Empirical Education

RESEARCH REPORT

The Effect of Ongoing Professional Development on Interactive Whiteboard Use:

A Report of a Randomized Experiment in Forsyth County Schools

Jessica V. Cabalo

Boya Ma

Minh-Thien Vu

Andrew Jaciw

Gloria I. Miller

Empirical Education Inc.

March 1, 2007

Empirical Education Inc.

www.empiricaleducation.com

425 Sherman Avenue, Suite 210

Palo Alto, CA 94306

(650) 328-1734

Acknowledgements

We are grateful to the people at the Forsyth County Schools for their interest and assistance in conducting this research. Logical Choice Technologies has collaborated with the Forsyth County Schools on providing all the training materials. The research was funded by a grant (#R305E040031) to Empirical Education Inc. from the US Department of Education. The purpose of this grant is to improve our ability to run small scale experiments to assist local decision-makers. The US ED is not responsible for the content of this report.

About Empirical Education Inc.

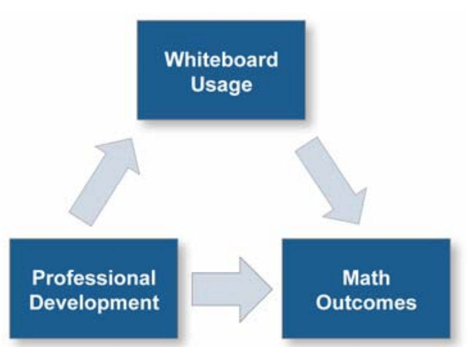
Empirical Education Inc. was founded to help school districts, publishers, and the R&D community assess new or proposed instructional and professional development programs through scientifically based pilot implementations. The company draws on the expertise of world-class researchers and methodologists assuring that the research is objective and takes advantage of current best practice in rigorous experimental design and statistical analysis. The company's findings let educators quantify the value of programs and help them partner with providers to implement those most effective for their students.

Executive Summary

Introduction. The Forsyth County School District sought scientifically based evidence for the effect of ongoing, district-developed professional development (*PD*) on the usage of Promethean Interactive Whiteboards. A randomly selected subset of teachers received *PD* in addition to basic training sessions offered in Fall 2005 to all district teachers for using interactive whiteboards (IWBs). District decision-makers were interested in learning about the impact of the additional *PD* on student achievement outcomes in core subjects as well as on teacher classroom practices.

The district, considered part of metro Atlanta, GA has an annual growth rate of 8-10% and serves mainly White students. The *PD* is an ongoing, systematic training structured to include 1- to 2-hour workshops and supplemental resources from the school IT specialist. The objective was to train teachers to perform advanced IWB operations to enhance their instruction and engage their students in learning. We conducted an experiment comparing outcomes for classes taught by teachers randomly assigned to receive the *PD* and classes taught by teachers assigned to only the basic training on IWBs that all teachers received. Interviews, surveys, and observations allowed us to characterize the use of IWBs both qualitatively and quantitatively.

Findings. We found that students of the teachers who received the *PD* had lower scores on Georgia's state assessment (CRCT) in Math than students of the other teachers. We also found a negative impact on the amount of IWB usage. These unexpected results led us to additional exploratory analyses and considerations. The figure below shows the analysis approach. To understand the mechanism through which the *PD* resulted in both lower IWB usage and lower Math scores, we investigated whether the difference in IWB usage was associated with the Math outcomes. Here we found some indication of a positive relationship. It therefore became very important to understand how it was possible for additional *PD* to have a negative effect on the two outcomes.



We considered two different phenomena. First, the control students' better performance could perhaps be attributed to non-*PD* teachers' eagerness to compensate for not having the additional *PD*, making a greater effort to obtain information from other sources in the school.

The second is what researchers call "contamination" of the control group that resulted from *PD* teachers making all the materials to support whiteboard use available on the school's computer network (the "shared drive"). Contamination is unfortunate from a research standpoint because it eliminates the differences between treatment and control groups. On the other hand, the sharing of resources is consistent with the culture of cooperation within schools, which appears to be an effective tool for multiplying the impact of the *PD* and spreading information and expertise, allowing teachers to maximally utilize their resources within their schools.

Although our goal in this research was to provide the district with evidence that would be useful in determining the impact of their *PD* program, we found that the value of the *PD* appeared to be eclipsed by the culture of sharing in the schools. Insofar as 1) the *PD* assisted the schools in developing IWB resources for shared use, and 2) the resources helped to increase overall usage for all teachers, the *PD* can be considered a success. The suggestive finding from the exploratory analysis that whiteboard use may have an impact on achievement remains to be confirmed with properly controlled comparisons.

Our sample of teachers was small (15) and for the analysis involving test scores, the study could not use seven of the teachers in grades where no state test results were available. Because few teachers participated, there is a danger of bias being introduced by, for example, more enthusiastic teachers falling by chance into the *PD* or control group. With a larger pool of teachers, the likelihood of this bias occurring would have been lower. We were able to eliminate technology expertise and interest as potential sources of bias, but others may have been present.

Analysis. Statistical analyses involving CRCT scores were based on eight teachers/classes (four *PD* and four control) and 125 students in grades 3 to 5. We investigated three separate outcomes.

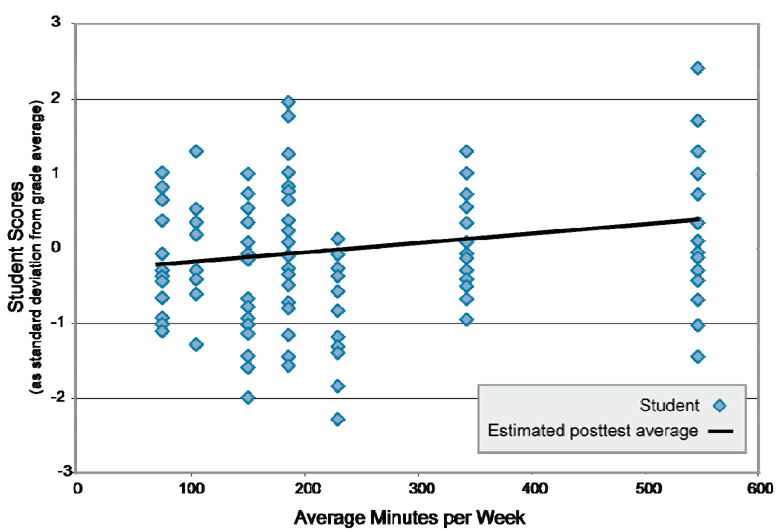
1. The primary topic of our experiment was the impact of additional professional development. We first looked at the impact of *PD* on the CRCT. We found negative effect sizes in English Language Arts, Reading, and Math, indicating that the students in control classes performed better than their counterparts in *PD* classes. (Note that nearly all students, regardless of condition, improved on the CRCT in Reading.)

2. Second, we examined whether teachers given the additional *PD* used IWBs more or less than control teachers. We found a statistical difference between *PD* and control teachers in IWB usage, control teachers using them considerably more than *PD* teachers. As noted, we attribute this to the culture of sharing often found in successful schools. Surveys revealed that *PD* teachers made their materials available to control teachers through their schools' shared drive. The table shows that, while *PD* teachers shared more of their materials, control teachers accessed these materials more than *PD* teachers.

3. Third, we asked whether the amount of IWB usage was related to the student outcomes, in particular in Math. If there were an impact of *PD* on student outcomes, we wanted to understand the extent to which that impact was mediated by IWB usage. The significant differences we found in student performance appear to be related to the degree of IWB usage. A test of the correlation reveals a positive relationship between IWB usage and student outcome. This is evident in the graph, which shows a positive slope in the predicted relationship between the z-transformed scores and IWB usage. The *p* value for this effect is .14, which gives us some confidence that the true slope is in fact different from zero.

Differences in Access to Shared Drive Between *PD* and Control Groups

Use of "shared drive"	*PD* teachers	Control teachers
Share self-created flipcharts	40%	25%
Access flipcharts others created	32%	38%
Other	28%	37%



Overall Teacher Impressions. Surveys and interviews revealed overall satisfaction with IWBs and increased student interest and engagement compared to the previous year before their introduction. One teacher referred to the IWB as a "magnet" that "gets the focus of the whole class. All eyes are on the whiteboard." When asked how they had their students use the IWBs, *PD* group teachers said their students used them more for basic presentation than for critical thinking activities, whereas the opposite was true for control teachers. Teachers in both groups, however, expressed frustration over the lack of time for planning and practicing with their IWBs and indicated that integrating such technology into their classroom instruction would be a lengthy process.

Effect of On-Going Professional Development on Interactive Whiteboard Use:

A Report of a Randomized Experiment in Forsyth County Schools

Table of Contents

Introduction

The Forsyth County Schools and Empirical Education Inc. had undertaken research to produce scientifically based evidence of the effectiveness of professional development on the use of Interactive Whiteboards (IWB). The question specifically addressed is whether the ongoing, district-developed professional development for the usage of Promethean Interactive Whiteboards is more effective than the basic training on their use received at the beginning of the year. In this final report we will describe the impact of the ongoing professional development (*PD*) on student achievement outcomes in English Language Arts, Reading and Math as well as on teacher classroom practices based on surveys, observations and a debrief session.

The study consists of a randomized experiment or randomized controlled trial (RCT) at the Forsyth County School District. The research focuses on 322 K-5 students in 15 classrooms, comparing the student achievement of two groups of classrooms randomly assigned to two conditions. The pilot group consists of the teachers receiving ongoing professional development (*PD)*, whereas the control group consists of the teachers who had only received the initial training at the beginning of the school year. The district was particularly interested in the following questions.

- Does an ongoing *PD* model (for IWBs) lead to improved instructional practice in the classroom? If so, *how* does this ongoing systemic training affect the kinds of teaching that goes on in the classroom?
- Do the teachers perceive the training to be beneficial to their teaching?
- Do students seem more engaged through the use of the interactive whiteboard?
- Does the addition of *PD* impact student achievement?

Methods

Research Design

This randomized study is a comparison of outcomes for classes being taught by teachers receiving the *PD* program and classes being taught by teachers not receiving the *PD* program (they have had only the standard introductory training on IWBs that all teachers received). A randomized controlled trial eliminates selection bias, which is a major source of bias that without randomized assignment is very hard to correct for. While pretest and outcome measures were recorded at the student level, randomization occurred at the teacher level. Participating teachers were paired and a coin was tossed to determine which teacher would receive the *PD* program and which teacher would not receive the *PD* program. The experiment started at the beginning of the school year of 2005.

We based our statistical analyses on 8 teachers/classes (4 *PD* and 4 control) and 125 students only in grades 3 to 5. The analyses focused on this subset of students because the content areas of interest were only available for these grades in the CRCT.

The Professional Development Program on Interactive Whiteboard Use

The Forsyth County School District created a team that included school and district-level Instructional Technology Specialists. Together with Logical Choice Technologies, this team developed an ongoing, systematic professional development program on the use of Promethean interactive whiteboards. This *PD* program was structured to include: three 1 to 2 hour workshops (in addition to the initial two 1.5 hour basic training sessions of IWBs from Logical Choice Technologies) and supplemental resources from the school Instructional Technology specialist (ITS). The program aimed to cover a variety of topics on how to integrate IWB functionality with instructional lessons. The objective of this *PD* model was to train teachers to perform advanced IWB operations to enhance their instruction and engage their students in learning. The schedule, content and structure of the *PD* workshops was determined by the progress of the teachers as well as the individual ITS at each school.

Data Sources and Collection

The research for this experiment encompasses a multiple methods approach. In addition to pre and posttest scores, we collected demographic information on students and teachers. We integrated the class roster and teacher background information and all the data from the district into a standard data warehouse for the study. Additional data collection methods included teacher surveys, teacher and ITS interviews and training observations. Through these methods, we measured and documented teachers' usage time and interactions with the IWBs. These qualitative data were used as descriptive information only. The ITS and teacher interviews were used to gain further information about the *PD* program while the survey data were processed and analyzed to report teacher practices with IWBs. The research study is designed to take maximum advantage of the opportunity for meaningful difference including contrasts between teachers' levels of experience using technology.

Training Observations

Observations in January and February 2006 focused on the training at two of the three schools. The purpose of these observations was to document the instructional content and structure of the trainings as well as to observe the teachers' interaction with the training materials. The district sought assistance in integrating ongoing observations into their *PD* program. With the combined efforts of the district coordinator of instructional technology and the researchers, a preliminary *PD* observation tool was developed. This tool was intended to be used as a guideline to understand how teachers are using their IWBs. Specific categories of interest included teaching strategies that promote student learning, flexibility of instructional style, student engagement and functions during IWB usage.

Interviews

Structured phone interviews were conducted with five control teachers and five *PD* teachers and all three Instructional Technology Specialists in March 2006. Each interview lasted between 20 to 30 minutes. The purpose of teacher interviews was to gain an understanding of their experiences about integrating IWB use in their lessons, how they are specifically using their IWB, and the kinds of supplemental resources they use. For *PD* teachers, additional questions were asked relative to their attitudes and beliefs about the ongoing training. The purpose of the ITS interviews was to gather detailed information about the content of their training as well as any insight they had on the implementation of IWB usage.

Surveys

Eight web-based surveys were administered on a bi-weekly schedule from December 2005 to April 2006. The content of the surveys included any factors that may influence the results. Examples of these include instructional time with IWBs, supplemental materials, outside resources, student engagement, IWB functions, training hours, technology, purpose of IWB use, student usage and engagement, access to school's shared drive and technical difficulties among others (see Appendix A). A final survey was administered at the debrief meeting with the participating teachers in May 2006. This final survey addressed questions about comfort with technology, interest in technology, technology skills, and use of computers in high school. We considered these to be important background characteristics that could influence the amount of IWB use.

Achievement Test Scores

The primary outcome measures are student-level test scores on the Georgia's standardized Criterion-Referenced Competency Tests (CRCT). The CRCT is a state-mandated end-of-year assessment, designed to measure how well students acquire, learn, and accomplish the knowledge and skills described in the Georgia Performance Standards (GPS) and Quality Core Curriculum (QCC). The CRCT is also designed to diagnose students' individual strengths and weaknesses related to instruction of the GPS/QCC (Georgia State Department of Education, 2005). We obtained the CRCT scores in English Language Arts (GPS-based), Reading (GPS-based) and Math (QCC-based). The CRCT scores in 2005 are reported as a pretest measure, and the CRCT scores administered in April 2006 are used as the posttest measure.

Because CRCT scores for English Language Arts, Reading and Math are available for grades 3 to 5, a portion of the total sample size (125 students out of the total sample size of 344 students) was used in the achievement outcome analyses.

Site Description

Forsyth County

Forsyth County, Georgia, considered part of the 28-county metro-Atlanta, is located approximately 40 miles northeast of downtown Atlanta. According to the 2005 statistics from the U.S. Census Bureau, the total population is 140,393. Of the adult population, 89.5% have a high school diploma and 44.2% have a Bachelor's degree. The median age in Forsyth County is 34 and the median household income is $82,478. The ethnic make-up is 91.6% White, 7.5% Hispanic and 3.2% Asian.

Forsyth County Schools

The Forsyth County School District consists of over 25,000 students across 25 schools—fourteen elementary, six middle, and three high schools, as well as one alternative school for middle and high school students and a charter non-traditional high school.

Of the student population in this district, 12.9% of the students (compared to 47.9% in the state) are economically disadvantaged, 3.5% (compared to 2.9% in the state) are English Language Learners, and 13.2% (compared to 12.3% in the state) are students with disabilities.

The district has an annual growth rate of 8 to 10%. Because all of the Forsyth County Schools are equipped with IWBs, the district sought to understand the impact of the professional development program they developed for the use of IWBs.

Materials

At the initial training session, all teachers were supplied with an IWB and the supporting software and tools (e.g. computer/laptop, LCD projector, etc.), IWB training packets produced by Logical Choice Technologies, hands-on activities and an orientation of how to use the IWB.

Additional materials distributed to *PD* teachers varied across the three school sites. Some *PD* teachers received IWB peripheral equipment (e.g. keypads and wireless keyboards), advanced software, lessons, templates, ITS support and interactive lessons.

Sample and Randomization

Recruiting

The district Coordinator of Instructional Technology had garnered the support of three principals who agreed to select 20 teachers in each of their schools. The researchers distributed an explanation of the research to the principals who then relayed this information to selected teachers. After understanding the details of the study, 17 total teachers attended the initial meeting on September 22, 2005 at one of the participating schools. Also in attendance were the three ITS's from each school and the Coordinator of Instructional Technology. The meeting included an explanation of the *PD* study and a discussion about the planned research procedures. After a question-and-answer period, 15 out of the 17 teachers who decided to participate in the study filled out a teacher background information form.

Randomization

Fifteen teachers were assigned using a coin toss to either the PD condition or to control. Randomization ensures that, on average, characteristics other than treatment, which affect the outcome, are evenly distributed between treatment and control. This prevents us from confusing treatment with some other factors, technically called 'confounders', that are not evenly distributed between groups and that affects the outcome. For example, through randomization we try to achieve balance between treatment and control on the average years of teaching experience – a factor that presumably also affects the outcome.

There are various ways to randomize teachers to conditions. We used a matched pairs design whereby we first identified pairs of similar teachers, and then, within each pair, we randomized one teacher to treatment and the other to control. Similarity was based on whether teachers were in the same grade level and whether they shared common meeting times. A pairing strategy often results in a more precise measurement of the treatment impact.

Sample Size

One concern we had was with sample size. Sample size is one of the things that determines how precisely we can measure an effect of a given size. With smaller samples we are usually only able to detect larger effects. We usually measure the size of an effect in terms of standard deviation units – which tells us how big the effect is, controlling for the spread in observed scores. We were only able to use 8 teachers in the actual analysis. Based on the available sample size and certain assumptions about other parameters that affect the size of the effect that we can detect, we computed that we can detect an effect size as small as .94. This is computed assuming false-positive and false-negative error rates of .05 and .20 respectively. Raising the false positive rate to .20 reduces the size of the effect that we can detect to .62. We emphasize that the matching design that we used further lowers this value. From this we see that the experiment is not powered to detect a very small effect which may be real but not discernable given the number of teachers in the study.

Grades K through 2 were not used because CRCT pretest scores are not available for these students. This reduced the sample size to 4 teachers each in the *PD* and control conditions. With eight teachers total, we realized that we did not have as large a sample as was called for by our initial design. Because the importance of the information warranted gathering the available data even if the results ultimately proved inconclusive, the district in consultation with the researchers decided to move forward with the experiment.

Statistical Analysis and Reporting

The basic question for the statistical analysis was whether, following the intervention, students in the *PD* classrooms had higher CRCT scores in English Language Arts, Reading and Math than those in the control classrooms. The mean impact is estimated using multi-level models that account for the clustering of students in teachers, which provides a more accurate, and often more conservative, assessment of the confidence we should have in the findings. We use SAS PROC MIXED (from SAS Institute Inc.) as the primary tool for this work. To increase the precision of our estimate, we include students' pretest scores in the analysis. In our experience this is a good predictor of achievement; including them as covariates in the impact analysis reduces the error variance, which makes it easier to discern the treatment impact. We also analyzed the impact of PD on classroom use of the IWBs and conducted an exploratory analysis of whether IWB use itself is correlated with student performance

Our analyses produce several results: among them are the coefficients for fixed effects, effect sizes, and *p* values. These are found in all the tables where we report the results of the statistical models except for fixed effect estimates for teachers.

Coefficients. The coefficient can be thought of as the estimated size of an effect. Specifically, it is how much we would predict the outcome to change for a one-unit increase in the corresponding variable. We are often most interested in the coefficient associated with the condition, which is the estimated difference in the outcome between those in the treatment and control groups (holding the values of the other covariates constant and assuming no interactions).

Effect sizes. We also translate the difference between treatment and control into a standardized effect size by dividing the difference by the pooled standard deviation of the outcome, which is a measure of how variable the outcome is. This allows us to compare the results with results from other studies that use different measurement scales. In studies involving student achievement, effect sizes as small as 0.1 (one tenth of a standard deviation) are sometimes found to be important educationally. The unadjusted effect size is the difference between treatment and control, controlling for dependencies of observations within randomized units. (This has implications for *p* values, but it also affects the estimate of the difference: it weights some cluster averages more than others – therefore we can expect inconsistency between the estimated difference and the raw difference.) The adjusted effect size adjusts for the pretest as well as other fixed and random effects used in the models with interactions that follow.)

***p* values.** The *p* value is very important because it gives us a gauge of how confident we can be that the result we are seeing is not due simply to chance. Specifically, it tells us what the probability is that we would get a result as large or larger than the absolute value of the observed estimate, when in fact there is no effect. Roughly speaking, it tells us the risk of concluding that the treatment has had an effect, when in fact it hasn't. Thus a *p* value of 0.1 gives us a 10% probability of that happening. We can also think of it as the level of confidence, or the level of belief we have that the outcome we observe is not simply due to chance. While ultimately depending on the risk tolerance of the user of the research, we suggest the following guidelines for interpreting *p* values:

- We have a high level of confidence when $p <= .05$. (This is the level of confidence conventionally referred to as reaching "statistical significance.")
- We have some confidence when $.05 < p <= .15$.
- We have limited confidence when $.15 < p <= .20$.
- We have no confidence when $p > .20$.

Results

Formation of the Experimental Groups

Groups as Initially Randomized

The randomizing process does not guarantee that the groups will be perfectly matched. It simply guarantees that there is no intentional or unintentional selection bias. It is important to inspect the two groups to determine whether, in spite of randomization, there are any significant differences on factors that affect the outcome. (Randomization ensures lack of selection bias, but we are interested in knowing whether the particular estimate resulting from this randomization may be far from the true value as a result of chance imbalances on factors that affect the outcome.) The following tables address the nature of the groups in each of the sites. Tables 1 and 2 show the distribution of teachers, classes, grades, and students between *PD* and control conditions. This is the complete number of students in the experiment at the time that the experiment began in September 2005.

Table 1. Distribution of the *PD* Group by Schools, Teachers, Grades, and Counts of Students

School ID#	Teacher ID#	Class ID#	Grade	Number of students
1	4	6	K	20
	5	1	K	19
	10	3	5	26
2	2	15	4	17
	6	13	1	22
	8	11	2	20
	14	12	3	15
3	12	10	5	25
	Total *PD* Teachers=8	**Total *PD* Classes=8**		**Total *PD* Students=164**

Table 2. Distribution of the Control Group by Schools, Teachers, Grades, and Counts of Students

School ID#	Teacher ID#	Class ID#	Grade	Number of students
1	1	4	3	21
	9	7	2	21
	11	5	5	27
	13	2	5	26
2	3	14	K	21
3	7	8	1	20
	15	9	3	22
	Total Control Teachers=7	**Total Control Classes=7**		**Total Control Students=158**

Post Randomization Composition of the Experimental Groups

In checking for balance in the composition of the experimental groups, we examine teacher characteristics (such as years of teaching experience, computer use, and interest/comfort with technology), student characteristics (SES and ethnicity), as well as student pretest outcomes.

Teacher Variables

Although teachers were randomized to condition, with a very small number of teachers, it is useful to see whether the groups are balanced on variables that could be important factors in the results. Table 3 shows that the control teachers were generally more experienced than the PD teachers (one control teacher did not answer the question).

Table 3. Years Teaching Experience

Condition	Number of Teachers	0 to 3 years	4 or more years
PD	7	3	4
Control	7	1	6
Totals	14	4	10

However, when asked about their experience, comfort, and expertise with technology, the two groups were quite similar (two *PD* teachers did not complete this survey). All teachers except for one in each condition selected the high technology skill category, which indicates balance. We used Fisher's Exact Test (Table 6) to check whether there is an association between treatment and the dichotomous covariate, and we ran both a *t* test and used the Mann-Whitney test to check whether a difference exists between *PD* and control on the technology scale. The high *p* values in both the results shown below indicate randomization resulted in balance on these important variables.

Table 4. High School Computer Use of *PD* and Control Teachers

	Computer use in high school		
	No	**Yes**	**Totals**
PD	4	2	6
Control	3	4	7
Totals	7	6	13
Fisher's exact test			***p* value**
			.59

Note. Due to the small number of cases in some of the cells, we use Fisher's exact test.

Table 5. Technology Interest and Comfort of *PD* and Control Teachers

Descriptive statistics: technology interest and comfort	Raw group means	Standard deviation	Number of teachers	Standard error	Effect size
PD	7.5	1.80	6	0.22	-0.16
Control	7.29	0.55	7	0.68	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*PD* – control)	-0.21		11	-0.28	.79

Student Variables

Socio-Economic Status

Table 6 shows the distribution of the socio-economic status (SES) of the students in each group, as determined by participation in the Free/Reduced-price Lunch program. Randomization resulted in SES being evenly balanced between *PD* and control. We confirmed this formally by way of a chi-square test which yielded a *p* value of .71.

Table 6. SES for *PD* and Control Groups

Condition	In the Free Lunch program		
	No	**Yes**	**Totals**
PD	111	53	164
Control	110	48	158
Totals	221	101	322
Chi-square statistics	**DF**	**Value**	***p* value**
	1	0.14	.71

Ethnicity

Table 7 summarizes the distribution of student ethnicity. A majority of the students is Caucasian, which coincides with the general ethnicity of Forsyth County. This implies that this sample is a good representation of the community. As a result of random assignment, the ethnicity of the students is evenly distributed across the *PD* and control groups. The result of the statistical test is consistent with this assertion.

Table 7. Ethnicity for *PD* and Control Group

Condition	Ethnicity						
	Asian	Hispanic	Native Indian	Multi-racial	Black	White	Totals
PD	5	26	1	1	4	127	164
Control	1	25	0	1	3	128	158
Totals	6	51	1	2	7	255	322
Fisher's exact test	***p* value**						
	0.63						

Note. Due to the small number of cases in some of the cells, we use Fisher's exact test to test for a difference between conditions in the distribution of ethnicity.

Pretest Scores

We also checked whether randomization resulted in balance on pretest scores, a variable that we include in most of our analyses to increase the precision of our estimates. Tables 8 through 10 show the results for non-disability students in grades 3 to 5 for which pretests were available. In each of these tables, the statistical test indicates that balance on the pretest scores was achieved.

Table 8. Difference in Reading Pretest Scores Between Students in the *PD* and Control Groups

Descriptive statistics: CRCT Reading pretest	Raw group means	Standard deviation	Number of students	Standard error	Effect size
PD	-0.02	0.99	65	0.12	-0.04
Control	0.02	1.00	71	0.12	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*PD* – control)	-0.04		134	0.21	.83

Notes. There is no discernible difference between the means of the two groups (a very high p value and a very low effect size). The results are expressed in terms of z-scores which are centered on zero. The rationale for and details of this transformation are described in the Quantitative Results section of the report.

Table 9. Difference in Math Pretest Scores Between Students in the *PD* and Control Groups

Descriptive statistics: CRCT Math pretest	Raw group means	Standard deviation	Number of students	Standard error	Effect size
PD	-0.09	0.93	65	0.11	-0.17
Control	0.08	1.05	71	0.12	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*PD* – control)	-0.17		134	0.98	.33

Table 10. Difference in ELA Pretest Scores Between Students in the *PD* and Control Groups

Descriptive statistics: CRCT English Language Arts pretest	Raw group means	Standard deviation	Number of students	Standard error	Effect size
PD	0.08	1.09	65	0.13	0.15
Control	-0.07	0.90	71	0.11	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*PD* – control)	0.15		134	-0.87	.38

Subsets of Students Used in Analysis

Our analysis of the impact of *PD* on teacher usage of the IWB was based on the full set of teachers in the sample. However, our analysis of the student outcomes was restricted to those students who took the state tests (CRCT). We also removed students labeled as having a disability since we did not have any information on how their disability may affect their performance.

Attrition

Based on the cases of students without disabilities from grades 3, 4, and 5, a low percentage of students did not take the posttests (Spring 2006 CRCT ELA and Reading). Out of a total enrollment of 155 based on fall class rosters, 19 students (or 12.2%) did not have pretest scores (Spring 2005 CRCT ELA and Reading). Of these remaining 136 students, posttest scores are missing for 10 or 7.3%. There are also 12 students who have posttest scores but did not have a pretest score. We found a similar attrition rate for CRCT Math. Table 11 shows the breakdown by the *PD* and control groups. An exact test of differences in proportion indicates that there is no relationship between level of attrition and experimental condition (or, there is no differential attrition). This is important because it means that attrition does not introduce bias. With low attrition that was not differential, we did not analyze the differences further.

Table 11. Students Missing Test Score Data

Condition	Categories of missing data[a]		
	Having both pre and posttest scores	Having only pretest score	Totals
PD	59	6	**65**
Control	67	4	**71**
Totals	**126**	**10**	**136**
Fisher's exact test[b]	**DF**	**Value**	***p* value**
	1	0.19	0.52

[a] Refers to scores on the 2005 CRCT ELA and Reading and 2006 CRCT ELA and Reading

[b] Due to the small number of cases in some of the cells, we use Fisher's exact test to test for a difference between conditions in the proportion of students with pretest scores.

Implementation Results

Methods and Response Rates

We used the following questions to guide our descriptions and analysis: How differently are *PD* teachers and control teachers using their IWBs with respect to both *IWB functionality* and *instructional time*? Also, how are students and teachers responding to the use of IWBs?

Surveys were deployed to both *PD* and control group teachers beginning on December 12, 2005 and continued on a bi-weekly basis until April of 2006. Response rates were calculated using a simple percentage calculation based on the ratio of actual received responses to the number of expected responses. There were 8 teachers in the *PD* group and 7 teachers in the control group. A total of nine surveys were deployed with an overall response rate of 94.81% for both groups, a 90.28% response rate for the *PD* teachers, and a 100% response rate for the control teachers.

The survey topics were developed to account for the various aspects of teacher and student actions associated with instruction and learning. In order to characterize the average time teachers and students spent using IWBs, we used a repeated question strategy. On surveys 2 through 8, we repeated these questions: a) Last week, how many average daily minutes did you use your IWB for instruction? and b) In an average week, how much of your own time do you spend utilizing outside resources to enhance your effectiveness teaching with an interactive whiteboard (e.g. . individual study on the Internet etc, one-on-one time with your school's ITS, help from other teachers, etc)? These questions, together with questions regarding the types of activities, allow us to draw inferences about how time was devoted to IWB use in both the PD and Control groups. Survey 9 focused on the content covered and teachers' overall experience with the materials.

Table 12. Survey Response Rates

Survey number	Date	Topic	Treatment response rate	Control response rate	Overall response rate
Survey 1	December 12-16	Training	100%	100%	100%
Survey 2	January 16-20	Planning and Preparation Time	100%	100%	100%
Survey 3	February 6-10	IWB Functions	87.50%	100%	93.33%
Survey 4	February 20-24	IWB Functions	100%	100%	100%
Survey 5	March 6-10	Access and Use of School's Shared Drive	100%	100%	100%
Survey 6	March 20-24	Instructional Use with IWB	100%	100%	100%
Survey 7	March 27-31	Student Engagement and Interactions	62.50%	100%	80.00%
Survey 8	April 17-21	Technical Difficulties	87.50%	100%	93.33%
Survey 9	May 1	Teacher Relationship with Technology	75.00%	100%	86.67%

Implementation of IWB—IWB Functionality

Surveys and interviews revealed that all teachers, in general reported increased student interest and engagement compared to the previous year before the introduction of IWBs, as well as overall satisfaction with the IWBs. One teacher stated "students have reacted very positively to the IWB. I get the focus of the whole class. All eyes are on the IWB because it's such a magnet." Another teacher commented "I love using the IWB. My students are engaged in learning and it is a very valuable tool for accessing information and for student interaction."

Table 13 displays the different functions teachers performed when using their IWBs. As indicated, both groups of teachers used their IWBs for flipcharts and internet exploration. There was a slight increase in IWB functions among the *PD* group.

Table 13. Differences in IWB Functions Between *PD* and Control Groups

Teacher usage	*PD* teachers	Control teachers
Flipcharts	7	6
Internet exploration	7	6
Storytelling	2	1
Movies: showing and creating	7	4
Graphics: showing and creating	5	2
Demo of computer functions	4	3
Testing and quizzing students	2	4
Class management/calendar	1	3
Lesson planning	0	2

The school district was also interested in how teachers had their students use the IWBs. In Table 14, *PD* group teachers reported students using the IWBs more for basic presentation (28.75%) than for critical thinking activities (10%) or other uses (10%). By contrast, control group teachers reported students using the IWBs more for critical thinking activities (15.71%) than for other uses (13.57%) or basic presentation (12.86%).

Table 14. Student Use of IWBs for *PD* and Control Groups

Student usage	*PD* teachers	Control teachers
Basic presentation	28.75%	12.86%
Critical thinking activities	10%	15.71%
Other uses	10%	13.57%

Most teachers (in both groups) expressed frustration in the lack of time available to plan and practice with their IWBs. They expressed that integrating such technology into their classroom instruction would require a considerable amount of time that they just don't have. Other teachers (in both groups) revealed the extent to which they access outside and in-house resources, regardless of the study parameters. Teachers who utilized outside resources did so apart from their instructional planning time via the Internet, the county-sponsored Digital Conference workshops, and/or other technology support. Teachers who utilized in-house resources accessed support from other teachers or their school ITS. Table 15 indicates why and how often teachers used their school's shared network or shared drive.

Table 15. Differences in Access to Shared Drive Between *PD* and Control Groups

Use of "shared drive"	*PD* teachers	Control teachers
Share self-created flipcharts	40%	25%
Access flipcharts others created	32%	38%
Other	28%	37%

As a result of *PD*, the *PD* teachers would create and share their flipcharts 40% of the time and the control teachers would access these flipcharts 38% of the time.

All three ITSs confirmed having provided support to all teachers (including control teachers). Such support included trouble shooting their technical difficulties, guiding their IWB processes, and providing resources to them via their school's network shared drive.

ITSs also revealed how differently the *PD* workshops are structured in each school. In one school, for example, the *PD* workshops were open to any teacher who wanted to participate. In another school, the *PD* workshops were exclusive to a "core team" which was composed of grade-level representatives and media leaders. Grade-level representatives are then required to report what they have learned from the *PD* workshops to other teachers in their grades, including control teachers.

ITSs further explained how teachers at their school were interacting with their IWBs. Their descriptions revealed that some of the control teachers were proactive in their whiteboard use. For instance, one of the control teachers participated in the Forsyth County Schools Digital Conference as a presenter of making movies, an advanced IWB function. Meanwhile, some *PD* teachers were intimidated by such technology. These teachers were embarrassed at times about demonstrating difficulties with their IWBs in front of their students.

Overview of Quantitative Results

We report the results in three parts as illustrated in Figure 1.

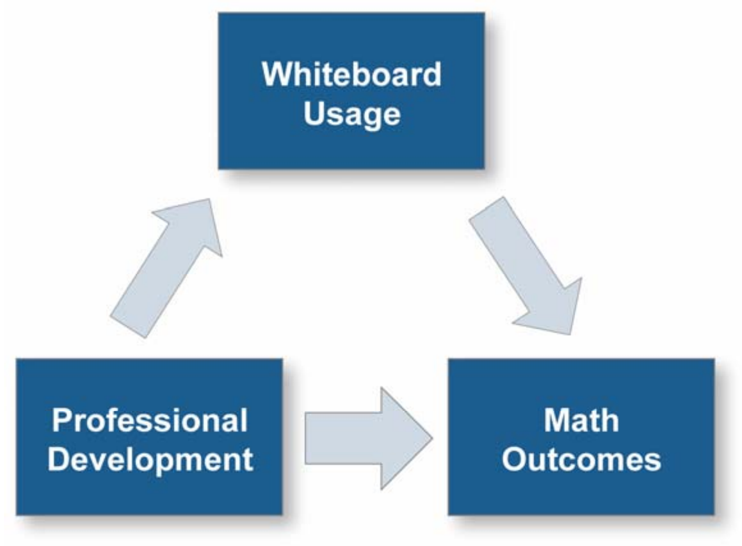


Figure 1. Outcomes Investigated in the Statistical Analysis

The primary topic of our experiment was the impact of additional professional development. We first look at the impact of *PD* on the state test scores (CRCT). This is broken into separate analyses for English Language Arts, Reading and Math. Second, we examine the impact of *PD* on the use of IWBs, that is, on teacher classroom behavior. Did the teachers who were given the additional *PD* use the IWBs more than the control teachers? Third, we ask whether the amount of IWB usage was related to the student outcomes, in particular in Math, as explained below. In this way, if there is an impact of *PD* on student outcomes, we can understand the extent to which that impact was mediated by IWB usage[1].

We emphasize that the small number of teachers in the experiment limits the generality of these findings. In particular, given the statistical techniques used[2], we are not able to generalize beyond this particular sample of teachers. The other limitation of the small number of teachers is that it prevents us from detecting small effects – we might conclude that we have no confidence that the effect is different from zero when the effect is in fact different from zero, but only slightly so. Such a result should not be confused with a conclusion that there was no impact.

Impact of *PD* on CRCT Scores

Our overall outcome measures were the scores on the Georgia CRCT State Assessments in English Language Arts, Reading and Math. The basic question for the statistical analysis was whether, following the *PD* program, students in *PD* classrooms had higher scores than those in control classrooms. In addition to looking at the main effect of *PD* and doing related analyses of covariance

[1] We were unable to perform a formal test of the mediating effect of IWB usage on Math outcomes. Assignment was at the teacher-level, IWB usage was at the teacher-level and the outcome, math performance, was at the student-level. This precluded use of the Baron and Kenny approach, which applies to one-level analyses. We did however test the strengths of the relationships between all three variables – the three arms in the schematic figure above– to provide preliminary evidence concerning a possible mediating effect of IWB usage.

[2] We modeled teachers as fixed effects.

(ANCOVAs), we estimated the interactions of condition (*PD* versus control) with the pretest of the students. In particular, we were interested in whether *PD* was differentially effective for low- and high-performing students. We first address the impact of *PD* on CRCT outcomes in English Language Arts, Reading and Math. Within each content area we provide a statistical analysis of the impact of *PD* controlling for pretest and examine the interaction of *PD* with pretest, that is, we examine whether students initially scoring higher or lower on the pretest differentially benefited from *PD.* The following tables present the results of our statistical modeling in each content area.

Since there is no information regarding whether the tests were vertically scaled for grades, we were not able to analyze the students test scores directly. A within-grade z-transformation was performed on both pre and post test scores. That is, we standardized the score distributions within each grade level by centering them on zero and dividing them by the pooled standard deviation for the grade level. Assuming that the treatment has the same effect at each grade level (in terms of standard deviation units) transforming the scales this way allows us to put the results for all teachers and students into one analysis, thereby increasing the power of the experiment.

CRCT Scores in English Language Arts

Table 16 shows the estimated impact of *PD* on students' performance on the CRCT score in English Language Arts. The first part of the table shows the contrast between the *PD* and control groups in terms of their mean scores and includes other descriptive details including the unadjusted effect size of -0.26. The adjusted effect size is -0.31 with a *p* value of 0.19. (The adjusted effect size essentially controls for chance imbalances on the pretest.)

The row in the table labeled "Impact of *PD* for a student with an average pretest" gives us information about whether *PD* made a difference in the CRCT scores in English Language Arts for a student who has an average score on the pretest. The coefficient associated with the treatment is -0.31. This shows a small negative difference associated with *PD* for a student with an average score on the pretest. The *p* value of 0.20, indicates that we can expect to see a difference, as large or larger than the absolute value of the estimate, 20% of the time when in fact there is zero impact. Using the criteria outlined earlier in the report, we conclude that we have limited confidence that the true impact is different from zero.

We also estimated the moderating effect of the prior score on the treatment impact to see whether the intervention was differentially effective for students at different points along the pretest scale. The *p* value for this effect is .81. We have no confidence that the estimated effect is different from zero.

Table 16. The Impact of PD on the CRCT Score in English Language Arts

Descriptive statistics: CRCT ELA outcome	Raw group means	Standard deviation	No. of students	No. classes	No. of teachers	Unadjusted effect size
PD	-0.14	1.00	62	4	4	-0.26
Control	0.12	0.98	76	4	4	
Mixed model: fixed factors related to test overall outcomes	**Estimate of coefficient**	**Standard error**	**DF**	***t* value**		***p* value**
Predicted value for a control student with an average pretest	0.06	0.16	108	0.37		.71
Predicted change in control outcome for each unit increase on the pretest	0.86	0.09	108	9.67		<.01
Impact of *PD* for a student with an average pretest	-0.31	0.24	108	-1.30		.20
Interaction of pretest and *PD*	-0.03	0.14	108	-0.25		.81

Notes.

1) The unadjusted effect size is based on z-transformed scores.
2) In addition to the fixed effects shown in the table above, we included indicators for teachers and assignment pairs. We do not display the corresponding fixed effects estimates. The prior score was centered at the mean; therefore, the effect estimates apply to a student who had an average score on the pretest.
3) 322 students were originally included in the sample. We removed students from Grades K, 1, and 2. We also removed cases identified as disabled from Grades 3, 4 and 5. Together we removed 167 cases. Of the remaining cases, 138 had a posttest and were included in the calculation of the unadjusted effect size. 126 students had pretest and posttest scores. Of these we removed 7 because they were influential points (where the determination of influential points was based on the model in the table above.) The remaining 118 students were used to compute the adjusted effect size and the estimates in the table above. The number of teachers for the reduced sample of 118 students was 8 (4 control and 4 treatment) and the number of classes was 8 (4 control and 4 treatment.)

As a visual representation of the result described in Table 16, we present a scatterplot in Figure 2, which graphs student pre- and post-test CRCT scores in English Language Arts. This graph shows where each student fell in terms of his or her starting point (horizontal x-axis) and his or her outcome score (vertical y-axis.) We remind the reader that both the pre- and posttest scores have been z-transformed within grades and so are approximately centered on zero. The darker dots represent PD students; the lighter dots, control students.

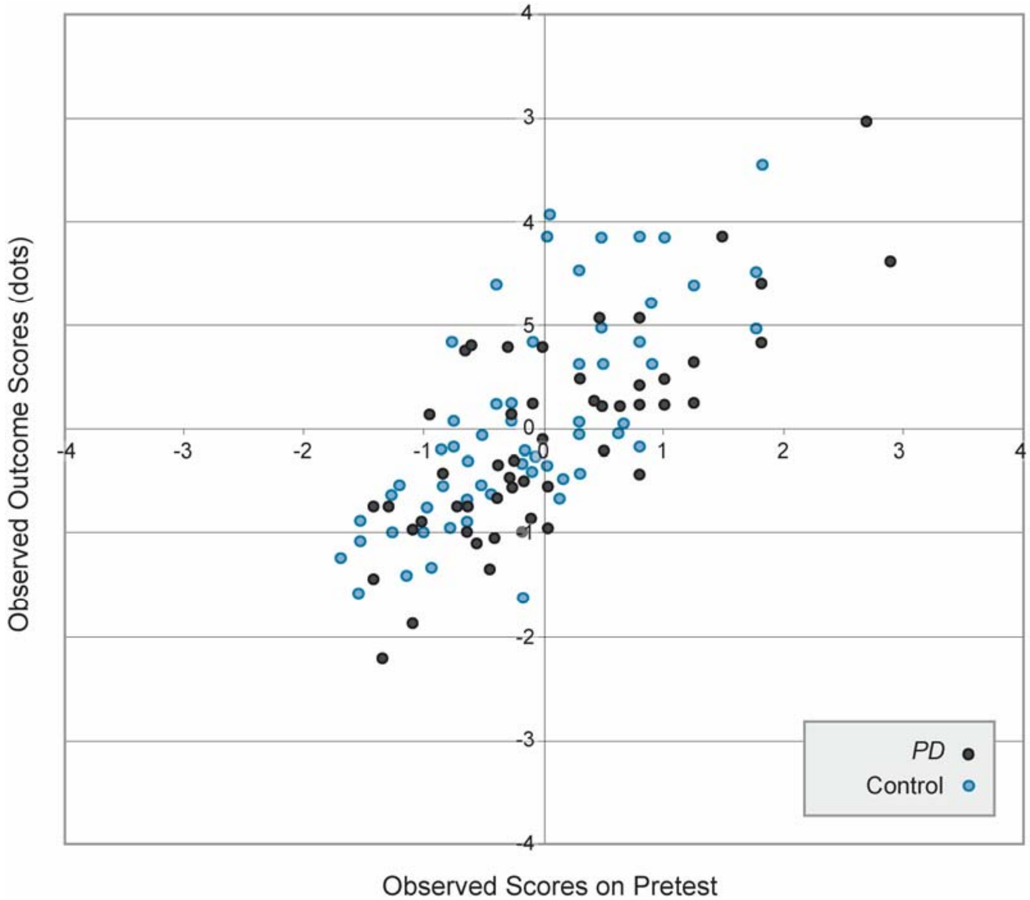


Figure 2. CRCT Outcomes in English Language Arts: Differences Between *PD* and Control Group Students

CRCT Scores in Reading

Table 17 shows the estimated impact of *PD* on students' performance on the CRCT score in Reading. The first part of the table shows the contrast between the *PD* and control groups in terms of their mean scores and includes other descriptive details including the unadjusted effect size of -0.11. The adjusted effect size was -0.25 with a *p* value of.39. (The adjusted effect size essentially controls for chance imbalances on the pretest.)

The row in the table labeled "Impact of *PD* for a student with an average pretest" gives us information about whether *PD* made a difference in the CRCT scores in reading for a student who has an average score on the pretest. The coefficient associated with the treatment is -0.24. This shows a small negative difference associated with *PD*. The large *p* value of.40 indicates that the difference that we observe could easily be the result of chance.

We also estimated the moderating effect of the prior score on the treatment impact to see whether the intervention was differentially effective for students at different points along the pretest scale. The *p* value for this effect is .54. We have no confidence that the actual effect is different from zero.

Table 17. The Impact of *PD* on the CRCT Score in Reading

Descriptive statistics: CRCT in Reading outcomes	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
PD	-0.06	0.97	62	4	4	-0.11
Control	0.05	1.02	76	4	4	
Mixed model: fixed factors related to CRCT in Reading	**Estimate of coefficient**	**Standard error**	**DF**	***t* value**		***p* value**
Predicted value for a control student with an average pretest	0.08	0.20	115	0.41		0.68
Predicted change in control outcome for each unit increase on the pretest	0.68	0.10	115	7.10		<.0001
Impact of *PD* for a student with an average pretest	-0.24	0.28	115	-0.84		0.40
Interaction of pretest and *PD*	-0.08	0.14	115	-0.61		0.54

Notes.

1) The unadjusted effect size is based on z-transformed scores.
2) In addition to the fixed effects shown in the table above, we included indicators for teachers and assignment pairs. We do not display the corresponding fixed effects estimates. . The prior score was centered at the mean; therefore, the effect estimates apply to a student who had an average score on the pretest.
3) 322 students were originally included in the sample. We removed students from Grades K, 1, and 2. We also removed cases identified as disabled from Grades 3, 4 and 5. Together we removed 167 cases. Of the remaining cases, 138 had a posttest and were included in the calculation of the unadjusted effect size. 126 students had pretest and posttest scores. Of these we removed 1 because it was an influential point (where the determination of influential points was based on the model in the table above.) The remaining 125 students were used to compute the adjusted effect size and the estimates in the table above. The number of teachers for the reduced sample of 125 students was 8 (4 control and 4 treatment) and the number of classes was 8 (4 control and 4 treatment.)

As a visual representation of the result described in Table 17, we present a scatterplot in Figure 3, which graphs student performance relative to the z-transformed distributions on the CRCT Reading pre- and posttests. This graph has the same interpretation as Figure 2.

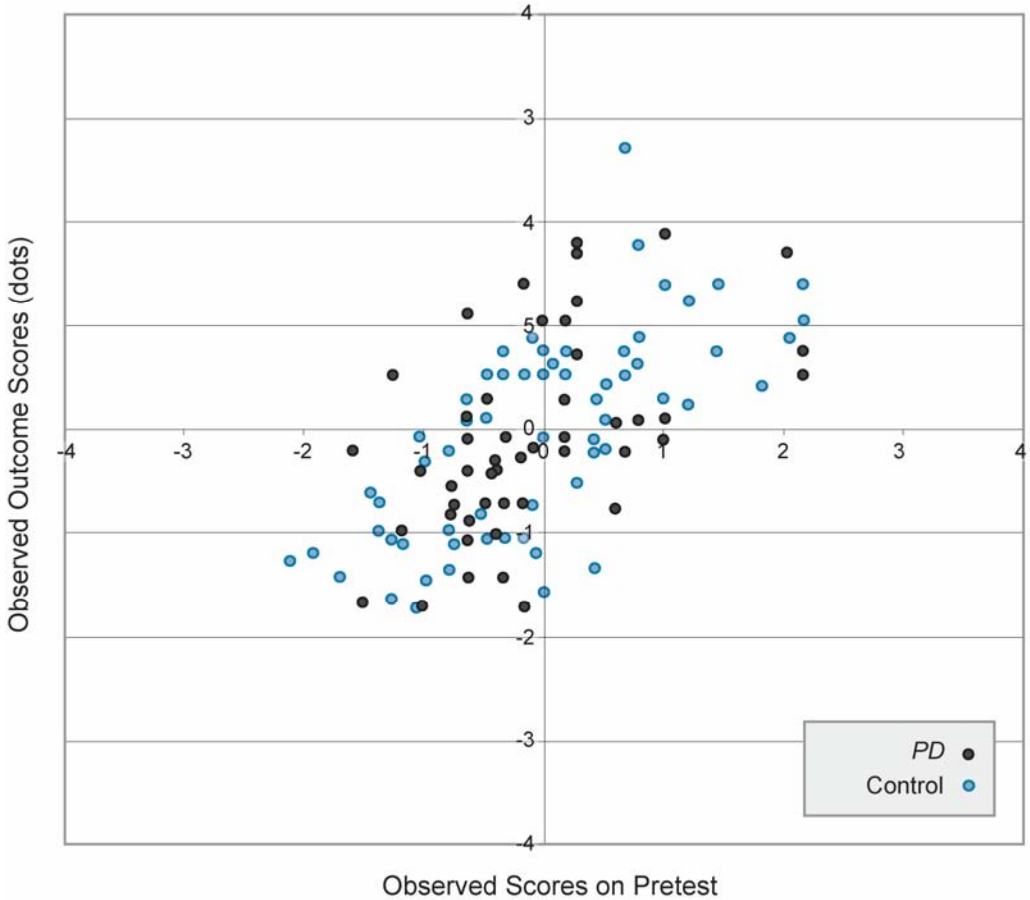


Figure 3. CRCT Outcomes in Reading: Differences Between *PD* and Control Group Students

CRCT Scores in Math

Table 18 shows the estimated impact of PD on students' performance on the CRCT in Math. The unadjusted effect size is -0.35. The adjusted effect size is -0.30 with a *p* value of 0.25

The row in the table labeled "Impact of *PD* for a student with an average pretest" gives us information about whether PD made a difference in the CRCT scores in math for a student who has an average score on the pretest. The coefficient associated with the treatment is -0.32. This shows a small negative difference associated with PD. The *p* value of 0.15, gives us some confidence that the true impact is different from zero.

We also modeled the interaction between the prior score and the treatment effect. The *p* value for this effect is .26. We have no confidence that the actual effect is different from zero.

Table 18. The Impact of PD on the CRCT Score in Math

Descriptive statistics: CRCT math outcome	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
PD	-0.19	1.01	63	4	4	-0.35
Control	0.16	0.96	76	4	4	
Mixed model: fixed factors related to test overall outcomes	**Estimate of coefficient**	**Standard error**	**DF**	***t* value**		***p* value**
Predicted value for a control student with an average pretest	0.28	0.16	108	1.77		.08
Predicted change in control outcome for each unit increase on the pretest	0.59	0.07	108	8.33		<.01
Impact of *PD* for a student with an average pretest	-0.32	0.22	108	-1.43		.15
Interaction of pretest and *PD*	0.13	0.11	108	1.14		.26

Notes.

1) The unadjusted effect size is based on z-transformed scores.
2) In addition to the fixed effects shown in the table above, we included indicators for teachers and assignment pairs. We do not display the corresponding fixed effects estimates. . The prior score was centered at the mean; therefore, the effect estimates apply to a student who had an average score on the pretest.
3) 322 students were originally included in the sample. We removed students from Grades K, 1, and 2. We also removed cases identified as disabled from Grades 3, 4 and 5. Together we removed 167 cases. Of the remaining cases, 138 had a posttest and were included in the calculation of the unadjusted effect size. 126 students had pretest and posttest scores. Of these we removed 1 because it was an influential points (where the determination of influential points was based on the model in the table above.) The remaining 125 students were used to compute the adjusted effect size and the estimates in the table above. The number of teachers for the reduced sample of 118 students was 8 (4 control and 4 treatment) and the number of classes was 8 (4 control and 4 treatment.)

As a visual representation of the result described in Table 18, we present a scatterplot in Figure 4, which graphs student performance relative to the z-transformed distributions on the CRCT Math pre- and posttests.

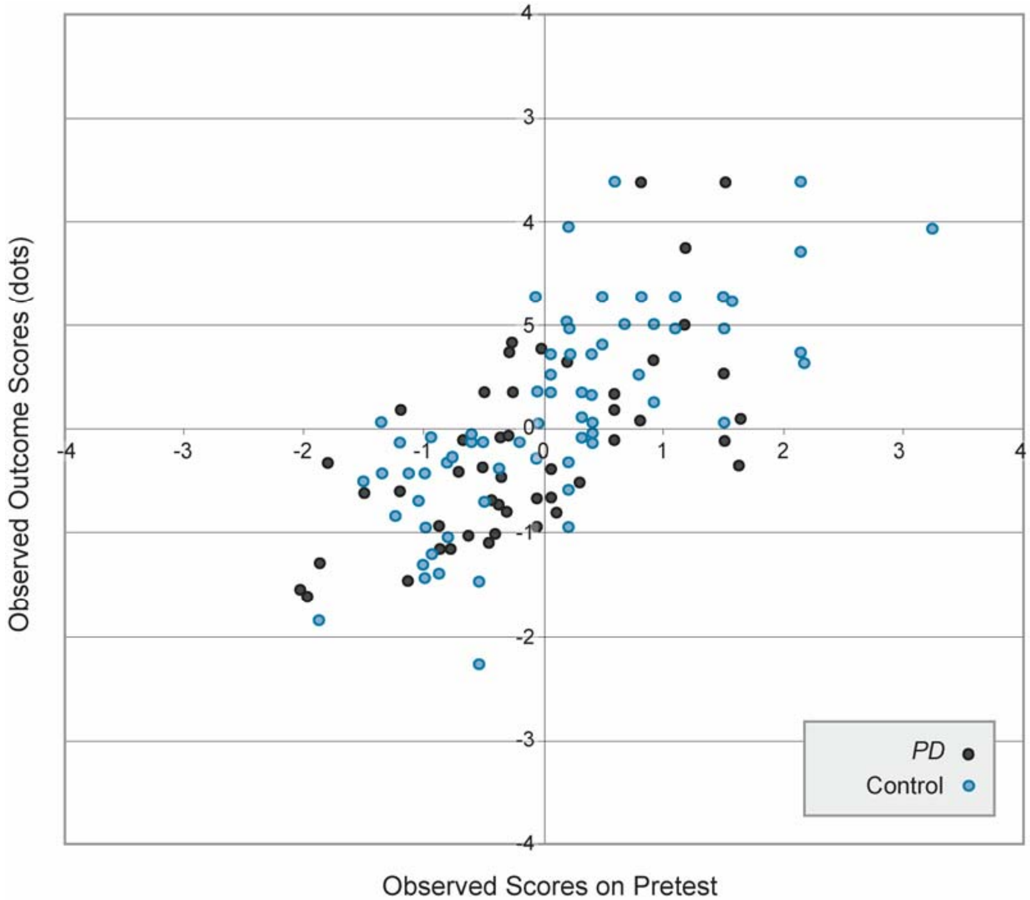


Figure 4. CRCT Outcomes in Math: Differences Between *PD* and Control Group Students

Impact of *PD* on Whiteboard Usage

In addition to understanding the impact of *PD* on student outcomes, we wanted to explore the impact of *PD* on IWB usage. The Forsyth County Schools developed the *PD* program with the expectation that the process would influence instructional practices, specifically the amount of IWB use. The surveys provided data on this variable. Our outcome is average time using IWBs. Instructional time was measured by each teacher's self-report of the number of minutes s/he spent using the IWB per week. Results were averaged across eight surveys that were administered every two weeks.

Instructional time with IWB can be considered a variable that potentially mediates the relationship between the exposure (*PD*) and the outcome (student achievement). That is, *PD* may encourage more time spent with IWB which, in turn, affects learning. (*PD* may also work through other mechanisms to produce an effect on achievement.) At this point it is important to measure the impact of *PD* on time spent using IWB because it represents a preliminary step in a potential causal chain. Because teachers were randomized to conditions (i.e., *PD* and control groups), the difference in instructional time with IWB is considered a causal effect of the intervention (*PD*). Table 19 reports an analysis comparing the amount of time spent by teachers in the *PD* group compared to teachers in the control group. Because this analysis does not depend on student test scores, we were able to use all the teachers including those teaching K through 2nd grades.

Table 19. Mixed Model for Average IWB Usage

Descriptive statistics: average IWB usage outcome	Raw group means	Standard deviation	Number of teachers		
PD	189.55	92.78	8		
Control	298.39	151.68	7		
Model: fixed factors related to average usage outcome	**Estimate of coefficient**	**DF**	**Standard error**	***t* value**	***p* value**
Intercept	298.39	1	46.68	6.39	<.01
Impact of *PD* on IWB usage	-108.85	1	63.92	-1.70	.11

Survey analyses revealed that *PD* group teachers reported using their IWBs for 189.55 average weekly minutes compared to control teachers using their IWBs for 298.39 average weekly minutes. With a *p* value of.11, we have some confidence that there is a difference between the groups in average IWB usage.

Figure 5 shows that the control group teachers spent more time using their IWBs than the *PD* teachers.

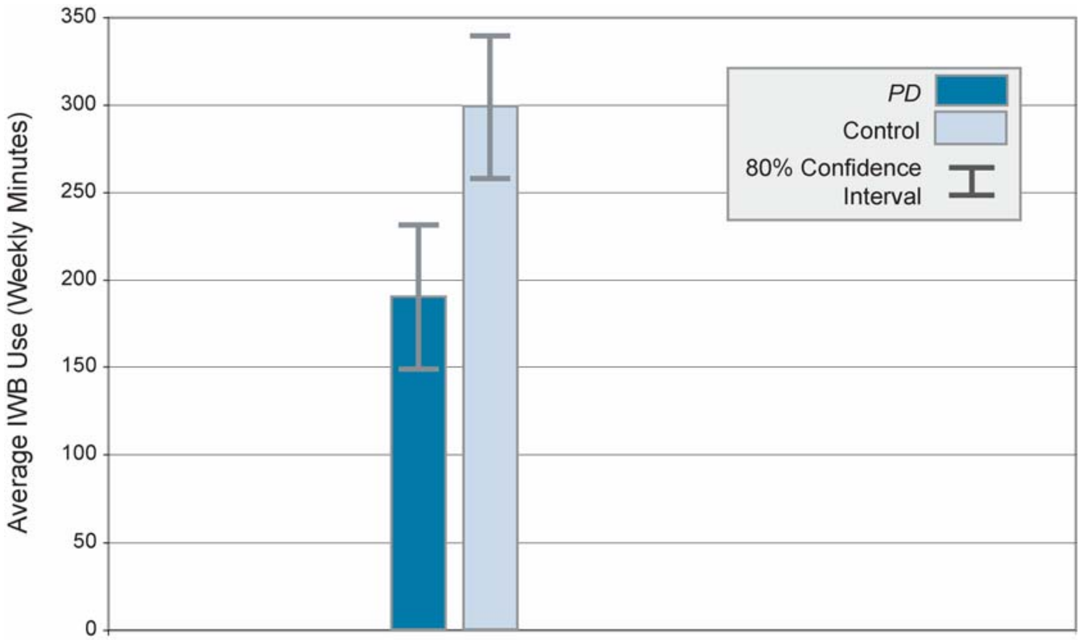


Figure 5. Differences in IWB Usage Between *PD* and Control Groups

PD teachers, on average, spent 109 minutes less per week using the IWB than their counterparts in the control condition. This corresponds roughly to a one-third reduction in the time spent using whiteboard. Since the sample size of teachers is relatively small, we also used a non-parametric test (the Mann-Whitney test) to compute the probability of getting a result as large as the one observed if in fact there was no difference. The *p* value associated with this test is .12.

With such a small sample, it is very possible for an imbalance to occur between teachers in the *PD* and control conditions on some variable that is related to the outcome. In this respect it is important to note the analysis of survey questions that established that there was no systematic difference between *PD* and control group teachers on technology expertise, interest and experience. Other, unmeasured variables may still have influenced the result.

Relationship Between IWB Usage and Student Outcome

Since we have some degree of confidence that the differences between the *PD* and control groups on the CRCT outcomes in both Math and Reading and IWB usage are not simply due to chance, we sought to explore the relationship between IWB usage and student outcomes on the CRCT in Math and Reading, the content areas where we found the largest differences between *PD* and control on student achievement. The relationship between IWB usage and the student outcome is correlational – we have not assigned teachers to levels of IWB usage so we cannot be sure whether it is IWB usage or some other variable which is correlated with IWB usage (e.g., teacher enthusiasm) that is the true cause of the student outcome. A test of the correlation[3] reveals a positive relationship between IWB usage and the student outcome.

This is evident in Figure 6, which shows a positive slope in the predicted relationship between the z-transformed CRCT Math scores and IWB Usage. The *p* value for this effect is .14, which gives us some confidence that the actual relationship between these two variables (i.e., the true value of the slope) is different from zero. The graph depicts CRCT Math outcomes only; however, we received the same effect for Reading.

---

[3] We modeled the estimated impact of *IWB usage* on students' performance on the CRCT score in math controlling for dependencies of observations within randomized units. The prior score was centered at the mean, therefore, the effect estimates apply to a student who had an average score on the pretest.

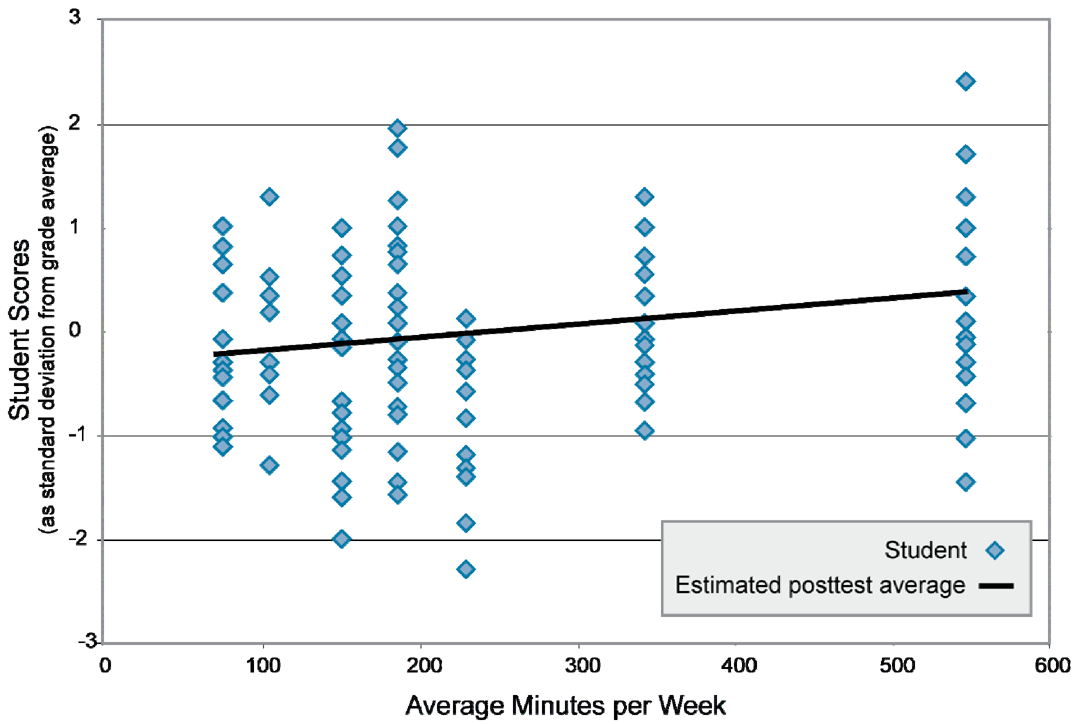


Figure 6. Correlation between IWB Usage and CRCT Outcomes in Math for *PD* and Control Groups

Discussion

Our experiment yielded several findings. First, we found that we have some confidence that the additional professional development had a negative impact on CRCT Math outcomes and we have limited confidence of the negative impact on the CRCT Reading outcomes. Second, we found that control teachers used their IWBs considerably more than the *PD* teachers. And third, the significant differences we found in student performance appear to be related to the degree of IWB usage.

To understand the trend in the negative direction on the *PD* students, we considered two different phenomena. The first is the contamination that resulted from the sharing of materials with control teachers by *PD* teachers. While contamination is unfortunate from a research standpoint, the sharing of resources is consistent with the culture of cooperation within schools that appears to be an effective tool for spreading resources, information, and expertise. “Extraordinarily successful schools… feature smaller, more personalized, and less fragmented structures...Teacher collaboration in these settings promotes knowledge sharing and communication that focuses ‘a faculty's collective technical expertise on specific problems within the school’” *(*Linda Darling-Hammond, 1995). Providing a shared school network or drive reflects that the district maximally utilized their resources within their schools.

While contamination may explain a finding of no difference between *PD* and control, other factors must have been at work to explain the negative result both for the *PD* group’s students’ achievement and for the use of the IWB’s themselves. Although the randomization process we used for forming teacher groups guaranteed that bias was not introduced intentionally, the conclusions must be treated with caution given the small number of teachers involved. With a larger pool of teachers, the likelihood of this imbalance occurring on some unmeasured variable would have been lower. We established that both groups were well matched with respect to interest in and expertise with technology, although the control group tended to have somewhat more experienced teachers.

The control students’ better performance could perhaps be attributed to what is known as the John Henry Effect. John Henry is the legendary steel driver who, knowing his work would be compared to a steam drill, worked so hard that he outperformed the drill and died of exertion (Saretsky, 1972). Public assignment of units to experimental and control conditions may cause social competition, whereby the control group tries to show it can do as well as the treatment group despite not getting benefits *(*Shadish, Cook, & Campbell, 2002).

Since the control teachers used the IWBs significantly more and had students who tended to score higher on the CRCT, it makes sense to explore whether IWB usage and achievement are related. We found that there was a correlation between the two suggesting as a subject for further research that IWB use may result in improved academic achievement.

Our goal in this research was to provide the district with evidence that would be useful in determining the impact of their *PD* program. We found that the value of the *PD* appeared to be eclipsed by the culture of sharing in the schools, which provided the control teachers with access to the same resources. Insofar as first the *PD* assisted the schools in developing IWB resources for shared use, and second, the resources helped to increase the overall usage for all teachers, the *PD* can be considered a success. The suggestive finding from the exploratory analysis that IWB use may have an impact on achievement remains to be confirmed with properly controlled comparisons.

References

Baron R.M., & Kenny D.A. (1986).The moderator-mediator variable distinction in social psychological research. *Journal of Personality and Social Psychology*, 51:1173-1182.

Bloom, H. S., Bos, J. M., & Lee, S., (1999) Using cluster random assignment to measure program impacts. *Evaluation Review, 23*, 445-469.

Darling-Hammond, L. (1995). Restructuring schools for student success. *Daedalus*. 124:4, 153+.

Fullan, M., & Promfret, A. (1977). Research on Curriculum and Instruction Implementation. *Review of Educational Research*, 47, 2, 335-397.

Georgia State Department of Education (2006). Retrieved on August 30, 2006 from http://public.doe.k12.ga.us/ci_testing.aspx?PageReq=CITestingCRCTDesc.

Mukti, N. (2000). Computer Technology in Malaysia: Teacher's Background Characteristics, Attitudes, and Concerns. *The Electronic Journal on Information Systems in Developing Countries* 3, 8, 1-33.

Raudenbush, S. W., (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*, 173-185.

Saretsky, G. (1972). The OEO PC experiment and the John Henry effect. *Phi Delta Kappan, 53* 579-581.

Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and Quasi-Experimental Designs forGeneralized Causal Inference*. Boston, MA: Houghton Mifflin.

Supovitz, J., & Turner, H. (2000). The Effects of Professional Development on Science Teaching Practices and Classroom Culture. *Journal of Research in Science Teaching,* 37, 9, 963-980.