Jl. of Computers in Mathematics and Science Teaching (2016) **35**(2), 131-151

Classroom Connectivity and Algebra 1 Achievement: A Three-Year Longitudinal Study

KAREN E. IRVING
The Ohio State University, USA
irving.8@osu.edu

STEPHEN J. PAPE
Johns Hopkins University, USA
Stephen.Pape@jhu.edu

DOUGLAS T. OWENS
The Ohio State University, USA
owens.93@osu.edu

LOUIS ABRAHAMSON
The Better Education Foundation, USA
la@bedu.com

DAVID SILVER
RTI International, USA
dsilver@rti.org

VEHBI A. SANALAN
Erzincan University, Turkey
Sanalan.1@osu.edu

The research reported here was supported by the Institute of Education Sciences, U. S. Department of Education, through Grant R305K050045 to The Ohio State University. The opinions expressed are those of the authors and do not represent views of the U. S. Department of Education.

Findings from three years of a longitudinal randomized control trial involving a national U. S. sample of Algebra 1 teachers and students are reported. The study examines the effects of a connected classroom technology (CCT) professional development and classroom intervention on student achievement when compared to classroom instruction with graphing calculators only. The theoretical framework suggests that active learning environments facilitated by CCT are likely to broaden the representational infrastructure of the classroom and to provide timely, targeted, and accurate feedback loops to improve formative assessment opportunities and ultimately student achievement in Algebra 1. In the first three years of this study, significant effect sizes on student achievement ranged from 0.19 to 0.37. These medium-sized effects are relatively rare for large-scale randomized experiments in education.

Algebra I learning poses a key hurdle for students' mathematical development (National Mathematics Advisory Panel [NMAP], 2008; RAND Mathematics Study Panel & Ball, 2003). Algebra completion by the ninth grade is a "leading indicator" for future mathematics achievement, completion of a bachelor's degree, and higher life-long wage earning (Munsen, 2010; Snipes & Finkelstein, 2015). The role of student success in algebra has been highlighted as an important factor in increasing college access for low-income students (USEOP, 2014). Various relationships point toward potential explanations for students' hurdles with algebra including difficulties with algebraic notation (MacGregor & Stacey, 1997) and students' representational fluency (Gunpinar & Pape, 2016; Sert, 2014).

Recent reform movements in mathematics education such as the Common Core State Standards-Mathematics (CCSS-M; National Governors Association Center for Best Practices, Council of Chief State School Officers [NGACBP-CCSSO], 2010) have emphasized the importance of communication, which represents the one unifying theme that crosses all disciplines. Defined as a process by which meaning is conveyed in order to promote shared understanding, communication plays a central role in teaching and learning. The CCSS-M Standards for Mathematical Practices emphasize the critical role of classroom communication. For instance, they state that mathematically proficient students …

> justify their conclusions, communicate them to others, and respond to the arguments of others. They reason inductively about data, making plausible arguments that take into account the context from which the data arose. Mathematically proficient students are

> also able to compare the effectiveness of two plausible arguments, distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is. (NGACBP-CSSO, 2010, para. 4)

Improved classroom communication has been facilitated in recent years by the emergence of classroom connectivity technology (CCT), which refers to classroom wireless communication systems that connect student handhelds and teacher computers (Roschelle, Penuel, & Abrahamson, 2004). In this study, we describe findings from a three-year randomized control trial, *Classroom Connectivity in Promoting Mathematics and Science Achievement* (CCMS) (IESR305K050045) project, of the impact of CCT in Algebra I teaching and learning.

Communication, CCT, and Formative Assessment

Although most U.S. states employ accountability measures to track student achievement, these infrequent summative tests often fail to provide useful information for classroom teachers (Council of Chief State School Officers, 2000). Unlike summative assessment, formative assessment occurs daily as teaching and learning unfold. Popham (2008) defines formative assessment as "a planned process in which assessment-elicited evidence of students' status is used by teachers to adjust their ongoing instructional [practices] or by students to adjust their current learning tactics" (p. 6). In a meta-analysis, Black and Wiliam (1998) identified effect sizes from 0.4 to 0.7 in their analysis of 43 research studies that involved evidence of student learning and formative assessment practices.

From the perspective of the student, formative feedback is "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning" (Shute, 2008, p.154). By defining formative feedback in terms of improved student learning, feedback that does not contribute to improving learning is excluded from Shute's definition. Her comprehensive review of the formative feedback research identified student achievement gains ranging from 0.40 *SD* (Guzzo, Jette, & Katzell, 1985) to 0.80 *SD* (Azevedo & Bernard, 1995; Kluger & DeNisi, 1996) in controlled experiments.

Formative assessment by classroom teachers, however, is one of the weakest aspects of teacher practice (Assessment Reform Group, 1999; Daws & Singh, 1999). Effective feedback depends on three variables: (a) motive (student need); (b) opportunity (timeliness); and (c) means (student willingness to use the feedback) (Shute, 2008). The multiple demands on a teacher's time and attention during a typical lesson challenge teachers to gather accurate and timely data about student learning in order to make on-the-spot, data-driven decisions about classroom instruction.

Improving classroom communication by enhancing the speed and accuracy of information flow between students and teachers suggests the possibility of raising the quality of formative feedback (Shute, 2008). Developing shared understandings in the classroom depends on teacher skill to make concepts explicit (Bell & Cowie, 2001). When students and teachers differ in their understandings, teacher decision making is constrained by incomplete knowledge that potentially impedes teaching and learning progress. The ability of teachers to adapt instruction based on evidence of student understanding and learning needs embodies the heart of formative assessment practice (Wiliam & Thompson, 2008). Since CCT facilitates formative assessment and prior research has demonstrated that formative assessment practice has been linked to improved student achievement, the current study explores the impact of CCT in Algebra 1 classrooms on student achievement.

CCT as a Mediating Tool in Classroom Interactions

Stroup, Ares, & Hurford (2005) argue that CCT serves to fundamentally change student interactions with mathematics. Hegedus and Moreno-Armella (2009) suggest the term *representational expressivity,* to describe the transformation of traditional communication forms through the use of software that broadens the representational infrastructure of the classroom. CCT technology such as that used in this study provides multiple representations of mathematical objects as well as accurate and timely collection and aggregation of data/expressions contributed by students to the discourse space. Students engage in the public domain with mathematical ideas and collective critique, which supports a stronger sense of agency and identity as a mathematical thinker through student-initiated actions (Hegedus & Moreno-Armella, 2009; Hegedus & Penuel, 2008; Pape et al., 2013; Stroup et al., 2005).

CCT transfers responsibility for mathematical thinking to the students by facilitating public display of mathematical tasks that invites student participation and discourse and providing a forum for examination of one's ideas in comparison to others, timely and accurate formative feedback to teachers and students for classroom decision making, and the opportunity to learn by including most students in the classroom work through both technology and verbal communication pathways (Pape et al., 2013). In addition, White (2006) suggested that the ability for students to participate anonymously through the use of CCT effectively "broadens the 'bandwidth' of classroom collaboration" (p. 359) by masking student social status thus reducing potential impediments to classroom participation by some students. Students perceive that CCT facilitates an open and comfortable learning environment, enables the teacher and students to better assess student understanding and make use of feedback as will be discussed in the next section, and understand mathematics concepts more deeply (Herman, Meagher, Abrahamson & Owens, 2013; Owens, Demana, Abrahamson, Meagher, & Herman, 2004).

CONNECTED CLASSROOM TECHNOLOGY

The CCT utilized in this study, the Texas Instruments Navigator™, is a second-generation audience response system (ARS) technology. A discussion of the development of ARS technology may be found in Abrahamson and Brady (2014). The version of CCT employed in this study connects the teacher's computer with individual student's handheld graphing calculators such that four handhelds are wired to a transmission hub that communicates with the teacher's computer through a wireless access point (see Figure 1).

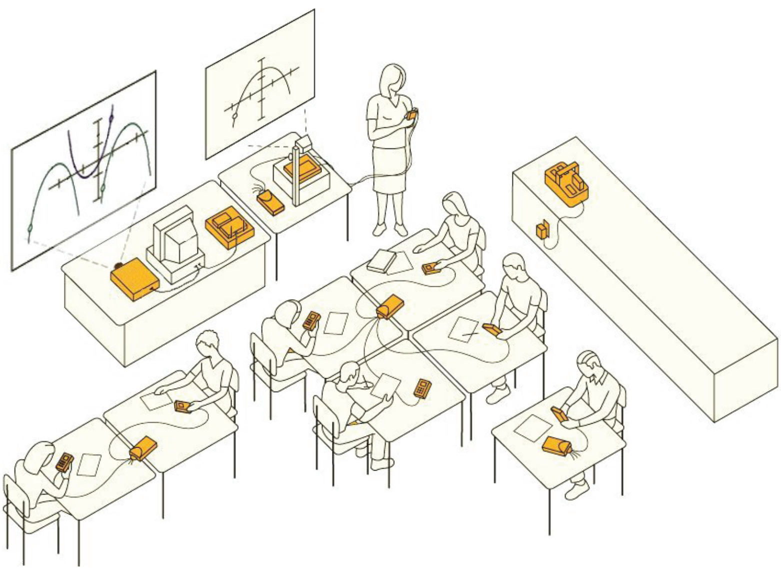

Figure 1. The Connected Classroom.

Simpler ARSs predominantly support the use of multiple-choice questions and are sometimes referred to as clicker systems. The more sophisticated CCT employed in this study, however, consists of four components: *Quick Poll, Learning Check, Screen Capture,* and *Activity Center. Quick Poll* and *Learning Check* allow teachers to send either individual or groups of questions to each handheld device for student response. The resulting data may be displayed publically for whole-class review, which provides formative assessment opportunities. Teachers may choose to reveal student names or provide anonymous displays. *Screen Capture* presents teachers with a 'snapshot' of individual student calculator screens for review and potential class display and/or discussion. Each of these components provides the teacher with data necessary to make instructional decisions and students receive immediate feedback in a supportive way that can encourage them to reflect and discuss their understanding or methods of solution in small groups and/or with the class as a whole (Roschelle, Penuel, & Abrahamson, 2004).

Using the Activity Center, a teacher can involve students in discovery lessons through display and interaction with a coordinate system. Students may submit individual points, equations for lines, or data lists to a shared workplace. Simultaneous display of multiple mathematical representations (equations, graphs, data tables) creates opportunities for rich mathematical discourse and supports the design and implementation of inquiry lessons. The richer discourse and display of students' mathematical thinking is hypothesized to support greater formative assessment opportunities for teachers and students.

During the professional development, participating teachers were introduced to the potential of CCT to improve student achievement through two mechanisms. First, CCT can serve as a mediating tool between mathematical content and classroom culture (Stroup, Ares, & Hurford, 2005). Increasing the interaction of students with mathematical content opens the door to improved opportunities for learning (Gee, 2008). Second, CCT enhances teacher opportunities for formative assessment and increases the potential for teachers to gain deeper understanding of student learning during classroom instruction.

The Present Study

The current study presents findings from the first three years of a longitudinal random control trial and explores whether the first year findings

(Pape et al., 2013) were replicated in successive years. In the first intervention year, teachers in the treatment groups attended professional development and then used CCT in their classroom teaching while the control group teachers used graphing calculators only. In the summer before the second year of the study, the control group teachers participated in the professional development and implemented the technology within their classrooms during year two of the study. Analysis of the first year data from this study showed an increase in student achievement in Algebra 1 with an effect size of 0.30, a relative rare finding for randomized educational experiments (Pape et al., 2013).

The present study seeks to follow-up on the first year of data by exploring the following research question: What is the impact of CCT on Algebra 1 achievement across three years of implementation? Both cross-sectional and longitudinal analyses were conducted. Cross-sectional models examine achievement within treatment classrooms across three years versus control classrooms; longitudinal models examine one group of teachers comparing their students' control classroom achievement versus achievement in these same teachers' classrooms for the next two years following implementation of the CCT intervention.

METHOD

Research Design

The research design for this study is a randomized cross-over trial where the control group was exposed to the intervention in the second and third years of the study. Cohort 1 treatment group participants were trained to use the CCT during a one-week, residential summer institute and participated in follow-up professional development at an international technology conference in February/March each year they participated in the study (see Pape et al., 2013 for more details regarding the summer professional development). During year 1 of the study, Cohort 2 teachers used graphing calculators with their students and completed tests and surveys. The first year data from Cohort 2 provide the control year data for comparison with treatment groups. Cohort 2 teachers participated in similar professional development activities starting in the summer before their second year of participation in the study. While the participating teachers remained the same throughout the three-year study, they worked with a new set of Algebra I students each academic year.

The following convention will be used to distinguish between the two groups of teachers who participated in the study. Cohort 1 participated in professional development and implemented the CCT in the first year of the study (Treatment 1.1) and continued to work with the technology for two additional years, which are denoted as Treatment 1.2 and Treatment 1.3. Cohort 2 teachers served as a control group during their first year of participation (Control 2.1). In the second and third years of the study, Cohort 2 teachers implemented CCT in their classrooms, which we denote as Treatment 2.2 and Treatment 2.3, respectively. Data from Cohort 2 during the first academic year of the study (Control 2.1) serves as control data for both cohorts.

Participant Recruitment and Selection

A list serve of teachers using graphing calculators and personal solicitation by project personnel at professional conferences were used to recruit teachers for participation in this study. Interested teachers completed an application and were screened for prior use of graphing calculators in their teaching practice. Administrators from participating schools were asked for a formal statement of support for the project. Only those applicants who verified proficiency with graphing calculator use and with support from an administrator were selected into the study. A random number generator was used to assign participants to either treatment (Cohort 1) or control group (Cohort 2). Teacher volunteers from the same school building were selected into the same cohort to prevent potential contamination effects. Schools provided participating teachers with compatible laptops, handheld graphing calculators, and release time to attend the winter professional development conferences. The grant supported participating teachers to attend a one-week residential summer professional development experience the summer before they entered the treatment condition and paid for participants' travel to an international technology conference yearly.

Participants

Algebra 1 teachers from 28 U.S. states and 2 Canadian provinces and their students participated in this study. At the outset of the project, a total of 127 teachers (66 control, 61 treatment; 74.0% female) were randomly assigned to either control or treatment groups. From this initial sample, 19 participants failed to complete the first year of the project due to a variety of reasons: incorrect or inadequate human subject's compliance (n = 2); lack of administration support (n = 4); inappropriate teaching assignments (n = 2); quit teaching (n = 2); failed to respond after initial training (n = 3); and health or personal reasons (n = 6). Data from teachers in Canada were not analyzed due to curriculum differences with U.S. schools. In addition, each year some teacher participants neglected to submit data and some data sets were lost in transit from the school sites to the data recording center (e.g., ripped and empty envelopes arrived). Table 1 indicates the number of teacher participants who submitted pretest and posttest Algebra achievement data for analysis and details regarding racial and gender characteristics as well as their teacher preparation and years of teaching experience.

Table 1

Teacher Demographic Data

Group	Years of CCT use	N*	% White	% Female	% Math degree	Years Teaching Experience Mean(Median)	% Free/ Reduced Lunch (Median)☛☛
Control 2.1	0	41	87.8	76.7	79.1	15(15)	22
Treatment 2.2	1	32	84.4	80.0	77.1	10(17)	22
Treatment 2.3	2	28	89.3	76.7	73.3	10(13)	15
Treatment 1.1	1	39	100	74.4	61.5	7(12)	9
Treatment 1.2	2	29	100	73.5	61.8	8(11)	9
Treatment 1.3	3	19	100	77.3	59.1	14(12)	9

*Number of teachers who reported demographic data
☛☛This column uses school averages as a proxy for classroom composition and should be interpreted with caution

Teacher participants identified themselves as white (84% to 100% throughout the study years), female (74-80%) and reported holding degrees in mathematics (59-80%). Across the cohorts, years of teaching experience ranged from 1 to 36 years with mean years of teaching experience between 7 and 15 years and median years of experience from 11 to 17 years (Table 1).

Students enrolled in Algebra I sections taught by participating teachers were invited to participate each year across the three-year study. Student sample sizes varied from a low of 532 (Treatment 2.2) to a high of 696 (Treatment 2.3) (Table 2). On average, about 60-75% of the student-participants identified themselves as White, and 43-57% of the student participants identified as female.

Table 2

Teacher and Student Participants, Years 1-3, All Cohorts

	Cohort 1		Cohort 2		Total	
Year	Teachers	Students	Teachers	Students	Teachers	Students
1	39	654	43	570	82	1224
2	34	560	35	532	69	1092
3	22	577	30	696	52	1273
Totals		1791		1798		3589

Data Sources

Student-level measures. Students completed an *Algebra Pretest* (National Center for Research on Evaluation, Standards, and Student Testing (CRESST), 2004), *Algebra Posttest* (Abrahamson et al., 2006), and *Student Views about Mathematics* (Pape, Kaya, Owens, Irving, & Boscardin, 2006) survey. Data from the Algebra Pretest and Posttest only will be analyzed for this study.

Algebra Pretest. The Algebra Pretest was administered at the beginning of the school year to measure initial achievement levels of participating students. The pretest measure included 32 pre-algebra and algebra multiple choice, short-answer, and extended constructed-response format items (α = 0.81; CRESST, 2004). This previously validated pretest was based on released items from the National Assessment of Educational Progress (NAEP) and the California Standards Test. Three-parameter logistic (3PL) Item Response Theory (IRT) analysis using the statistical package BILOG MG was conducted to verify the technical quality of the instrument. These data served as a covariate in some models included in our results.

Algebra Posttest. The development of the Algebra I Posttest (Abrahamson et al., 2006) began by comparing mathematics content standards of 13 states (e.g., Ohio, Texas, New York, Virginia) from which a majority of CCMS participants were drawn. The IRT analysis (3PL using BILOG MG) conducted to ensure the technical quality of the measure resulted in the exclusion of five items. Thirty-five questions aligned with these standards were selected from released items from California and Virginia mathematics tests, and from the Trends in International Mathematics and Science Study (TIMSS) assessment and NAEP released items. The final instrument included 24 multiple-choice items, 5 extended-response items, and 1 three-part short-answer question (α = 0.85). Eleven items were common to the pre-test and posttest.

Teacher Demographic Survey. During the first year of the study, teachers completed a demographic survey intended to gather information about their teacher preparation (degree types), years of classroom teaching, teaching assignments, teaching licensure, racial demographics, and gender.

PROCEDURE

Data Analysis

Algebra Open-Response Scoring. The Algebra Pretest and Posttest contained seven and six open-ended items, respectively, five of which occurred as identical items on both instruments. Scoring of these items was conducted by teams of raters who had been trained using a detailed protocol developed by members of the research team. Inter-rater reliability was calculated on 10% of the scored papers that were randomly selected and independently scored by a graduate student who was trained on the scoring rubric. Inter-rater reliability ranged between 0.88 and 0.98 demonstrating a high degree of consistency in the scoring of open-ended items.

Statistical Analyses. Two-level hierarchical linear modeling (HLM) was employed to model the impact of the intervention for students while accounting for the nested structure of the data (i.e., students nested within teachers' classrooms). This data analytic technique reflected the fact that students were not statistically independent of each other, but rather were situated within classrooms, and that HLM offers the opportunity for within class analysis (level 1) as well as between class analyses (level 2). Standardized effect sizes**,** δ, were estimated according to the formula suggested by Raudenbush, Martinez, and Spybrook (2007). As is the case with traditional effect size statistics (e.g., Cohen's *d*) multilevel effect sizes are estimates of population differences between treatment and control, divided by the standard error of the outcome. The only difference in the multilevel (cluster-randomized design) framework, is that total variance is the sum of the level 1 and level 2 variance, which is a pooled standard deviation.

$$\delta = \frac{\gamma_1}{\sqrt{\tau^2 + \sigma^2}}$$

Where $\gamma_1 = \mu_E - \mu_C$
μ_E is the population mean for the experimental group
μ_C is the population mean for the control group

This report includes multilevel comparisons of six groups of classrooms: Control (cohort 2 Year 1); Treatment 2.2, Treatment 2.3, Treatment 1.Treatment 1.2, and Treatment 1.3. Treatment groups 1.1 to 1.3 are analyzed in three separate cross-sectional models, each using the same control.

LEVEL 1 MODEL

$$\text{POSTTOTA} = \beta_0 + \beta_1 \, (\textit{PRETOTAL}) + \tau$$

LEVEL 2 MODEL

$$\beta_0 = \gamma_{00} + \gamma_{01} \, (\text{TREAT}) + \gamma_{02} \, (\text{Y2YTEACH}) + u_0$$
$$\beta_1 = \gamma_{10} + u_1$$

POSTTOTA= posttest algebra score

PRETOTAL = pretest score for student *i* in class *j* centered around the grand mean

TREAT is an indicator variable that takes on a value of 1 for treatment group and 0 for control group.

Y2YTEACH is teaching experience in years

Treatment groups 2.2 and 2.3 are analyzed with two longitudinal models in which they serve as their own Year 1 controls.

LEVEL 1 MODEL

$$\text{POSTTOTA} = \beta_0 + \beta_1 \, (\textit{PRETOTAL}) + \beta_2 (\text{TREAT}) + \tau$$

LEVEL 2 MODEL

$$\beta_0 = \gamma_{00} + u_0$$
$$\beta_1 = \gamma_{10} + u_1$$
$$\beta_2 = \gamma_{20} + \gamma_{21} + (\text{Y2YTEACH}) + u_2$$

All models were estimated once with student pretest scores included as a covariate and once without pretest scores. The rationale for excluding the pretest is that pretest administration dates were inconsistent, and late pretesting may have led to inflated pretest scores; hence, entering the pretest as a covariate may have depressed effect sizes in the second and third year of the study. Because of the RCT design of the study, pretest control should not be required to determine accurate estimates of treatment effects.

Models that control for pretest are labeled below with “p”; those that do not control for pretest are labeled with “np.” In all models, only teachers who participated in both year 1 and in the comparison year were included for analysis. Teachers with incomplete data were not included in the analysis.

RESULTS

Since the year 1 findings for Treatment 1.1 showed increased algebra achievement in the CCT classrooms (Pape et al., 2013), both cross sectional and longitudinal analyses were completed on the year 2 and 3 data sets. Table 3 provides a summary of Algebra achievement comparisons over the three years of the study[1].

Cross-Sectional Analysis

Model 1p (Table 3) used HLM analysis to compare student outcomes for Treatment 1.1 and Control 2.1 groups while accounting for years of teaching experience (*Y2YTEACH*) and pre-test score *(PRETOTAL)*. The treatment group significantly outperformed the control group (p = .034; ES = 0.30) by approximately 1.9 points ($\gamma 01$ = 1.86) even after adjusting for pretest scores and teacher experience. Years of teaching experience approached significance in this model ($\gamma 02$ = 0.11, p= .076). This finding was reported in a previous publication (Pape et al., 2013) and is re-reported here for comparison with the later findings.

Model 2p examined differences in student outcomes for Treatment 1.2 versus Control 2.1 groups while taking into account years teaching experience *(Y2YTEACH)* and pre-test score *(PRETOTAL)*. The treatment group did not significantly outperform the control group although a modest effect size is consistent with the effect found in the other models (p = .154; ES = 0.23). Treatment groups scored about 1.3 points higher than control groups ($\gamma 01$ = 1.27). Years of teaching was a significant covariate in this model ($\gamma 02$ = 0.14, p = .020).

1 Detailed statistical analyses for each of the models separately may be obtained from the corresponding author, Karen Irving (Irving.8@osu.edu).

Table 3

Summary of Algebra Achievement Comparisons Over Three Years

	Groups compared	Pretest Controlled	Effect size	*p*-value
Cross sectional Models				
Model 1p	Treatment 1.1 versus Control 2.1	Yes	0.30	.034
Model 2p	Treatment 1.2 versus Control 2.1	Yes	0.23	.154
Model 2np	Treatment 1.2 versus Control 2.1		0.36	.044
Model 3p	Treatment 1.3 versus Control 2.1	Yes	0.24	.198
Model 3np	Treatment 1.3 versus Control 2.1		0.37	.080
Longitudinal Models				
Model 4p	Treatment 2.2 versus Control 2.1	Yes	0.20	.009
Model 4np	Treatment 2.2 versus Control 2.1		0.20	.004
Model 5p	Treatment 2.3 versus Control 2.1	Yes	0.13	.190
Model 5np	Treatment 2.3 versus Control 2.1		0.19	.046

Model 2np compared Treatment 1.2 and Control 2.1 groups controlling for years of teaching experience *(Y2YTEACH)* only. The treatment group significantly outperformed the control group (p = .044; ES = 0.36) by approximately 2.6 points in this model ($\gamma 01$ = 2.59). As in the previous model, years of teaching was a significant covariate in this model ($\gamma 02$ = 0.22, p = .018).

Models 3p and 3np compare data collected during the third year of the study for Cohort 1, Treatment 1.3, versus control data from the first year of the study (Control 2.1). Although the treatment in Model 3p was not significant ($\gamma 01$ = 1.34, p = *.198*), when pretest scores and years of teaching experience were accounted for in the model, the comparison produced a modest effect size (ES = 0.24). In Model 3np contrasting Cohort 1 teachers in their third year in connected classrooms with the control group when the pretest score was not included in the model, the effect of treatment approached significance (p = .080) with the treatment group students outperforming the control group by about 2.65 points. The effect size associated with Model 3np (ES = 0.37) was stronger than for Model 3p in which the pretest was included as a covariate. Years of teaching experience was a significant covariate when pretest was included as a covariate ($\gamma 02$ = 0.12, p = .043), but only approached significance when pretest was not accounted for in the model ($\gamma 02$ = 0.14, p = .074).

Longitudinal Analyses of Cohort 2 Teachers

The longitudinal models indicate fairly consistent effects (~*ES* = 0.20), with significant treatment effects (Table 3). Models 4p and 4np compare data for Cohort 2 using Year1 data as baseline (Control 2.1) and Year 2 data as treatment (Treatment 2.2). When controlling for the pretest and years of teaching experience, the treatment group shows significant improvement ($\gamma 20 = 1.08$, $p = .009$) with a moderate effect size (*ES* = 0.20). Students in the experimental group scored about 1.1 points higher than their control group counterparts when controlling for these two covariates. When the pretest is removed as a covariate, the comparison still produces a significant difference between treatment and control ($\gamma 10 = 1.35$, $p = .004$) with an effect size of 0.20. In this model, the treatment group students scored about 1.3 points higher than the control group. Years of teaching experience was a non-significant covariate when the pretest was included as a covariate in Model 4p ($\gamma 21 = 0.04$, $p = .205$) but was a significant covariate when the pretest was excluded as a covariate in Model 4np ($\gamma 11 = 0.07$, $p = .047$).

Longitudinal models 5p and 5np compare data for Cohort 2 using Year1 data as baseline (Control 2.1) and treatment data from Year 3 (Treatment 2.3). When controlling for the pretest and years of teaching experience, the treatment group showed no significant improvement ($\gamma 20 = 0.69$, $p = .190$) with a very small effect size (*ES* = 0.13). When the pretest was removed as a covariate, the comparison produced a significant difference between treatment and control ($\gamma 10 = 1.30$, $p = .046$) and an effect size of 0.19. In this model, the treatment group students scored about 1.3 points higher than the control group. Years of teaching experience was not a significant covariate in either Model 5p or Model 5np ($\gamma 21 = 0.07$, $p = .185$ and $\gamma 11 = 0.11$, $p = .123$, respectively).

DISCUSSION AND CONCLUSIONS

In a national randomized control trial, both the cross-sectional comparisons of Cohort 1 with the control group and the longitudinal comparisons for Cohort 2 across the three years showed that the treatment groups outperformed the control groups in algebra achievement. Five of the nine models resulted in significant differences between treatment and control groups with a sixth model nearing significance (Table 3). The results from HLM models were largely consistent across years and cohorts and support a conclusion that the previously reported results from year 1 are applicable across the three years of this study (Pape et al., 2013). All five treatment groups outperformed the control group on the algebra posttest in at least one of the two modeling frameworks (Table 3). Effect sizes for the comparisons ranged from 0.19 to 0.37.

Evidence from teacher participants indicated that the pretest was administered later than expected in the second and third year of the project, which suggests that the pretest scores in those years might be inflated. If teachers delayed administering the pretest until after several weeks of instruction, student scores would not provide reliable baseline data for comparison. As a result, models with pretest as a covariate and without pretest as a covariate are presented.

Years of teaching experience was a consistently significant covariate in the study with an impact of between one additional posttest point for every five to ten years of teaching experience. This puts the connected classroom effect in a favorable light: the impact of the intervention is equivalent to the observed impact of approximately ten to fifteen years of teacher experience in teaching Algebra I. Other covariates that were explored but showed inconsistent or non-significant effects included classroom composition variables such as ethnicity and grade level of students; teacher gender and degree; baseline classroom achievement; and school location and SES. None of these variables contributed significantly to any of the models tested after controlling for teacher experience. Teacher attrition across the program may attenuate the significance of the teacher experience effects in some of the longitudinal models, but the observed effects are similar in magnitude to the effects evidenced in the cross-sectional models, even when non-significant. Results from these models imply that teacher experience interacts with the treatment effect, rather than influencing students in all classrooms equally.

The steady effect sizes from .20 to .37 are considered medium, a relatively rare finding for a large, national RCT (Bloom, Hill, Black & Lipsey, 2008). The implementation of random assignment with a true control group provides strong evidence to support the inference that the use of CCT by treatment teachers caused the increased algebra achievement. Note that we argue that the intervention includes *both* the CCT *as well as* the teacher professional development, which included a summer institute and follow-up in addition to yearly attendance at an annual technology-related conference. Simply providing equipment to teachers has not been demonstrated to make a difference in teaching or learning. Our argument posits that the CCT increases the opportunity to learn by serving as a mediating tool to produce a classroom environment that supports student examination and analysis of patterns as well as supporting collaborative work and justification of mathematical generalizations. In addition, CCT facilitates students' and teachers' immediate feedback in a public, supportive forum that may restructure the difficult task of formative assessment and may promote productive mathematical discourse in a classroom community. Additional evidence from close analysis of classroom conversation in CCT classrooms when compared to patterns in non-CCT classrooms indicates that some disruption of the traditional Initial-Response-Evaluate discourse pattern of classroom communication occurs in the CCT classrooms (Pape et al., 2010).

Educational stakeholders within the United States and around the world recognize the importance of increasing Algebra achievement (USEOP, 2014). In contrast to this study, another reported large and expensive experimental study of multiple educational technologies found no significant effects (Dynarski et al., 2007). Within this context, the present experiment is noteworthy due to the use of randomization to allow valid causal inferences and a suitable sample size to detect effects. The study took place over three full years of instruction under realistic conditions in a variety of schools in the United States. The CCT used in the study is relatively affordable, broadly scalable and based on graphing calculators, a technology already in use in many U.S. mathematics and science classrooms. We hypothesize the possibility of a mechanism explaining these gains based on CCT-facilitated active learning environments that broaden the representational infrastructure of the classroom (Pape et al., 2013) as well as the improved formative assessment practice possible in connected classrooms (Irving, Sanalan & Shirley, 2009) and changes in classroom discourse patterns (Pape, et al., 2010).

LIMITATIONS AND FUTURE WORK

The teacher participants in the current study represent a unique set of individuals, with richer mathematics background knowledge and stronger initial graphing calculator skills than might be expected of Algebra I teachers. While participants were expressly chosen with strong graphing calculator skills, a degree in mathematics was not a requirement for participation in the study. Thus the findings of this longitudinal random control study may be generalized to a similar group of Algebra I teachers rather than Algebra I teachers in US schools writ large. As might be expected in a longitudinal study of this magnitude, teacher attrition may have confounded the results of the study. Teachers who persisted with the study may have been those with stronger mathematics content knowledge or those who were more comfortable and expert at using technology. A replication of the study with a broader sample of Algebra I teachers who possess less technology experience and greater ethnic diversity is needed. Additional study of the rich qualitative data set collected during the implementation of this project will shed greater light on the mechanisms involved in classroom practice within a connected classroom that may lead to improved algebra achievement.

Acknowledgments

The authors wish to thank Prof. Frank Demana for his freely given time throughout the study, and for insight, enormous experience, and patience he brought to bear in helping us configure and monitor the teacher professional development, which took place within this project. Also, our grateful thanks to Dr. Jeremy Roschelle of SRI International for his invaluable ongoing advice, perspectives and suggestions, and Dr. Joan Herman of CRESST at the University of California Los Angeles for her help and wisdom in structuring this project.

The authors wish to acknowledge the generous contributions Texas Instruments to supporting the professional development efforts and for providing the Navigator systems to school districts at a reduced cost.

The research reported here was supported by the Institute of Education Sciences, U. S. Department of Education, through Grant R305K050045 to The Ohio State University. The opinions expressed are those of the authors and do not represent views of the U. S. Department of Education.

References

Abrahamson, A. L., & Brady, C. E. (2014), A brief history of networked classrooms to 2013 - Effects, cases, pedagogy, and implications with new developments. *International Journal of Quality Assurance in Engineering and Technology Education, 3*(3), 1-54.

Abrahamson, A. L., Sanalan, V., Owens, D. T., Pape, S. J., Boscardin, C., & Demana F. (2006). *Classroom Connectivity in Mathematics and Science Achievement: Algebra I Posttest.* Columbus, OH: The Ohio State University.

Assessment Reform Group (1999). *Assessment for learning: Beyond the black box.* Cambridge: University of Cambridge, School of Education.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer based instruction. *Journal of Educational Computing Research, 13*, 109-125.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*, 536-553.

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: King's College London. Retrieved from http://www.spd.dcu.ie/site/teaching_today/documents/raisingstandardsthroughclassroomassessment.pdf

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1,* 289-328.

Council of Chief State School Officers. (2000). State Educational Assessment Center, State student assessment programs database, 1997-1998 school year. Retrieved from https://nces.ed.gov/programs/statereform/aas.asp

Daws N., & Singh, B. (1999). Formative assessment strategies in secondary science. *School Science Review* 80, 71–78.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Gee, J. P. (2008). A sociocultural perspective on opportunity to learn. In P. A. Moss, D. C. Pullin, J. P. Gee, E. H. Haertel, & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn* (pp. 76-108). New York, NY: Cambridge University Press.

Gunipar, Y., & Pape, S. J. (2016). *Teachers' instructional practices within connected classroom technology environment to support representational fluency*. Manuscript submitted for publication.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology, 38*, 275-291.

Hegedus, S. J., & Moreno-Armella, L. (2009). Intersecting representation and communication infrastructures. *ZDM Mathematics Education, 41,* 399-412. DOI: 10.1007/s11858-009-0191-7

Hegedus, S. J., & Penuel, W. R. (2008). Studying new forms of participation and identity in mathematics classrooms with integrated communication and representational infrastructures. *Educational Studies in Mathematics, 68,* 171-183. DOI 10.1007/s10649=008-9120-x

Herman, M., Meagher, M., Abrahamson, L., & Owens, D. (2013). Student perceptions on use of a classroom communications system in mathematics classes. *International Journal of Technology in Mathematics Education 20*(2), 45-67.

Irving, K. E., Sanalan, V. A, & Shirley, M. L. (2009). Physical science connected classrooms: case studies. *Journal of Computers in Mathematics and Science Teaching, 28,* 247-275.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254-284

Munsen, L. (2010). *Pre-algebra and algebra enrollment and achievement*. Providence, RI: Annenberg Institute for School Reform at Brown University. Retrieved from http://files.eric.ed.gov/fulltext/ED533119.pdf

National Center for Research on Evaluation, Standards, and Student Testing (CRESST) (2004). *CRESST Algebra I Test*. Los Angeles, CA: Author.

National Council of Teachers of Mathematics [NCTM] (2014*). Principles and standards for school mathematics: Communication.* Retrieved from http://www.nctm.org/standards/content.aspx?id=26854,

National Mathematics Advisory Panel (NMAP). (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: US Department of Education.

United States. Executive Office of the President USEOP (2014). *Increasing college opportunities for low income students: Promising models and a call to action.* Retrieved from white_house_report_on_increasing_college_opportunity_for_low-income_students_1-16-2014_final.pdf

Owens , D. T., Demana , F., Abrahamson, A. L., Meagher, M., & Herman, M. (2004). *Developing Pedagogy for Wireless Calculator Networks--and Researching Teacher Professional Development. Final Report. Part 2--Technical Report and Research Description to the National Science Foundation* (National Science Foundation Grant Report by The Ohio State Univ. Res. Fdtn. & Better Education Inc. No. ESI 01-23391 & ESI 01-23284). Washington, DC: (ERIC Document Reproduction Service No. ED 479499).

Pape, S. J., Irving, K. E., Owens, D. T., Boscardin, C. K., Sanalan, V. A., Abrahamson, A. L., Kaya, S. Shin, H. S., & Silver, D. (2013). Classroom connectivity in Algebra I classrooms: Results of a randomized control trial. *Effective Education, 4*(2), 169-189. Doi 10.1080/19415532.2013.841059

Pape, S. J., Bell, C. V., Owens, S. K., Bostic, J. D., Irving, K. E., Owens, D. T., Abrahamson, L., & Silver, D. (2010, May). *Classroom interactions within connected mathematics classrooms.* A paper presented at the 2010 Annual meeting of the American Educational Research Association, Denver, CO.

Pape, S. J., Kaya, S., Owens, D. T., Irving, K. E., & Boscardin, C. (2006). *Classroom connectivity in mathematics and science achievement: Student views of mathematics*. Columbus, OH: The Ohio State University.

Popham, W. J. (2008). *Transformative assessment.* Alexandria VA: Association for Supervision and Curriculum Development (ASCD).

RAND Mathematics Study Panel, & Ball, D. L. (2003). *Mathematical proficiency for all student: Toward a strategic research and development program in mathematics education.* Santa Monica, CA: RAND

Raudenbush, S. W., Martinez, A, & Spybrook J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, 29, 5 - 29.

Roschelle, J., Penuel, W. R., & Abrahamson, L. (2004). The networked classroom. *Educational Leadership, 61*(5), 50-54.

Sert, Y. (2014). *The effects of using connected classroom technology on representational fluency*. Gainesville, FL: University of Florida.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*, 153-189.

Snipes, J. & Finklestein, N. (2015). *Opening a gateway to college access: Algebra at the right time*. San Francisco: REL West @ WestEd. Retrieved from https://relwest.wested.org/system/resources/196/REL_ResearchBrief_AlgebraReadiness(sm).pdf?1436833679

Stroup, W. M., Ares, N. M., & Hurford, A. C. (2005). A dialectic analysis of generativity: Issues of network-supported design in mathematics and science. *Mathematical Thinking and Learning,7*(3), 181-206.

White, T. (2006). Code talk: Student discourse and participation with networked handhelds. *Computer-Supported Collaborative Learning, I,* 359-382. Doi:10.1007/s11412-006-9658-5.

Wiliam, D., & Thompson, M., (2008). Integrating Assessment with Learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning*, (pp. 53-82). New York: Lawrence Erlbaum.