penuel, & Abrahamson, 2004). While feedback loops in the regular classroom are very slow, classroom networked technology has the capability to provide rapid cycles of feedback to improve ongoing activity in real time.

In prior studies (Mackay, Olson, & Slovin, 2006; Dougherty, Akana, Cho, Fernandez, & Song, 2005), researchers found that in a TI-Navigator[2] networked classroom formative assessment could be conducted efficiently and anonymously and results could be used more effectively because assessments were done in real-time. Mackay, et al found that the use of TI-Navigator assisted students in pinpointing areas for questioning and saved both student and teacher time.

Dougherty, et al reported that "the use of TI-Navigator technology supports the development of a collaborative classroom environment by enhancing student interactions, focusing students' attention on multiple responses, and providing opportunities for students to peer- and self- assess student work. The ability to display a full class set of data or task responses supports a problem-solving approach to developing skills and concepts" (p. 28).

Design of Professional Development

The professional development was designed to address what teachers need to know to use formative assessments effectively: (a) knowing the algebra content, (b) designing the tools, (c) using the tools, (d) interpreting the results of the gathered responses, and (e) taking action based on the results (Ayala & Brandon, 2008). Particular attention was paid to the type of tasks used for formative assessment, questioning and discourse strategies, developing students' awareness of assessments and helping students use information to become reflective of their own progress (Black & Wiliam, 1998a). The matter of content and tasks was addressed by focusing on identifying the characteristics of "worthwhile mathematical tasks" (NCTM, 1991) using a task analysis scheme adapted from Stein, Smith, Henningson & Silver, (2000) on how mathematical tasks differ with respect to their levels of cognitive demand. In

---

[2] TI-Navigator™ is a networking system developed by Texas Instruments that wirelessly connects each student's graphing calculator to a classroom computer.

working on the questioning and discourse components of the design, the professional development emphasized formulating questions that provide information about students' thinking and understanding. Rich questioning provides teachers not just with evidence about what their students can do, but also what the teacher needs to do next, in order to broaden or deepen understanding (Wiliam, 1999). The design included time provided for teachers to work with their own curriculum to plan using formative assessment when they teach the targeted algebra concepts (Black & Wiliam, 2005a).

METHODOLOGY

Research Question

The following research question guided the study reported on in this paper: What are the between-group differences in changes in teachers' knowledge about assessment and in their attitudes (confidence, self-efficacy, value and interest) toward technology after the first year of participation in a project examining two professional development models?

Procedure

Thirty-two teachers from 15 schools throughout the State were recruited and randomly assigned within each school to either the formative assessment (FA) or Navigator (NAV) group. This randomised-block design helped us control for extraneous variables, such as student background and school context (e.g., teachers' work load, curriculum, class equipment, and community support). Assigning the teachers to groups within schools is a matching procedure that reduces the number of units necessary for group-randomized studies (Murray, 1998).

The NAV group was trained in both the TI Navigator and formative assessment, and the FA group was trained in formative assessment only. The five-day training for both groups was designed to follow parallel models where possible. For the group using classroom networking we applied the formative assessment model, questioning strategies, and mathematics activities were connected to the use of TI-Navigator. Since participants in this group were to learn how to use the technology, significant time for hands-on experiences with the features of TI-Navigator was provided.

We conducted five, half-day follow-up sessions and periodic coaching assistance as mentors, thereby addressing the teachers' need for considerable practice to learn how to adapt formative assessment practices to their classrooms (Wiliam, Lee, Harrison, & Black, 2004; Yin, 2005).

Data Collection

Teacher data were collected with three instruments that were developed for the study: (a) the Teaching Practice and Perceptions Questionnaire, which includes two scales about teacher collaboration and four about teacher support; (b) the Assessment Knowledge, Self-Efficacy, and Practice Survey, which includes four scales about assessment knowledge (assessment in general, student learning, subject content, and formative assessment) and one about teacher self-efficacy in using formative assessment; and (c) the Using Technology Questionnaire, which includes four scales about teachers' perceptions of using technology.

We developed the initial questionnaires based on our NSF proposal, previous research project (Shavelson et al., 2008; Yin, 2005), and some existing questionnaires, such as the teacher questionnaires used in TIMSS. After developing the initial questionnaires, our research team (consisting of teachers, teacher trainers, content experts, evaluation experts, and research methodologists) reviewed, discussed, and revised the questionnaires multiple rounds. After our research team had approved the questionnaires, we pilot tested the questionnaires among mathematics teachers in four elementary schools and one middle school. Twenty-nine elementary school and 21 middle school mathematics teachers responded the questionnaires. Based on these teachers' responses to the questionnaires and their feedback about the questionnaire design, we further improved the questionnaires.

RESULTS AND DISCUSSION

Data were collected from 15 teachers in the FA group and from 16 teachers in the NAV group at the start of the first-year summer professional development and collected posttest data from all 16 of the FA group and 15 of the NAV group in May of the following school year. We conducted internal consistency analyses for each of the pretest and posttest datasets (yielding alpha coefficients) on each of the 15 scales that were included in the three instruments. We identified and deleted the items with low item-total correlations on both administrations of the instruments. Then we calculated total scores for each of the scales by averaging all the items measuring that scale. Of the 15 scales, each scale has 3 to 10 items and the alphas were all greater than .70 for both the pretest and posttest except for one scale. (The exception was the K*nowledge About Student Learning* Scale, which was .60 on the pretest and .65 on the posttest). We deemed these results adequate for further analyses of the scale scores.

The Teaching Practice and Perceptions Questionnaire includes two scales about teacher collaboration and four about support for the teachers. The project was not expected to affect either collaboration or support, and the validity of the study design is enhanced to the extent to which differences between groups are not found on either of the two sets of scales. The results of multivariate analyses of variance showed no significant differences between groups on either the pretest or posttest, supporting the validity of the design (in which teachers were randomly assigned to groups) and showing that some aspects of the context within which the teachers worked should not be expected to have affected the results on the other instruments that we administered to the teachers.

The data on the four assessment-knowledge scales were analysed with a split-plot ANOVA design with group as the independent variable and pre-post change scores as the dependent variables. The results showed a significant difference between the FA group and the NAV group *F (1, 26)* = 4.86, *p* = .036, with the FA group. These results suggest that, after the first year, the FA group came to know more than the NAV group about assessment in general, student learning, subject content, and formative assessment. This pattern was expected because the knowledge about these topics was the primary focus of the PD, the follow-ups, and coaching sessions for the FA group, while the PD for the NAV group focused more on using TI-Navigator for formative assessment.

For the teacher scale about self-efficacy in using formative assessment, we conducted an ANCOVA, with posttest self-efficacy as the dependent variable, group as the independent variable, and pretest self-efficacy as the covariate. The results showed no statistical differences between groups at the .05 level. While the FA group's knowledge about FA increased significantly, the two groups' perceived self-efficacy about their capabilities to organize and execute formative assessment and manage prospective situations did not differ significantly at the end of the first year. As Bandura (1994) has noted, the most effective way of creating a strong sense of efficacy is through mastery experience. Both the FA and NAV groups had a similar sense of mastery of FA even though there were differences in the technology that they used. Differences among groups on self-reports about knowledge, as found in the results on the assessment-knowledge scales, are not always confirmed by differences in attitudes.

For the results on teachers' use of technology, we conducted four ANCOVAs with post scores for the four scales on teachers' perceptions about using technology as dependent variables, group as the independent variable, and pretest scores as the covariates. The results show that after controlling for pretest, the FA group scored significantly higher on the value of using technology (*F* (1, 24) = 11.25, *p* = .003) and interest in technology (*F (1, 25)* = 15.73, *p* = .001) than the NAV group. Adjusted mean of value was 4.68 for the FA group and 4.09 for the NAV group. The adjusted mean of the interest was 4.95 for the FA group and 4.37 for the NAV group. However, the two groups did not differ on self-efficacy of using technology or confidence in using all kinds of technology. These results should be interpreted in light of the fact that the two groups were provided with different technology. Both groups were given a laptop, LCD projector, digital visualizer, and classroom set of graphing calculators while the NAV group was also given a TI-Navigator System. The results suggest that both groups had similar levels of confidence and self-efficacy with the use of the technology with which they were provided. However, the significant difference in interest and value of technology raises some interesting questions. Was that difference a result of the type of technology they were using? The NAV group had more technology to master the first year to see the value of the pay off of their investment of time and preparation for using the technology. Could the difference be related to the difference in demographic variables such as age and years of teaching? The FA group was younger and average 6.9 years of teaching experience compared to the NAV group at 12.6 years. Therefore, the FA group was closer to being 'digital natives' while the NAV group could be considered closer to 'digital immigrants'. How much did the PD, follow-ups, and coaching of the two groups' use of the technology and emphasis create an interest and value of the technology they were using? Further investigation is needed to determine the reasons for the difference between the two groups.

In summary, the results suggest that aspects of the schooling context did not affect the results and that the FA group showed greater gains in knowledge of assessment. The FA group did not show more positive attitudes toward using assessment--results that are not uncommon in studies of teacher learning and attitudes. The NAV group, which used more technology than the FA group, did not show more confidence in technology use than did the FA group at the end of the year. This finding might reflect the steep learning curve required to use the TI Navigator. It might also show that the FA had an advantage in using technology because of higher beginning skill and experience levels—a hypothesis that we plan to examine in the future.

References

Ayala, C. C., & Brandon, P. R. (2008). Building evaluation recommendations for improvement: Insights from student formative assessments. In N. L. Smith & P. R. Brandon (Eds.), Fundamental issues in evaluation (pp. 159–166). New York: Guilford.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998).

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*, 8-21.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*, 7-68.

Black, P., & Wiliam, D. (1998b). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Black, P., & Wiliam, D. (2005a). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal, 16*(2), 249-261.

Black, P., & Wiliam, D (2005b). Chapter 5: Developing a theory of formative assessment in J. Gardner (Ed*). Assessment for learning: practice, theory, and policy*, London, UK: Sage.

Dougherty, B., Akana, K., Cho, C., Fernandez, J. C., & Song, M. (2005). TI-Navigator Technology and Algebra I. Report submitted to Texas Instruments.

Mackay, I., Olson, J., & Slovin, H. (2006). Effects of TI-Navigator on student achievement, formative assessment and students' beliefs in high school reform classes. Report submitted to Texas Instruments.

National Council of Teachers of Mathematics (1991). The *Professional Standards for Teaching Mathematics,* Reston: Author.

National Research Council (NRC). (1999). *How People Learn*. Washington, DC: National Academy Press.

Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3&4), 237-263.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Shavelson, R., Yin, Y., Furtak, E., Ruiz-Primo, M., Ayala, C., Young, D.B., Tomita, M., Brandon, P., & Pottenger, F.M. (2008). On the role and impact of formative assessment on science inquiry teaching and learning. In J. Coffey & C. Stearns (Ed). *Assessing science learning: Perspectives from research and practice* (pp. 21–36). Arlington, VA: National Science Teachers Association.

Shavelson, R. J., Yin, Y., Furtak, E. M., Ruiz-Primo, M. A., Ayala, C. C., Young, D. B., et al. (2006). *On the role and impact of formative assessment on science inquiry teaching and learning*: National Science Teachers Association.

Stein, M. K., Smith, M. S., Henningson, M. A., & Silver, E. A. (2000). Implementing standards based mathematics instruction: A casebook for professional development, Teachers College Press.

Wiliam, D. (1999), Formative assessment in mathematics: Part 1:Rich questioning, Equals, 5(2), 14-18.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004) Teachers developing assessment for learning: impact on student achievement. *Assessment in Education, 11(1),* 49-65.

Yin, Yue (2005). *The influence of formative assessments on student motivation, achievement, and conceptual change*. Ph.D. dissertation, Stanford University, United States -- California. Retrieved January 15, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3186430).