

# Examining Post-Intervention Effects of a Randomly-Assigned School Leadership Professional Development Intervention

Robert Shand, Roger Goddard, and Yvonne Goddard

## Background and Context

Researchers and policymakers agree on the importance of quality school leadership for promoting positive school culture, fostering teacher effectiveness, and facilitating student achievement growth (Leithwood, et al., 2006, 2008; Robinson, et al., 2008). Less consensus exists on the mechanisms of those effects and whether school leader quality improvements can be induced by policy or intervention (Camburn, et al., 2007; Hallinger et al., 1996; Mulford, 2004). School leaders' professional development (PD) could serve as one such intervention. However, few studies have rigorously estimated the causal effects of school leader PD programs (Jacob et al., 2015; Camburn, et al., 2007; Goldring et al., 2007).

## Purpose and Research Questions

The purpose of this research is to examine lagged effects of leadership PD following the conclusion of a randomly assigned research-based PD for school principals. This proposal builds upon the initial findings of the program's impact immediately after the conclusion of the two-year training program (Authors, 2015). While RCTs provide rigorous estimates of an intervention's effect, decision-makers may not trust or rely upon research-based evidence (Henig, 2008; Slavin, 2002) because of concerns about applicability to individual contexts (Orland, 2009). Examining long-term effects of outcome measures make RCTs more informative and credible for decision-makers (Kyriakidi & Ioannidis, 2002; Llewellyn-Bennett, Bowman & Bulbulia, 2016; Rothwell, 2006).

Two concerns exist related to school leader PD. First, effects proximal to the intervention may fade without sustained follow-up. Second, long-term PD effects may be missed because they affect more distal outcomes or may take significant time to become established. These two countervailing effects exist in the literature on early childhood education, whereby immediate academic gains faded but more lasting effects appeared several years after the intervention concluded (Barnett & Husted, 2005).

This study extends previous work on a principal training RCT (Jacob et al., 2015) by examining principal, teacher, and student outcomes one year after intervention. To examine these effects, we pose two research questions:

- Do effects on principal attitudes and behaviors persist after intervention ceased?
- Do previously insignificant effects on teacher practice and student learning emerge after intervention ceased?

## Setting and Participants

The original study involved 126 public elementary schools in rural Michigan that were assigned to treatment and control groups. Ultimately, 95 schools participated (42 treatment, 53 control), and 91 schools concluded the study. One year after the study concluded, 82 school principals and

1,373 teachers in 90 schools participated in a follow-up survey. Baseline characteristics for treatment and control schools are shown in Table 1; the schools had above-average levels of students receiving subsidized lunch and below-average baseline test scores.

## **Intervention**

The intervention comprised ten, two-day professional development sessions over three years based upon McREL's Balanced Leadership framework (Waters et al., 2003). Participants engaged in PD that addressed leadership knowledge and skills in several domains. The intervention was hypothesized to have direct effects on the school's instructional climate, as well as principals' efficacy. It was also hypothesized to impact teacher practice and turnover, and, in turn, student achievement (Jacob et al., 2015).

## **Research Design**

This study extends the RCT via a one-year follow-up survey for principals and teachers. The primary issue facing a follow-up study of this nature is attrition. We examine follow-up survey results for the 82 principals who responded, which is approximately 10% lower than participants in the final year of the study itself. Table 2 compares baseline covariates for the 44 schools that attrited between random assignment and the follow-up survey to the 82 schools that participated. No statistically significant differences exist between the groups. We also compared values for treatment versus control groups among attriting schools to test for differential attrition and found only one difference that was marginally significant, in baseline math scores ( $p$ -value of 0.07). Analyses for principal and student test score outcomes were performed in a school-level regression framework including an indicator for treatment status and covariates to increase precision. Analyses for teacher-level outcomes use multi-level models, with teachers nested within schools and standard errors clustered at the school level.

## **Data Collection and Analysis**

Student learning outcomes comprise school-level results for math and ELA in grades 3-5 from the Fall 2012 state assessments, one full school year after the intervention concluded. Notably, as Jacob et al. (2014) demonstrated, when estimating the effects of school-level interventions in an RCT, the effect sizes and standard errors associated with school-level achievement data are substantively identical to those obtained with student-level data. Principal and teacher surveys were designed to capture elements of leadership, school climate, and practice. Twenty-one sub-scales were condensed into three aggregate factors supported by confirmatory factor analysis: principal leadership, school climate, and collaboration. Two additional sub-scales, principal efficacy and collective differentiated instruction, did not adequately fit the aggregated models so remained separate outcomes.

## **Findings**

Tables 3 and 4 illustrate one-year follow-up results on principal, teacher and student outcomes. Modest effects on principal attitudes and behaviors found at the intervention's conclusion (Jacob et al., 2015) faded substantially one year later, with no statistically significant intervention effects on behavior and effects on attitudes becoming only marginally significant at approximately half of the originally measured effect size. There are no statistically or

substantively significant lagged effects on teacher reports of leadership, school climate, or instructional practice. However, there are marginally significant lagged effects on fourth grade math scores, and statistically significant effects on fifth grade math scores. Effects on older but not younger students suggest that students present during the time of the intervention may have experienced lagged or cumulative effects, but these effects did not persist beyond the intervention to affect new students.

## **Conclusions**

These study results reinforce the view that “knowledge is not enough.” Achieving lasting changes in school leader quality that translate into school climate and instructional practice effects is challenging; without sustained effort, even modest results of PD on self-reported knowledge and attitudes fade. Methodologically, this study points to the importance of longer-term follow-up on RCTs, as concerns commonly raised about the long-term applicability of RCT results – fade-out and lagged effects –both appear in this study. The long-term impacts on student outcomes are noteworthy.

## References

- Barnett, W. S., & Hustedt, J. T. (2005). Head Start's lasting benefits. *Infants & Young Children, 18*(1), 16-24.
- Camburn, E. M., Goldring, E., May, H., Supovitz, J., Barnes, C., & Spillane, J. P. (2007). Lessons learned from an experimental evaluation of a Principal Professional Development Program. Retrieved from <http://www.cpre.org/lessons-learned-experimental-evaluation-principal-professional-development-program>
- Goldring, E., Camburn, E., Huff, J., Sebastian, J., & May, H. (2007). Effects of the National Institute for School Leadership: Early results from a randomized field trial. Retrieved from [http://www.cpre.org/images/stories/cpre\\_pdfs/NISL%20Effects%20Early%20Results%20Final%20AERA%2007\\_final.pdf](http://www.cpre.org/images/stories/cpre_pdfs/NISL%20Effects%20Early%20Results%20Final%20AERA%2007_final.pdf)
- Hallinger, P. & Heck, R. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly 32*, (1), 5-44.
- Henig, J.R. (2008). The evolving relationship between researchers and public policy. In Hess, F. M., editor, *When research matters: how scholarship influences education policy*, pages 41–62. Harvard Education Press, Cambridge, MA.
- Jacob, R. T., Goddard, R. D., & Kim, E. S. (2014). Assessing the use of aggregate data in the evaluation of school-based interventions: Implications for evaluation research and state policy regarding public-use data. *Educational Evaluation and Policy Analysis, 36*(1), 44-66.
- Jacob, R., Goddard, R., Kim, M., Miller, R., & Goddard, Y. (2015). Exploring the causal impact of the McREL Balanced Leadership Program on leadership, principal efficacy, instructional climate, educator turnover, and student achievement. *Educational Evaluation and Policy Analysis, 37*(3), 314-332.
- Kyriakidi, M., & Ioannidis, J. (2002). Design and quality considerations for randomized controlled trials in systemic sclerosis. *Arthritis Care & Research, 47*(1), 73-81.
- Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006 b). *Successful School Leadership - What it is and How it Influences Pupil Learning*. Nottingham: University of Nottingham.
- Leithwood, K., & Jantzi, D. (2008). Linking leadership to student learning: The contributions of leader efficacy. *Educational Administration Quarterly, 44*, 496–528.  
doi:10.1177/0013161X08321501

- Llewellyn-Bennett, R., Bowman, L., & Bulbulia, R. (2016). Post-trial follow-up methodology in large randomized controlled trials: a systematic review protocol. *Systematic Reviews*, 5(1), 214.
- Mulford, Bill. (2004) Congruence between the democratic purposes of schools and school principal training in Australia. *Journal of Educational Administration*, 42(6), 625-639
- Orland, M. (2009). Separate orbits: the distinctive worlds of educational research and policymaking. In Sykes, G., Schneider, B., and Plank, D. N., editors, *Handbook of education policy research*, pages 113–128. Routledge, New York, NY.
- Robinson, V. M., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly* 44(5), 635 - 674.
- Rothwell, P. M. (2006). Factors that can affect the external validity of randomised controlled trials. *PLoS Clinical Trials*, 1(1), e9.
- Slavin, R. E. (2002). Evidence-based education policies: transforming educational practice and research. *Educational Researcher*, 31(7):15–21.
- Waters, J. T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Aurora, CO: Mid-continent Research for Education and Learning.

Tables

Table 1. Descriptive Statistics (at random assignment)

Variable	Treatment mean (n=62)	Treatment SD	Control mean (n = 64)	Control SD
Enrollment	313.79	153.59	311.28	164.27
% minority students	0.10	0.16	0.12	0.14
% subsidized lunch	0.48	0.14	0.48	0.13
2007 4th grade math scores	427.38	6.26	425.53	9.84
2007 4th grade reading scores	429.55	10.83	430.36	14.16
Female principal	0.56	0.47	0.46	0.47
White principal	1	0	0.98	0.15
Principal with Master's degree or higher	0.92	0.28	0.89	0.31

Table 2. Attrition Analysis (126 at random assignment vs. 82 in year 4 follow-up)

Variable	Attritors (n=44)	Non- Attritors (n=82)	t value	p value
% minority students	0.12	0.11	0.47	0.64
% subsidized lunch	0.48	0.48	-0.19	0.85
2007 fourth grade reading	430.76	429.53	0.53	0.6
2007 fourth grade math	426.88	426.22	0.43	0.67

Table 3. One Year Post-Treatment Effects, Principal and Teacher Results

Outcome	b	SE	p value	Effect size
Principal reports (n=82)				
Principal efficacy	0.28~	0.16	0.078	0.2
Principal leadership	0.08	0.11	0.49	0.08
Collaboration	0.09	0.14	0.51	0.07
School climate	0.20~	0.12	0.09	0.19
Collective differentiated instruction	0.25	0.18	0.16	0.16
Teacher reports (1373 teachers in 90 schools)				
Principal leadership	-0.03	0.13	0.81	-0.03
Collaboration	-0.03	0.08	0.68	-0.04
School climate	0.01	0.07	0.83	0.02
Collective differentiated instruction	0.04	0.07	0.61	0.04

Notes: ~ p<0.1, \* p<0.05, \*\* p<0.01, \*\*\* p<0.001. All models include pre-treatment covariates: enrollment, percent of students who are minorities, and percent of students receiving free and reduced price lunch.

Table 4. Student Test Score Results

Outcome	n	b	SE	Effect size (ITT)
Grade 3 Math	108	-1.55	1.32	-0.103
Grade 4 Math	101	2.68~	1.49	0.177
Grade 5 Math	90	3.94*	1.97	0.195
Grade 3 Reading	107	-1.09	1.13	-0.08
Grade 4 Reading	101	1.20	1.50	0.08
Grade 5 Reading	90	1.85	1.40	0.13

Notes: ~ p<0.1, \* p<0.05, \*\* p<0.01, \*\*\* p<0.001. All models include pre-treatment covariates: Title I eligibility, Charter school, enrollment, percent minority, percent receiving free and reduced price lunch, and same grade, same- and other-subject pre-treatment scores.