Abstract Title Page

Title:

Making Sense of the "Zero Effect" of Comprehensive Teacher Induction Programs: A Mediational Analysis

Authors and Affiliations:

Yihua Hong and Guanglei Hong,
The University of Chicago

Abstract Body

Background / Context:

Teachers new to the profession may face various challenges and struggle with pedagogy and classroom management. They tend to be less effective in boosting student learning than their more experienced colleagues (Murnane & Phillips, 1981; Raymond, Fletcher, & Luque, 2001; Rivkin, Hanusheck, & Kain, 2001). Since the early 1980s, there has been an increasing recognition of the importance of providing induction support in forms of mentoring programs, workshops, orientation seminars, collaboration opportunities, and other support systems to new teachers in their initial years of teaching (Furtwengler, 1995). According to a policy paper by the New Teacher Center (Goldrick, Osta ,Barlin, & Burn, 2012), at present, 27 states require some forms of induction or mentoring support for new teachers, 22 states mandate completion of or participation in an induction program for advanced teaching certification, and 17 states provide dedicated funding for teacher induction.

While the general goal of teacher induction is to transform a student of teaching into a competent teacher of students, many evaluations in the past have focused on program impacts on novice teacher retention and professional well-being. Only a few studies have attended to instructional improvement as outcomes (see reviews by Ingersoll & Strong, 2011; Strong, 2009; Wang, Odell, & Schwille, 2008). Most studies (Davis & Higdon, 2008; Evertson & Smithey, 2000; Stanulis & Floden, 2009; Thompson, Paek, Goe, & Ponte, 2004) have suggested that more intensive mentoring and support from university-trained mentors might be associated with a higher rate of using effective instructional practices among new teachers. Yet one study (Roehrig, Bohn, Turner, & Pressley, 2008) reported that new teachers regardless of induction intensity declined in their use of effective teaching practices over the first year. These evaluations have been mostly non-experimental or quasi-experimental with a relatively small sample size. In contrast, a large-scale randomized study funded by the U.S. Department of Education and conducted by a research team from Mathematica Policy Research (Glazerman et. al, 2010) compared two prominent Comprehensive Teacher Induction (CTI) programs with standard district or school support for more than one thousand new teachers. Although teachers in the treatment group experienced more intensive, structured, and sequenced mentoring activities from trained external mentors, they exhibited surprisingly similar teaching practices as those in the control group in the spring of the first year such that a zero effect of the CTI programs was concluded.

In an attempt to reconcile the above mixed evidence, Ingersoll and Strong (2011) pointed out that almost all new teachers participated in some forms of induction regardless of treatment assignment. For example, in the Mathematica study, 83% of control teachers reported having a mentor, compared to 94% of the treatment group. In light of this observation, we consider alternative explanations for the zero-effect of the CTI programs.

- One possible explanation would be that neither the CTI programs nor the local induction services made any difference to the new teachers' classroom practices. This explanation is not highly plausible because it seems to be in contradiction with evidence from most research in the past.
- A second explanation would be that the participation rate in the CTI program activities in the treatment group was comparable to or even lower than the participation rate in local induction activities in the control group. Program effectiveness is expected to be offset by low participation (Ingersoll and Strong, 2011).

– A third explanation would be that the mentoring activities provided by the CTI programs to the treatment teachers, despite their higher intensity, were not as effective as the home-based mentoring activities provided to the control teachers. This is possible because the mentors selected, trained, and assigned by the externally administered induction programs did not possess school-specific knowledge and resources that might be particularly useful for new teachers (Rockoff, 2008). It is also possible that the mentors in the treatment group were new to the comprehensive induction programs while the home-based mentors tended to be familiar and experienced with the local programs (Ingersoll and Strong, 2011).

We argue that an investigation of these alternative causal mechanisms may reveal why the comprehensive teacher induction programs were not as "effective" as anticipated. Such understanding will have important implications for developing optimal induction programs for beginning teachers.

Purpose / Objective / Research Question / Focus of Study:

Reanalyzing data from the comprehensive teacher induction study, we aim to unpack the zero effect of the CTI programs on teaching practices by closely examining the content and activities of mentoring as potential mediators of the induction program effects on teaching practices. The content of mentoring includes teaching planning and preparation, management of classroom environment, instructional content and pedagogy, and professional responsibilities. Key activities for mentees include keeping record and analysis of teaching and student learning, working with a study group of teachers, observing other teachers' teaching, and meeting with local instructional leaders. We ask the following research questions:

1. Did treatment teachers and control teachers have different experiences with mentoring content and activities?
2. Did the differences in mentoring experiences mediate the program effect on teaching practices?
3. Was receiving mentoring from external mentors in the CTI programs as effective as receiving mentoring from home-based mentors under the control condition?

Setting:

The comprehensive teacher induction study selected large, urban, low-income public school districts in which the prevailing induction programs were not intensive, formal, or comprehensive. The two CTI programs chosen as the intervention were provided by either the Educational Testing Service (ETS) or the New Teacher Center at the University of California, Santa Cruz (NTC) for beginning teachers.

Population / Participants / Subjects:

The original sample includes 1,009 beginning teachers from 418 elementary schools in 17 districts. This study focuses on 614 reading/language arts teachers who responded to the survey about induction activities during 2005-2006 and who had classroom observation data available in spring 2006. Among them, 336 teachers (54.72%) received comprehensive teacher induction while 278 (45.28%) did not.

Intervention / Program / Practice:

The CTI programs featured intensive, structured, and sequenced support delivered by experienced, trained full-time mentors. In fall 2005, beginning teachers in the treatment group

were assigned to a full-time mentor with a 12 to 1 ratio. Mentors received ongoing training and a curriculum of materials to support the mentees' development. Mentees were provided weekly meetings with the mentor, monthly professional development sessions, opportunities to observe veteran teachers, and continuing evaluation of their practices. It was assumed that the breadth, intensity, and nature of induction services would differ on average between the CTI programs and the home-based informal arrangements for supporting beginning teachers.

Research Design:

This is a multi-site cluster randomized trial. Seventeen school districts in thirteen states participated in the study. Within each district, elementary schools participating in the study were assigned at random to either the treatment group, in which case eligible teachers at the school were offered comprehensive teacher induction provided by either ETS or NTC based on district preferences, or to the control group, in which case teachers at the school took part in the districts' usual set of induction services. Ten districts received one year of comprehensive induction services; the other seven districts received two years of such services.

Data Collection and Analysis:

The outcome of interest is classroom teaching practices. Literacy teachers received a one-time classroom observation in late spring 2006 from raters who were blind to the treatment status of the classrooms. The instrument (the Diagnostic Classroom Observation) measures the teacher practices that current research suggests are essential to good teaching or that have been linked to student achievement growth. Observers scored teachers on five-point scales in each of three domains — implementation of a literacy lesson, content of a literacy lesson, and classroom culture. The inter-item reliability coefficients for the three summary scores are .89, .80, and .93.

Beginning teachers received a background teacher survey in fall 2005 and were requested permission to obtain their college entrance exam scores. Yearly surveys of teacher induction activities were administered to both treatment and control teachers. We consider induction content and activities reported by each beginning teacher as mediators. Measures include number of mentors, dosage of induction content, intensity of mentoring activities including class observation and feedback provided by the mentor, principal, or other local instructional leaders. These measures were derived from the teacher survey administered in spring 2006.

Our analysis consists of three phases. In phase I, we examine the variation in the outcome and mediator measures between the treatment group and the control group and within each group. In phase II, we examine whether the content and activities of induction experienced differently by beginning teachers in the treatment group and those in the control group mediate the program effects on teaching practices. In phase III, we compare the effectiveness of specific mentoring elements provided by the CTI programs with the effectiveness of their counterparts provided by the local induction programs.

The analyses in phase I are protected by the clustered randomized design. We apply multivariate hierarchical linear models to the three summary measures of classroom practices and compare the program effects across the three domains of classroom practices. To estimate the program effects on the multi-item measures of induction content and activities, we embed measurement models at level 1 and compare the program effects on different latent constructs of mentoring at the teacher level. In these analyses, we additionally examine the variation in program effects across districts because schools were randomized within districts.

The internal validity of the analyses in phases II and III is threatened by potential mediator-outcome confounding. Utilizing rich information obtained from beginning teachers' background survey, for each mediator measure, we estimate a teacher's conditional probability of displaying a certain mediator value under the treatment condition that the teacher has been actually assigned to and that under the counterfactual condition as functions of the pretreatment covariates. Using the ratio-of-mediator-probability weighting (RMPW) method (Author, 2010a; Authors, 2011; Authors, 2012), in phase II we decompose the total effect of the CTI programs on teaching practices into a natural indirect effect channeled through the program-induced change in a particular mediator and a natural direct effect. The RMPW method allows the mentoring effect on the outcome to be different across the treatment conditions without making assumptions about the functional form of the outcome model. This is an important strength not shared by conventional methods such as structural equation modeling.

In phase III, we use marginal mean weighting through stratification (MMW-S) (Author, 2010b, 2012; Authors, 2009), to remove selection bias in estimating the effects of specific mentoring elements on teaching practices in the treatment group and compare with the effects of similar mentoring elements in the control group. We put particular emphasis on detecting thresholds in estimating the dosage effects.

Findings / Results:

Here we present the preliminary results from the phase I analysis. We have identified significant differences in the number of mentors (coefficient =.34, *se* = .07, $t = 4.79$. $p < .001$) and in the intensity of mentoring activities (coefficient = 16.90, s*e* = 7.54, $t = 2.24$. $p < .05$) between the treatment and control groups. On average, treatment teachers received about 1.42 mentors (*sd* = .92) and had 62.79 (*sd* = 74.61) minutes per week meeting with the mentors. Control teachers had1.08 mentors (*sd* =.84) and spent 45.90 (*sd* = 110.89) minutes per week with mentors. Tables 1-3 show that the treatment and control groups also differed significantly in dosage of induction content, engagement in mentoring activities, and exposure to class observations and feedback.

Conclusions:

The results from our preliminary analysis indicate that treatment teachers and control teachers had different experiences with mentoring content and activities. Clearly, beginning teachers assigned to the CTI programs tend to receive a higher dosage of induction content and a higher intensity of mentoring activities. Therefore, we can rule out the second explanation for the zero effect of the CTI programs given that the treatment teachers displayed an equal or higher rate of participation than did the control teachers. However, we notice that a higher level of participation rate in the treatment group apparently did not lead to superiority in teaching practices in comparison with the control group. One would wonder, had the treatment teachers participated in the CTI programs at a lower rate that becomes equal to the control teachers' participation rate in their local induction programs, whether the teaching practices of the treatment group would become inferior to that of the control group. The causal mediation analysis in phase II will provide a direct answer to this question. The analyses in phase III will identify the thresholds in estimating the mentoring effect under each program and will compare the effectiveness of support from external mentors in the CTI programs with that from home-based mentors under the control condition.

Appendices

Appendix A. References

Author. (2010a).
Author. (2010b).
Author. (2012).
Authors. (2009).
Authors. (2011).
Authors. (2012).
Davis, B., & Higdon, K. (2008). The effects of mentoring/induction support on beginning teachers. *Journal of Research in Childhood Education*, *22*, 261–274.
Evertson, C. M., & Smithey, M. W. (2000). Mentoring effects on protégés' classroom practice: An experimental field study. *Journal of Educational Research*, *93*, 294–304.
Furtwengler, C. B. (1995). Beginning teacher programs: Analysis of state actions during the reform era. *Education Policy Analysis Archives* 3(3), 1-20.
Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., Jacobus, M. (2010*). Impacts of comprehensive teacher induction: Final results from a randomized controlled study* (NCEE 2010-4027). Washington DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
Goldrick, L., Osta, D., Barlin, D., & Burn, J. (2012). *Review of State Policies on Teacher Induction.* Santa Cruz, CA: New Teacher Center.
Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, 81(2), 201-233.
Murnane, R. J., & Phillips, B. R. (1981). What do effective teachers of inner-city children have in common? *Social Science Research, 10*, 83–100.
Raymond, M., Fletcher, S. H., & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas.* Stanford, CA: Stanford University, Center for Research on Education Outcomes.
Rivkin, S., Hanushek, E., & Kain, J. (2001). *Teachers, schools, and academic achievement*. Working Paper Number 6691. Cambridge, MA: National Bureau of Economic Research.
Rockoff, J. E. (2008). *Does mentoring reduce turnover and improve skills of new employees? Evidence from teachers in New York City* (Working Paper 13868). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/ w13868
Roehrig, A. D., Bohn, C. M., Turner, J. E., & Pressley, M. (2008). Mentoring beginning primary teachers for exemplary teaching practices. *Teaching and Teacher Education, 24, 684–702.*
Stanulis, R. N., & Floden, R. E. (2009). Intensive mentoring as a way to help beginning teachers develop balanced instruction. *Journal of Teacher Education*, *60*, 112–122.
Strong, M. (2009). *Effective teacher induction and mentoring: Assessing the evidence*. New York, NY: Teachers College Press.
Thompson, M., Paek, P., Goe, L., & Ponte, E. (2004). *Study of the impact of the California Formative Assessment and Support System for Teachers: Report 2: Relationship of*

BTSA/CFASST engagement and teacher practices (ETS-RR-04–31). Washington, DC: Educational Testing Service.

Wang, J., Odell, S. J., & Schwille, S. A. (2008). Effects of teacher induction on beginning teachers' teaching: A critical review of the literature. *Journal of Teacher Education*, *59*, 132–152.

Appendix B. Tables and Figures

Table 1 *Dosage of Induction Content Received by Mentees*

Mentoring Guidance	Comprehensive Teacher Induction *M (SD)*		
	Participants (n = 336)	Non-participants (n = 278)	All (n = 614)
Planning and preparation			
assessing district and community resources	.73 (.91)	.45 (.76)	.60 (.86)
working with other teachers to plan instruction	.60 (.87)	.58 (.87)	.59 (.87)
selecting or adapting curriculum materials	.85 (.92)	.52 (.81)	.70 (.89)
understanding/teaching toward state or district standards	.93 (.98)	.53 (.83)	.75 (.94)
planning lessons	.86 (.97)	.55 (.88)	.72 (.94)
using student assessments to inform teaching	.86 (.89)	.46 (.78)	.68 (.86)
Classroom environment			
managing classroom activities, transitions and routines	1.06 (.97)	.62 (.88)	.86 (.96)
managing student discipline and behavior	1.07 (1.05)	.68 (.94)	.90 (1.02)
motivating students	.94 (.98)	.53 (.81)	.75 (.93)
Instruction			
teaching reading/language arts	.97 (.95)	.59 (.86)	.80 (.93)
teaching mathematics	.76 (.88)	.54 (.85)	.66 (.87)
teaching children with varying levels of achievement/ability	.97 (.99)	.57 (.86)	.79 (.96)
using multiple instructional strategies/techniques to teach students	.99 (.99)	.57 (.87)	.80 (.96)
reviewing and assessing student work	.95 (.96)	.47 (.80)	.73 (.92)
teaching English language learners	.54 (.81)	.27 (.56)	.42 (.72)
teaching special needs students	.61 (.89)	.33 (.74)	.49 (.84)
teaching students of varying ethnic/racial and socioeconomic backgrounds	.73 (.92)	.46 (.83)	.60 (.89)
Professional responsibilities			

understanding the school's culture, policies, and practices	.96 (1.01)	.72 (.91)	.85 (.97)
completing paperwork	.84 (.95)	.61 (.84)	.74 (.91)
working with other school staff	.60 (.83)	.46 (.81)	.54 (.82)
reflecting on instructional practices	1.20 (.96)	.57 (.88)	.91 (.98)
communicating with parents	.61 (.87)	.47 (.80)	.55 (.84)
Total	18.97 (17.18)	10.21 (14.05)	14.96 (16.39)

Table 2 *Mentees' Engagement in Learning Activities*

	Comprehensive Teacher Induction %		
Engagement in Learning Activities	Participants (n = 336)	Non-participants (n = 278)	All (n = 614)
Keep a written log or record of reflections on your teaching practices	37.16	26.71	32.40
Keep a portfolio or record of student work and an analysis of that work	77.34	75.45	76.48
Work with a study group of new teachers	70.78***	28.16	51.40
Work with a study group of new and experienced teachers	50.00***	35.02	43.16
Observe other teachers teaching in their classrooms	71.30***	39.35	56.74
Observe someone else teaching your class	48.64**	37.18	43.42
Meet with the principal to discuss your teaching	67.67	69.68	68.59
Meet with a literacy or mathematics coach or other curricular specialist	71.52	72.2	71.83
Meet with a resource specialist to discuss needs of particular students	64.13	65.7	64.85

* p<.05 **p<.01 ***p<.001

Table 3 *Times of Class Observation or Feedback Received by Mentees over 3 months*

	Comprehensive Teacher Induction *N (SD)*		
	Participants (*n*=336)	Non-participants (*n*=278)	All (*n*=614)
Observation by mentor	2.81 (1.53) ***	1.29 (1.65)	2.12 (1.76)
Observation by principal	1.80 (1.38)	1.81 (1.44)	1.80 (1.41)
Feedback on teaching (not as part of a formal evaluation process)	2.20 (1.50) ***	1.79 (1.47)	2.01 (1.50)
Feedback on teaching (part of a formal evaluation process)	1.61 (1.28)	1.45 (1.18)	1.53 (1.23)
Feedback on lesson plans	1.35 (1.46)	1.41 (1.58)	1.38 (1.51)

* $p<.05$ **$p<.01$ ***$p<.001$