EFFECTS OF TWO DIFFERENT MODELS OF PROFESSIONAL DEVELOPMENT ON STUDENTS' UNDERSTANDING OF ALGEBRAIC CONCEPTS

Judith Olson, Seongah Im, Hannah Slovin, Melfried Olson, Michael Gilbert, Paul Brandon
University of Hawai'i at Mānoa
jkolson@hawaii.edu, seongahi@hawaii.edu, hslovin@hawaii.edu, melfried@hawaii.edu, mjgilber@hawaii.edu, brandon@hawaii.edu

Yue Yin
University of Illinois-Chicago
yueyin@uic.edu

This paper examines the effects of the first year of a two-year project to investigate formative assessment in a networked classroom. Participants were divided into two groups; one group receiving professional development on formative assessment with networked technology while the second group received professional development only on formative assessment. Data were gathered on participants' knowledge of formative assessment, teacher pedagogical content knowledge, mathematics background, and attitudes toward technology. Student data were collected and analyzed to examine the effects of teacher variables on student achievement.

Past studies indicate professional development (PD) was the difference between teachers using technology to emphasize critical-thinking and problem-solving skills versus skill and drill (Wenglinsky, 1998; Brannigan, 2002). Thus, PD for integrating technology into the teaching of mathematics must include a strong focus on pedagogical approaches that have potential for impacting learning. This paper describes two models of PD for formative assessment and classroom-connected technology and reports on results from the first year of implementation. Building upon theory and practice that showed positive outcomes, Project FANC (Formative Assessment in a Networked Classroom)[1] was designed to address concerns raised about both the use of formative assessment and integration of technology. Teachers, recruited from 15 middle schools, were assigned to two models of PD. Each group participated in a five-day summer PD with five follow-up sessions during the school year along with in-school coaching. The design of both models took into consideration that one of the most salient characteristics of effective PD is providing teachers with opportunities to work with colleagues, both in their school building and beyond, giving them opportunities to learn from one another's successes and failures and to share ideas and knowledge (National Center for Research on Teacher Learning, 1995).

Theoretical Framework

Connected Classroom Technology

In *How People Learn* (NRC, 1999), one of the most promising technology-based innovations noted for transforming the classroom environment was the use of networks. In prior research on implementation of a networking system, use of TI-Navigator was found to support development of a collaborative classroom environment by enhancing student interactions, focusing students' attention on multiple responses, and providing opportunities to peer- and self-assess student work. The ability to display class data or responses supports a problem-solving approach to developing skills and concepts (Dougherty, Akana, Cho, Fernandez, & Song, 2005; Mackay, Olson & Slovin, 2006). One of the most promising uses of TI-Navigator is its potential to overcome a significant hurdle to improving formative classroom assessment: the collection, management and analysis of data (Roschelle, Penuel, & Abrahamson, 2004). While feedback loops in a regular classroom are slow, classroom networked technology can provide rapid cycles of feedback in real time. Beaty and Gerace (2009) remind us that technology is a tool and that pedagogical approaches aid or impact learning. Yet, although Owens, Demana, Abrahamson, Meagher & Herman (2004) found teachers in TI-Navigator classes perceived as more responsive to individual needs, more focused on knowledge building and assessment, and more community centered, they may not have changed their instruction based on information obtained.

There are four distinct, intertwined, functions of the TI-Navigator system particularly helpful for implementing formative assessment: 1) Quick Poll—to immediately collect and display all the students' answers to a question; 2) Screen Capture—to monitor individual students' work at anytime; 3) Learn Check—to administer quick, frequent formative assessments and provide timely feedback; and 4) Activity Center—allowing students to work collaboratively to contribute individual data to a class activity. While these tools are directly appropriate for formative assessment, teachers who may make significant changes in the use of technology in their classes do not necessarily make full use of the potential of the connected classroom for formative assessment (Owens, Pape, Irving, Sanalan, Boscardin, & Abrahamson, 2008).

Formative Assessment

Similarly, evidence has shown that appropriately implemented formative assessments can produce substantial learning gains at various ages and across subjects (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998b; Wiliam, Lee, Harrison, & Black, 2004). Black and Wiliam 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). Formative assessment for learning includes activities such as questioning, discussion, seatwork, and student self-assessment. Teachers' preparation in assessment is often non-existent and teachers' content knowledge may be insufficient for deep understanding of concepts and principles they teach and assess (Heritage, 2007). Moreover, Heritage suggests that while there is no best way to carry out formative assessment, assessment and teaching are reciprocal activities, and need to be firmly situated in the practice of educators. Black and Wiliam (1998a) demonstrated that formative assessment improves student achievement if it guides changes in day-to-day practice. 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: 1) on-the-fly formative assessment, when teachable moments unexpectedly arise; 2) planned-for-interaction formative assessment, used during instruction but prepared deliberately before class to align with instructional goals; and 3) formal and-embedded-in-curriculum formative assessment designed to be implemented at the end of a unit of instruction to ensure students reach important goals before moving on. Despite their variety, when formative assessments are used, common steps are explicitly or implicitly involved: 1) determining achievement goals students are expected to reach—the expected level; 2) collecting information about what students know and can do—the actual level; 3) identifying the gap between the actual level and expected level; and 4) taking action to close the gap.

Many formative assessment strategies, however, are challenging because they take too much time to be used practically (Black & Wiliam, 1998b). It is time-consuming to count students' responses, and almost impossible to provide specific feedback on each student's work in a typical teaching load of four to six classes and 20 to 30 students per class. Even if teachers believe the time investment for formative assessment will yield reward in the future, Black and Wiliam (1998b) noted that teachers' workloads are often overwhelmed by agendas and dominated by district curriculum specifications and high-stake tests.

Methodology

Project FANC is a three-year research project funded by the National Science Foundation to investigate the use of formative assessment in a networked classroom as it affects 7th grade student learning of algebra concepts. Thirty-two teachers from 15 schools were recruited and randomly assigned within schools to the Formative Assessment (FA) or Navigator (NAV) group. This randomized-block design helps control for extraneous variables, such as student background and school context (e.g., teachers' work load, curriculum, class equipment, and community support). In the first year of the project, students' understanding of algebraic concepts taught by 7th grade teachers participating in the NAV group were compared to students' understanding of algebraic concepts taught by 7th grade teachers participating in the FA group.

While Project FANC provided PD to participants over a two-year period, only the first year of the PD is described since results are being reported only for that period. Input from the eight Advisory Board members along with numerous articles on PD, formative assessment, technology and related topics (Ayala & Brandon, 2008; Black & Wiliam, 1998a, 1998b, 2005; Gearhart, & Saxe, 2004; Guskey, 2007/2008; Wiliam, Lee, Harrison, & Black. 2004) were taken into consideration to design the two models of PD. The formative assessment model (Stiggins, Arter, Chappuis, & Chappuis, 2004, p.42), questioning strategies, and rich mathematics activities were essential components in both PD experiences, although the actual delivery and focus varied. The PD models delivered to two groups, FA and NAV, are briefly described below.

FA teachers participated in five days of PD on strategies for using formative assessment in their classrooms (not using TI-Navigator). In addition, five follow-up days and coaching were conducted throughout the academic year. The formative assessment model, questioning strategies, and mathematics activities focused on designing appropriate tasks where students could demonstrate thinking. In-depth discussions were held on techniques to analyze student work and interpret results in terms of revealing student understanding. FA participants were provided with opportunities to examine tasks and lessons and discussed how the tasks fit expected learning progressions. Participants worked on selecting and modifying tasks using processes of reversibility, flexibility, and generalization to deepen students' understanding of content. Ideas on creating a learning environment that included an assessment conversation with student-to-student interactions were shared with participants. At follow-up sessions, participants shared progress on implementation, examples of student work, and changes in the classrooms climate and were given opportunities involving classroom activities and questioning strategies.

NAV teachers participated in five days of PD on formative assessment and strategies for using TI-Navigator in their classrooms. In addition, five follow-up days and coaching were conducted throughout the academic year. The formative assessment model, questioning strategies, and activities were connected to the use of TI-Navigator. Significant time for hands-on experiences with TI-Navigator was provided. As participants practiced using handheld technology and TI-Navigator, emphasis was on creating and choosing good tasks and asking good questions. The PD focused on the 'added value' of handheld technology in a connected classroom. During follow-up, participants shared how they implemented formative assessment with TI-Navigator and discussed troubleshooting tips and changes in classroom climate.

At the beginning of the project all participants were given MacBook laptops, LCD projectors, Elmo visualizers (document cameras), and a classroom set of TI-73 graphing calculators. NAV participants were provided TI-Navigator Systems for their classrooms.

Data Collection

The results in this paper focus on student growth in achievement on a student assessment of algebraic concepts with a focus on patterns and relations, teacher's responses on questionnaires, and assessment of teachers content knowledge. A student assessment was developed to determine the extent to which participating students showed pre-post gains after implementation of the intervention. As much as possible, algebra items were drawn from existing sources such as state assessments and released NAEP, TIMSS, and SAT items. Items were reviewed for content, perceived alignment with the project's target content, perceived difficulty, and quality (e.g., clearness of item stems, etc.). Eighty-five items were selected from these sources. Once standards and knowledge levels were reconciled, items were selected for pilot testing. Two versions of a pilot-test were created from 48 items and piloted with over 1000 students in Grades 7 and 8. Item Response Theory was used to select items included on FANC Student Assessment.

To measure the growth of participants' content knowledge for teaching (CKT), the University of Michigan's *Learning Mathematics for Teaching* (LMT) instrument was administered at the beginning of the summer institute in year one, and then again after one year of participation. The LMT measures have been shown to be a significant predictor of student achievement (Hill, Rowan, & Ball, 2005). Teacher data were collected with three instruments that were developed for the study: a) the Teaching Practice and Perceptions Questionnaire includes two scales about teacher collaboration and four about teacher support; b) the Assessment Knowledge, Self-Efficacy, and Practice Survey 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) Using Technology Questionnaire includes four scales about teachers' perceptions of using technology.

Results

Proposed Model

The analysis reported here is for 23 participants (FA (11), NAV (12)) for whom complete student data (1,629 students) was collected. A two-level hierarchical model (Luke, 2004) included one student covariate, the pretest score, and eight teacher variables. The six continuous variables included: a) teacher pretest score (labeled "TeaPre"); b) teachers' years of teaching experience ("TeaExp"); c) teachers' efficacy in formative assessment ("FAefficacy"); d) teachers' efficacy in technology ("TCHefficacy"); e) teachers' perception about their students' motivation ("StuMotiv"); and f) teachers' perception about overall support from school, colleagues, parents, and students ("Support"). The two discrete variables included a) the experiment group (FA vs. NAV) and b) teachers' mathematics or mathematics education major (major and minor vs. no math background in university, labeled "NoMath" and "Math").

The level 1 units were students taught by teachers who participated in FANC. The level 2 units were the teachers in FANC. Multilevel modeling was conducted using SAS PROCMIXED, Version 9. At the student level, the intercept was the only random component. A model in which the student pretest score was entered as a random variable failed to converge; therefore, the decision was made to treat the covariate as fixed effect. No interactions were hypothesized at the teacher level. Interactions between the level 2 predictors as well as between the level 1 and level 2 predictors were not significant. As anticipated, the interaction effect between the pretest score and the experiment group was not significant. This means students' pretest scores did not differ across the FA and NAV groups. Another interaction effect between the pretest score and teachers' mathematics education was also not significant, suggesting students' pretest scores did not differ by whether teachers had mathematics education in undergraduate and/or graduate schools. Based on these preliminary examinations, the following model is proposed:

Level 1: $Y_{ij} = \beta_{0j} + \beta_1(\text{StuPre}) + r_{ij}$, where $r_{ij} \sim N(0, \sigma^2)$, β_{0j} a random intercept for a teacher j, β_1 a fixed slope of students' pretest scores, and r_{ij} random error associated with student i/teacher j.

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{TeaPre})_j + \gamma_{02}(\text{ExpGp})_j + \gamma_{02}(\text{MathBack})_j + \gamma_{03}(\text{TeaExp})_j$$
$$+ \gamma_{04}(\text{FAefficacy})_j + \gamma_{05}(\text{TCHefficacy})_j + \gamma_{06}(\text{StuMotiv})_j + \gamma_{07}(\text{Support})_j + \mu_{0j}$$

where $\mu_{0j} \sim N(0, \tau_{00})$, γ_{00} the grand mean of students' posttest scores, γ_{0k} the slope of the variable k where k=1,2,..8, and μ_{0j} the random deviation of students' mean posttest score of teacher *j* from the mean.

Preliminary Analysis

Table 1 presents descriptive statistics for student pretest and posttest scores. Table 2 presents descriptive statistics for six continuous variables in the model at student and teacher levels.

Table 1. Descriptive Statistics for Students' Pretest and Posttest Scores by Teacher Group

Group	Statistic	Student pretest	Student posttest
TI-Nav	*M*	13.720	16.130
(*n*=692)	*SD*	5.916	8.514
FA	*M*	14.040	16.850
(*n*=937)	*SD*	6.083	6.969
Total	*M*	13.900	16.550
(*n*=1,629)	*SD*	6.103	7.501

Table 2. Descriptive Statistics for Teacher Variables

Group	*Statistic*	*TeaPre*	*TeaExp*	*FAefficacy*	*TCHefficacy*	*StuMotiv*	*Support*
TI-Nav	*M*	20.100	10.455	7.475	7.674	3.643	3.227
(*n*=12)	*SD*	5.820	5.716	1.163	1.146	0.503	0.858
FA	*M*	19.900	7.364	7.143	7.843	3.250	3.300
(*n*=11)	*SD*	5.195	4.945	0.649	1.322	0.707	0.461
Total	*M*	20.000	8.909	7.292	7.763	3.456	3.262
(*n*=23)	*SD*	5.370	5.450	0.905	1.213	0.626	0.682

The initial step of the multilevel analysis was to examine the unconditional means model to ascertain whether the data warranted a multilevel model. Estimated variance at teacher level (τ_{00}) was found to be 18.330 ($p<.01$), and estimated variance at the student level (σ^2) was 42.624 ($p<.01$) in this null model. Hypothesis tests indicated that both variance components were significantly different from 0. These estimates suggest students' scores differ among teachers in average posttest scores, although there is more variation among students within the teachers than between teachers. The intra-class correlation, 0.301, shows that about 30% of the total variance occurs at the teacher level—a substantial proportion. This cluster of variability between different teachers' students' scores suggests that an OLS analysis of these data would likely yield misleading results.

Random Effects

This proposed model was significantly better than one in which the intercept was included at the initial step, χ^2 (9) =10790.422-6891.167= 3899.255, p< 0.01. Table 3 presents random effects at level 1 and level 2. The conditional component representing variation among students' scores within the teacher level (σ^2) decreased from 42.624 in the unconditional means model to 20.484 (p<.01) after including the student level covariate, student pretest score. This indicated that inclusion of this student pretest score explained about 52 % ((42.624-20.484)/42.624 = 0.519) of variation within teachers. Even though the students' variation within the teacher level was substantially explained by student pretest score, the significant random variation of 20.284 shows possibilities to explain the remaining random variation using other student level variables.

The variance component representing variation between the teachers (τ_{00}) drastically decreased from 18.330 in the unconditional model to 3.369 (p<.05) after controlling for the student and teacher level covariates. The predictors included in this proposed model explained substantive amount of variation in students' posttest scores at the teacher level. The intra-class correlation was 0.155, meaning that the proportion of between teacher variance to the total variance was 15.5 % after controlling for the predictors. The change of intra-class correlations from 0.301 in the unconditional means model to 0.155 in the proposed model also demonstrates that the teacher level predictors explained well the variation of students' posttest scores at the teacher level. Even though these results showed the random variations at the student and teacher levels were substantially explained by the predictors, the remaining clusters of variability at the teacher level as well as at the student level warrants further development of the multilevel model.

Table 3. Random Effects at Level 1 and Level 2

Effect	Estimate	*SE*	Wald *Z*	*p* (1-sided)	95% confidence interval Lower	Upper
Random effect at Level 1						
Residual	20.484	0.854	23.999	0.000	18.878	22.227
Random effect at Level 2						
Intercept	3.369	1.442	2.337	0.019	1.457	7.793

Fixed Effects

Table 4 shows results of hypothesis tests for fixed effects of student level and teacher level predictors. The fixed effect of students' pretest score was significant after controlling for the teacher level predictors. One of the variables of interest was Group, indicating whether a teacher belonged to the FA or NAV group. The student posttest mean in the FA group was 3.462 points higher than that of NAV group after controlling for the students' pretest scores and other teacher background variables. Teachers' knowledge in the subject content was a significant predictor. When a teacher scored 1 point higher in the pretest, their students achieved 0.448 higher points after the influence from the other variables was partialled out. Teachers' years of experience in teaching and mathematics education did not significantly influence students' achievement.

The variables asking the teachers' perceived efficacy in formative assessment and in technology use were not significant, nor was the teachers' perception about their students' motivation in class a significant predictor, after taking into account the other student and teacher level predictors. Finally, teachers' perception about overall support from school, colleagues, parents, and students showed a positive effect on students' achievement. Teachers who differ by 1 point in their perception about overall support differ by 3.041 points in students' achievement.

Table 4. Fixed Effects of the Student and Teacher Predictors

Effect	Estimate	*SE*	Approx *df*	*t*	*p* (2-sided)	95% confidence interval Lower	Upper
Intercept	6.084	7.101	13.938	0.857	0.406	-9.152	21.321
StuPre	0.797	0.023	1166.851	34.572	0.000	0.752	0.842
Group	3.462	1.465	13.179	2.363	0.034	0.301	6.623
MathBack	2.593	1.446	14.150	1.793	0.094	-0.506	5.691
TeaPre	0.448	0.163	13.281	2.745	0.016	0.096	0.800
TeaExp	-0.259	0.126	13.865	-2.060	0.059	-0.529	0.011
FAefficacy	-1.190	0.700	15.036	-1.699	0.110	-2.681	0.302
TCHefficacy	-0.917	0.588	13.854	-1.558	0.142	-2.180	0.346
StuMotiv	-1.096	1.010	13.977	-1.084	0.297	-3.263	1.072
Support	3.041	1.225	13.813	2.482	0.027	0.410	5.672

Conclusions and Discussions

The purpose of the study was to use Year 1 data to examine if students' achievement in the FA and NAV groups was different, after taking into account students' pre-test scores and teachers' initial statues such as their knowledge, teaching experiences, efficacy, and content knowledge for teaching. Student's understanding of patterns in the FA group was significantly better than students in the NAV group. The teacher variables that significantly affected the students' achievement on the post-test were the teacher's content knowledge and the support from the school, parents, and students. Data from nine of the 32 Project FANC participant's student data were not included in the analysis because they were involved in the pilot testing of the student assessment. However there were no significant differences on the teacher variables between the 9 teachers and the 23 teachers who were included in the analysis. While PD can be designed to increase teacher's content knowledge, support from the school, parents, and students is a much larger issue. During Year 2, the FA group will have PD experiences with using TI-Navigator for formative assessment and the NAV group will have a continuation of PD using the Navigator for formative assessment with more emphasis on formative assessment strategies. Extensive analysis of Project FANC will be completed at the end of Year 2.

Endnotes

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

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

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.

Brannigan, Cara. (June 5, 2002). "Study: Missouri's ed-tech program is raising student achievement." *eSchool News*.

Beaty, I. & Gerace, W. (2009). Pedagogy for teaching science with classroom response

technology. Journal of Science Education and Technology, 18: 146 – 162.
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 (2005). 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.
Gearhart, M., & Saxe, G. (2004). When teachers know what students know: Integrating mathematics assessment. *Theory into Practice, 43(4),* 304-313.
Guskey, T. R. (2007/2008). The rest of the story, Educational Leadership, 65(4), 28-35.
Heritage, M. (January 2007). Formative assessment in the classroom. *EED Winter Conference: Informing Instruction, Improving Achievement.* Anchorage, Alaska.
Hill, H., Rowan, B., Ball, D. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. American Educational Research Journal. 42(2), 371407.
Luke, D.A. (2004). Multilevel Modeling. Thousand Oaks, California: Sage Publications.
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 Center for Research on Teaching and Learning. (1995). *Learning to Walk the Reform Talk: A Framework for Professional Development for Teachers.* Michigan State University.
National Research Council (NRC). (1999). *How People Learn*. Washington, DC: National Academy Press
Owens , D. T., Demana , F., Abrahamson, A. L., Meagher, M., & Herman, M. (2004). *Developing pedagogy for wireless calculator networks—and researching teacher professional development.* (Report by The Ohio State University & Better Education, Inc. National Science Foundation Grant No. ESI 01-23391 & ESI 01-23284). Washington, DC: ERIC (ERIC Document Reproduction Service No. ED479499).
Owens, D. T., Pape, S. L., Irving, K. E., Sanalan, V., Boscardin, C. K., Abrahamson, L. (2008). The connected algebra classroom: A randomized control trial, Proceedings *for Topic Study Group 22, Eleventh International Congress on Mathematics Education*. Monterrey, Mexico, Retrieved July 2, 2009 from http://tsg.icme11.org/document/get/249.
Roschelle, J., Penuel, W.R., & Abrahamson, L. (February 2004). The networked classroom. *Educational* Leadership.
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.
Stiggins, R. J., Arter, J. A., Chappuis, J. & Chappuis, S. (2004). Classroom assessment for student learning, Portland, Oregon: Assessment Training Institute, Inc.
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.
Wenglinsky, H. (1998). Does it compute? The relationship between educational technology and student achievement in mathematics. Princeton, N.J.: ETS Policy Information Center.