Abstract Title Page
Not included in page count.

Title: Using Data to Inform Decisions: How Teachers Use Data to Inform Practice and Improve Student Performance in Mathematics

Authors and Affiliations:

Cavalluzzo, Linda
Vice President, CNA Resource Analysis Division,
Principal Investigator and Project Director
cavallul@cna.org

Geraghty, Thomas M.
Research Scientist, CNA Education
geraghtyt@cna.org

Steele, Jennifer L.
Policy Researcher, RAND Corporation
jsteele@rand.org

Alexander, Jane K.
Senior Research Specialist, CNA Education
alexanderj@cna.org

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A1000445 to CNA, Corp. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

The last two decades have witnessed a vast expansion in the use of education data to improve classroom instruction and raise student achievement. Educators are making decisions using a wide variety of data about students, including state accountability test scores; interim progress test results; classroom tests, assignments, and homework; attendance, mobility, and grade-level progression rates; as well as dropout and graduation rates (Allensworth & Easton, 2007; Marsh et al., 2006; Hamilton et al., 2009). For those schools and districts that do manage to use data effectively, significant gains in student achievement are possible. For example, Carlson et al. (2011), in a multi-state randomized study covering 500 schools in 59 districts, found that a data-driven reform initiative resulted in statistically significant improvement in student math achievement. Faria et al. (2012) report that several research studies suggest that using a particular type of data–formative assessment–can result in student achievement gains (Black, Harrison, Lee, Marshall and Wiliam, 2002; Brookhart, 2001; Christman et al., 2009; Hayward, Priestley, & Young, 2004; Heritage, 2007; Shepard, 2005; Black and Wiliam 1998a, b).

Schools and districts face important challenges in implementing increased data use for instructional improvement. One key challenge is the need for teachers and administrators to have "data literacy"–the skills to analyze data, and to use a variety of data sources to refine and improve instruction (Halverson and Thomas 2007; Thessin 2007). Data systems and data initiatives have grown at a much faster pace than educator training around data use. This reality justifies the evaluation of a program such as TERC's Using Data, which aims to provide teachers with the needed training.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

Using school-level random assignment, this study seeks to estimate the causal impact of TERC's Using Data (UD) program on elementary teachers' understanding and use of data to inform mathematics instruction, and on the mathematics achievement of their fourth and fifth grade students. The main questions we seek to answer are:

1. At the end of the first year of the intervention, do fourth and fifth grade students in schools randomly assigned to the UD treatment program show greater levels of math achievement as compared to their counterparts in control schools?
2. After the first year of the intervention, do UD-trained teachers in treatment schools report greater data-use knowledge and skills and more-positive attitudes and beliefs about the value of data to inform instruction as compared to their same-grade counterparts in control schools?
3. After the first year of the intervention, do UD-trained teachers in treatment schools reportedly make more use of data or work more frequently with data in a collaborative setting, as compared to their same-grade counterparts in control schools? in comparison to control teachers?

We also consider a number of additional questions, including examining differences in effects on student achievement among different subgroups of students. A supplementary mixed-methods analysis will examine program implementation and will serve to contextualize the student achievement findings by describing the variation in program implementation across schools and suggesting some reasons for that variation.

Setting:
Description of the research location.

A large, urban district in the southeastern U. S. joined the study in March 2011, agreeing to implement the Using Data program. Mathematics test scores in the district lag behind those of the state overall, and behind those of other large urban districts in the state. More than half of students in the district are eligible for participation in the free or reduced price lunch program. Eight percent of students are Hispanic, and 15 percent have Individual Learning Plans.

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.

Sixty schools were randomized into the treatment group (30 schools) or the control group (30 schools), with an average of four teacher participants and one data coach from each school taking on the same treatment condition as their school. Our sample includes over 11,000 students and 800 math teachers in the 60 schools. Block randomization has given us a well-balanced sample across a number of important school characteristics, including students' prior test scores, racial composition, and eligibility for free or reduced price lunch (Table 2).

Intervention / Program / Practice:
Description of the intervention, program, or practice, including details of administration and duration.

The Using Data Program, which was developed with funding from an NSF grant by researchers at TERC and WestEd, provides professional development and technical assistance aimed at helping teachers improve instruction through the collaborative analysis of student data and development and implementation of action plans to address systemic learning problems. The UD program as tested in this study rolls out over a 2-year period. Data teams formed as part of the program consist of four participating teachers and a data coach. Data coaches are typically school-based service providers, such as instructional specialists, who work directly with a data team to lead them through the process of data analysis using collaborative inquiry.

Research Design:
Description of the research design.

The design is a block-randomized experiment with two levels of treatment—Using Data versus practice as usual. Randomization took place at the school level because in addition to the effect on the participating teachers, the intervention is designed to alter the school data-use culture by spreading knowledge, strategies, and perspectives on data use throughout the school. We chose to use a block-randomization design because of the diversity of the district. Schools were assigned an index score using principal components analysis with Title 1 status, percent of

students who are black, percent of students eligible for free or reduced-price lunch, and the school-wide 2010 state assessment math score as inputs. The schools were ranked according to index score, and randomized into the treatment or control groups by quartile.

Quantitative Analysis of Student Achievement. Quantitative analyses will include an estimate of the overall program effect on student math test scores for year 2, as well as subgroup analyses. The main confirmatory analyses will be conducted using a 2-level Hierarchical Linear Model (HLM) which nests students in schools and includes school-specific covariates and the student's prior test score. Since randomization was done at the school level, and the attrition rate is low, this is an experimental model. Additional subgroup and dosage analyses will be conducted using a quasi-experimental 3-level HLM that nests students with teachers, and teachers in schools, and includes teacher-level covariates. We also plan to use a 3-level instrumental variable model that uses the school treatment condition as an instrument for teacher participation in the program, in order to provide stronger causal conclusions with a model that includes teacher-level covariates (Gennetian et al. 2005).

Quantitative Analysis of Teacher Effects. The analysis plan will be similar to that for student achievement, using a quasi-experimental 2-level HLM with teachers nested in schools, except that the outcome variables will be teacher responses to surveys measuring their (a) attitudes and beliefs, and (b) knowledge and skills with respect to using data to improve instruction at the end of year 1, with treatment condition and teachers' baseline response included as model covariates, in addition to school-specific variables.

Qualitative Analysis of Implementation. Researchers have been reviewing and distilling interview, observation, and focus group data gathered from site visits to understand the degree to which the UD program is being implemented with fidelity. We plan to develop various metrics of implementation fidelity from qualitative data being collected from the 30 treatment schools, including observations of team meetings and interviews with key participants. Researchers have conducted site visits to 8 of the 30 treatment schools to understand the ways in which schools adapted the UD process to their unique contexts, as well as the contextual factors associated with their implementation approaches.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

Quantitative Student Achievement Data. We have received data from the district on both students (grades 4 and 5) and teachers. Student data includes demographic and socioeconomic variables and test scores on the state assessment math and reading exams, and identifies the math and reading courses taken by each student and their teacher for those courses. For teachers, we have information about their participation in the program and additional indicators such as experience and degree held (for participants).

Survey Data on Teacher Effects. CNA researchers developed and implemented two assessments pertaining to data use to improve instruction. The "Attitudes and Beliefs" instrument was developed to measure teacher agreement with core assumptions of the Using Data Process. The assessment also includes questions about data use, collaboration, and leadership support. The "Knowledge and Skills" assessment was developed to measure teacher knowledge of data use

concepts and application of data use skills when interpreting tables and graphs. These assessments were given at baseline and at the end of year 1 to teachers in the originally randomized sample and to replacement teachers. A final administration of the assessments will take place in May 2013. This final administration will also include an administration of the Learning Mathematics for Teaching (LMT) assessment, which measures teachers' ability to perform mathematics classroom tasks such as assessing student work, representing numbers and operations, and explaining common mathematical rules or procedures. LMT will be used to validate the Knowledge and Skills assessment.

Qualitative Implementation Data. Data gathered from site visits by researchers include observations of data team meetings, interviews of principals, data team members and district officials, as well as meeting logs from technical assistance sessions and data team meetings. Data that can be used to analyze implementation fidelity include participant attendance at all of the professional development workshops, facilitator ratings of team engagement at these workshops and at completed technical assistance training sessions, and the number, length, attendance, and subject matter at data team meetings.

Findings / Results:
Description of the main findings with specific details.

Preliminary analysis of pre/post survey data from teachers suggests that *collaborative* data use increased in the treatment schools relative to the control schools at statistically significant levels (Table 1). This finding is corroborated by interviews conducted over time in eight case study treatment schools, where teachers reported that the program had influenced their use of data primarily by giving them an analytic framework for discussing data with their colleagues.

Preliminary analysis after one intervention year suggests that student achievement in mathematics was no higher in the treatment schools than in the control schools (Table 2). We will have a more complete set of year 1 teacher and student results available by the date of the conference.

Conclusions:
Description of conclusions, recommendations, and limitations based on findings.

Preliminary results suggest that the Using Data intervention promoted increased collaboration among teachers. What is less clear is the amount of time that may be necessary for improved collaboration to promote improvements in practice that raise student achievement.

Both survey and interview data also revealed close alignment between the district's existing professional development emphasis on data use and the Using Data intervention. While such alignment may be desirable in advancing the cause of data use, experiential overlap between the treatment and control conditions (even if due to something other than participant noncompliance) complicates estimation of the treatment effect. These findings thus speak to the importance of documenting practices in both treatment and control settings. They also raise key questions about the amount of alignment that is desirable between an existing district culture and the professional development interventions deployed to strengthen that culture.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Allensworth, E., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago Public Schools: A close look at course grades, failures, and attendance in the freshman year*. Chicago: Consortium on Chicago School Research.

Black, P. J., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21.

Black, P. J., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice*, 5(1), 7–75.

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

Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education*, 8(2), 153–169.

Carlson, D., Borman, G.D., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Education Evaluation and Policy Analysis*, 33(3), 378-398.

Christman, J. B., Neild, R. C., Bulkley, K., Blanc, S., Liu, R., Mitchell, C., & Travers, E. (2009). *Making the most of interim assessment data: Lessons from Philadelphia*. Philadelphia: Research for Action. Retrieved January 22, 2012, from http://eric.ed.gov/PDFS/ED505863.

Faria, A., et al. (2012). *Charting Success: Data Use and Student Achievement in Urban Schools* Council of the Great City Schools and the American Institutes for Research. Summer 2012.

Gennetian, L. A., Morris, P. A., Bos, J. M., Bloom, H. S. (2005). Constructing instrumental variables from experimental data to explore how treatments produce effects. In H. S. Bloom, ed. *Learning More from Social Experiments*. New York: Russell Sage Foundation.

Halverson, R., & Thomas, C. N. (Eds.). (2007). *The roles and practices of student services staff as data-driven instructional leaders*. New York: Teachers College Press.

Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved January 22, 2012, from http://ies.ed.gov/ncee/wwc/pdf/practice_guides/dddm_pg_092909.pdf

Hayward, L., Priestley, M., & Young, M. (2004). Ruffling the calm of the ocean floor: Merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, 30(3), 397–415.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89(2), 140–145.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education: Evidence from recent RAND research*. Santa Monica, CA: RAND Corporation.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63(3), 66–71.

Thessin, R. A. (2007). Newton North High School gets smart about data. In K. P. Boudett & J. L. Steele (Eds.), *Data wise in action: Stories of schools using data to improve teaching and learning* (pp. 29-50). Cambridge, MA: Harvard Education Press.

Appendix B. Tables and Figures

Not included in page count.

Figure 1. Logic Model for Using Data Intervention

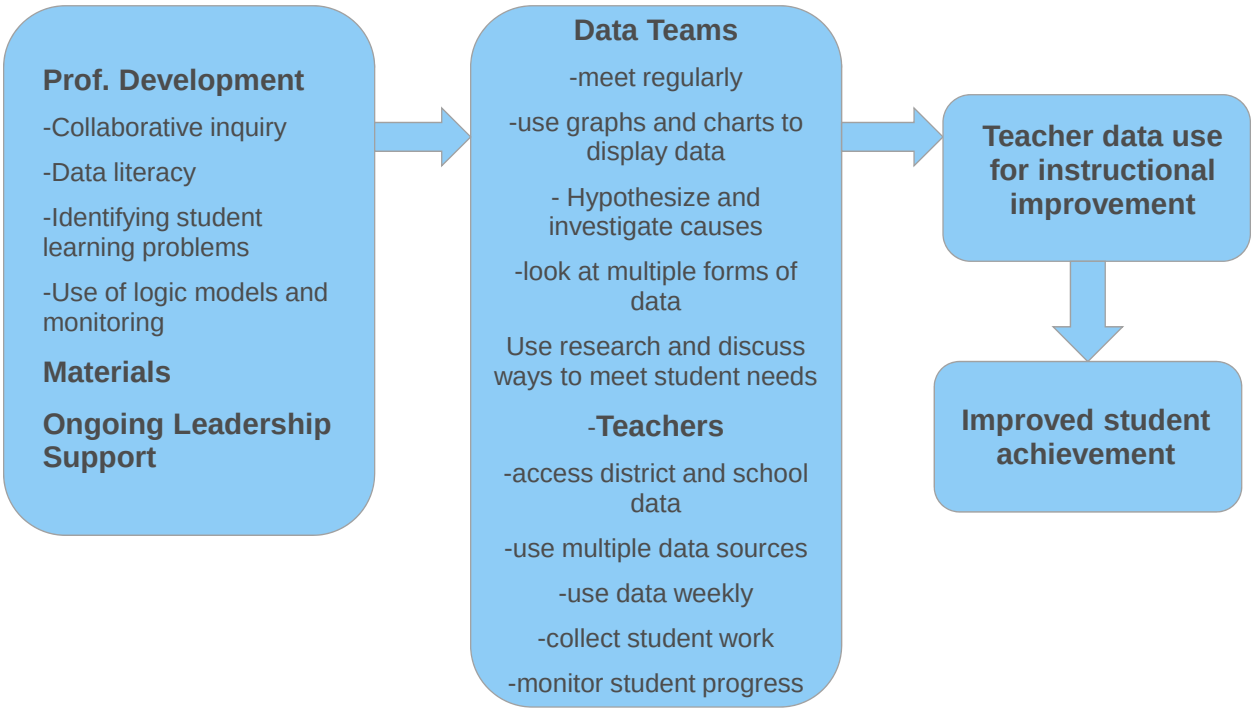

Table 1. Collaborative Data Use Questions – Spring Results

	Spring 2012 Survey			
In school year 2011-2012, how often did you work with data in the following contexts to make instructional decisions?	Control (n=109)	Treatment (n=115)	Diff.	p-value
a. on your own	3.51	3.67	0.16	0.05*
b. working with colleagues in your grade level	2.87	3.04	0.18	0.10
c. working with colleagues from other grade levels	1.91	2.30	0.40	0.00*
d. as part of a district-level activity with staff from other schools	1.65	1.93	0.28	0.00*
e. in another setting	1.63	1.86	0.23	0.02*
Data Use Collaboration Composite (Questions b to d)	6.42	7.27	0.85	0.00*

	Fall	Spring	Diff.	p-value
Treatment Composite	6.31	7.26	0.95	0.00*
Control Composite	6.32	6.47	0.15	0.39

Note: We asked teachers in both fall and spring to rank how often they use data in the following context on a scale from never, a few times, once or twice a month, or once a week or more. Answers were scored from 1 to 4 with one being never and four being once a week or more. In the fall there were no differences between treatment and control. In the spring, there were differences on items a, c, d, and e. Responses from items b, c, and d were added together to form a composite collaborative data use score. Higher scores represent more data use. Scores range from 3 to 9. There was no difference between treatment and control in the fall. There was a difference between the two groups in the spring; treatment participants scored almost one points higher than control participants.

Table 2. Student Characteristics by Treatment Condition, 2011-12

Student characteristic	**Percentage**	
	Control Schools	Treatment schools
Baseline variables		
Female	48.3%	48.5%
Currently receiving ELL services	3.0%	3.0%
Eligible for free or reduced-price lunch	56.1%	59.3%
Absentee rate	4.5%	4.5%
Outcome variables	**Mean**	
State Assessment Math score (2011-12)	220.0	219.0
n	5,511	5,713