

Who Knows if our Teachers are Prepared? Three Different Perspectives on Graduates' Instructional Readiness and the Features of Preservice Preparation that Predict them

AUTHORS

Matthew Ronfeldt

University of Michigan

**Kavita Kapadia
Matsko**

National Louis University

**Hillary Greene
Nolan**

University of Michigan

Michelle Reininger

Stanford University

ABSTRACT

This study follows 305 preservice teachers (PSTs) who student taught in Chicago Public Schools (CPS) in 2014-15 and were subsequently hired in CPS in 2015-16. Drawing on surveys of PSTs and their cooperating teachers (CTs) and CPS administrative data, we linked features of their preservice teacher education to three outcome measures for instructional readiness: (i) CTs' perceptions of their PSTs' preparedness to teach at the end of student teaching, (ii) PSTs' self-perceptions of preparedness at the end of student teaching, as well as (iii) PSTs' first-year observation ratings based on the district evaluation rubric. We first investigated whether these outcome measures were related to one another. We found that CTs' perceptions of PSTs' preparedness positively predict first-year observation ratings but PSTs' self-perceived preparedness were unrelated. We then examined which features of preparation were positively related to these three outcomes and whether or not the same features predicted all three outcomes. PSTs received stronger first-year observation ratings when their CTs had stronger observation ratings themselves, their CTs reported providing stronger coaching in specific instructional areas, they gained employment in the same school in which they had completed their student teaching, and when their placements had been in self-contained elementary classrooms. However, these same features were generally unrelated to our two other survey-based outcome measures, while others emerged as important. This study then provides suggestive evidence that different features of preparation are likely promising levers for program improvement but that program leaders and policymakers must first consider which outcome might determine "improvement," as different features appear to be associated with different measures of instructional readiness.

Acknowledgements: We are grateful to the Spencer Foundation for its generous financial support of this research. We are also grateful to the Chicago Public Schools and the University of Chicago Consortium on School Research for their ongoing partnership and contributions, and particularly to Matt Lyons, Molly Gordon, Josh Klugmann, and Jennie Jiang. All ideas expressed in this article should be attributed only to the authors.

VERSION

January 2018

Suggested citation: Ronfeldt, M., Matsko, K.K., Greene Nolan, H., & Reininger, M. (2018). Who Knows if our Teachers are Prepared? Three Different Perspectives on Graduates' Instructional Readiness and the Features of Preservice Preparation that Predict them (CEPA Working Paper No.18-01). Retrieved from Stanford Center for Education Policy Analysis: <http://cepa.stanford.edu/wp18-01>

Who Knows if our Teachers are Prepared?

Three Different Perspectives on Graduates' Instructional Readiness and the Features of Preservice Preparation that Predict them

Matthew Ronfeldt

University of Michigan

Kavita Kapadia Matsko

National Louis University

Hillary Greene Nolan

University of Michigan

Michelle Reininger

Stanford University

ABSTRACT: This study follows 305 preservice teachers (PSTs) who student taught in Chicago Public Schools (CPS) in 2014-15 and were subsequently hired in CPS in 2015-16. Drawing on surveys of PSTs and their cooperating teachers (CTs) and CPS administrative data, we linked features of their preservice teacher education to three outcome measures for instructional readiness: (i) CTs' perceptions of their PSTs' preparedness to teach at the end of student teaching, (ii) PSTs' self-perceptions of preparedness at the end of student teaching, as well as (iii) PSTs' first-year observation ratings based on the district evaluation rubric. We first investigated whether these outcome measures were related to one another. We found that CTs' perceptions of PSTs' preparedness positively predict first-year observation ratings but PSTs' self-perceived preparedness were unrelated. We then examined which features of preparation were positively related to these three outcomes and whether or not the same features predicted all three outcomes. PSTs received stronger first-year observation ratings when their CTs had stronger observation ratings themselves, their CTs reported providing stronger coaching in specific instructional areas, they gained employment in the same school in which they had completed their student teaching, and when their placements had been in self-contained elementary classrooms. However, these same features were generally unrelated to our two other survey-based outcome measures, while others emerged as important. This study then provides suggestive evidence that different features of preparation are likely promising levers for program improvement but that program leaders and policymakers must first consider which outcome might determine "improvement," as different features appear to be associated with different measures of instructional readiness.

ACKNOWLEDGEMENTS: We are grateful to the Spencer Foundation for its generous financial support of this research. We are also grateful to the Chicago Public Schools and the University of Chicago Consortium on School Research for their ongoing partnership and contributions, and particularly to Matt Lyons, Molly Gordon, Josh Klugmann, and Jennie Jiang. All ideas expressed in this article should be attributed only to the authors.

RECOMMENDED CITATION: Ronfeldt, M., Matsko, K.K., Greene Nolan, H., & Reininger, M. (2018). Who Knows if our Teachers are Prepared? Three Different Perspectives on Graduates' Instructional Readiness and the Features of Preservice Preparation that Predict them (CEPA Working Paper No.18-01). Retrieved from Stanford Center for Education Policy Analysis: <http://cepa.stanford.edu/wp18-01>

INTRODUCTION

In recent decades, teacher preparation programs have been faced with mounting pressures to demonstrate the efficacy of their particular approaches to preparation. The search for an empirical basis upon which to make the case for program effectiveness was heightened in the 1990s when alternative pathway policies gained momentum, and university-based preparation pathways faced increased scrutiny. At the time, no such empirical base existed to link preservice preparation to graduates' workforce outcomes. The past two decades, however, have seen a steady increase in large-scale studies which attempt to better understand the impact of teacher preparation on their graduates.

Initially, most studies of this kind tested whether graduates from different pathways or programs had, on average, better value-added to student achievement and yielded mixed results. Given substantial variation in preparation within pathways and programs (Boyd et al., 2012), these inconclusive findings are not altogether surprising. More recently, scholars have shifted focus away from whether certain pathways or programs are better than others to investigate instead which features of preparation -- varying within and across pathways and programs -- are associated with better instructional quality. Many of these studies linked features of preparation to survey-based measures of graduates' self-perceived preparedness to teach as the focal outcome. While these studies indicated which features of preparation made teachers *feel* better prepared, they failed to tell us whether or not these same features were related to actually becoming more instructionally effective. To our knowledge, no prior studies have examined convergent validity -- that measures of self-perceived preparedness are positively related to observable measures of instructional effectiveness.

The present study addresses these gaps in the literature in a few ways. Specifically, it tests whether (i) preservice student teachers' (PSTs) perceptions of their instructional readiness or (ii)

cooperating teachers' (CTs) perceptions of their PSTs' instructional readiness are positively associated with (iii) graduates instructional effectiveness as first-year teachers as measured on observation rubric-based district evaluations. We find that PSTs who say they feel better prepared are no more or less effective on their first-year district evaluations; on the other hand, those PSTs who were evaluated as more instructionally ready by their CTs received significantly stronger first-year district evaluations. In other words, our study provides some evidence of convergent validity for CT ratings of PST instructional readiness but not for PST self-perceived preparedness.

Additionally, we find that PSTs are evaluated more instructionally effective by school leaders using a district observation rubric during their first year of teaching (time = t) when their CTs were rated as more instructionally effective (time = t-1), when their CTs reported providing stronger mentoring around the instructional domains evaluated on the district rubric, and when they were hired into the same schools in which they completed their student teaching. However, these same features of preparation were mostly unrelated to our other, survey-based measures for PST instructional readiness, ultimately suggesting that different features of preparation may be related to different PST outcomes.

LITERATURE REVIEW

How should we determine whether or not teacher preparation makes a difference? Only in the past two decades have scholars tried to answer this question by linking different preparation programs or features of preparation to workforce outcomes among graduates, including employment, retention, and measures of instructional effectiveness (e.g., value-added and observational ratings). One reason these kinds of analyses are relatively new is that only recently has it become possible in some labor markets to link preservice programs to centralized administrative and evaluation data on program graduates. In this review, we describe research undertaken in

connection to three sets of outcomes: self-perceived preparedness (survey-based measures), value-added to student achievement measures, and observational ratings.

Self-Perceived Preparedness (Survey-Based Measures). Prior to studies linking preparation programs to graduates' administrative or evaluation data, a number of scholars used large-scale survey administrations to recent program graduates as a way to assess whether or not teacher preparation made a difference. Notably, Darling-Hammond, Chung and Frelow (2002) used survey data on 3,000 beginning teachers in New York City and found that graduates from traditional programs felt better prepared than graduates from alternative programs or teachers who had no formal training. A limitation of this work, though, is that it assumed that average self-reported feelings of preparedness are adequate measures for assessing program quality. The authors acknowledged that linking preparation to direct measures of instructional effectiveness would be preferable, but they also argued that self-perceived preparedness was a worthwhile outcome because they found it to predict measures of teacher efficacy which, in prior work, had been found to be related to student achievement gains (Hammond, Chung, & Frelow; 2002). In subsequent work, scholars have found survey-based measures of self-perceived preparedness to be positively related to teachers' career plans, self-efficacy, and early-career retention (Ronfeldt, Schwartz, & Jacob, 2014; Ronfeldt & Reininger, 2012) but none, to our knowledge, have linked self-perceived preparedness to observed instructional quality as measured, for example, by value-added to student achievement measures or observational evaluations.

Despite this limitation, a number of scholars, including authors on the present study, followed Darling-Hammond, Chung, and Frelow (2002) in using self-perceived preparedness among graduates to further assess the effects of preservice preparation. Rather than look for average differences between programs or pathways, though, subsequent studies investigated which features of preparation – varying within and across programs and pathways – were related to graduates'

average feelings of preparedness and self-efficacy. In Chicago, for example, Ronfeldt and Reiningger (2012) examined whether having more or better quality student teaching experiences predicted graduates feeling better prepared, having stronger self-efficacy, or planning longer teaching careers. They found that graduates who reported better quality clinical experiences fared better across outcomes but that number of weeks of student teaching were mostly unrelated to all outcomes. In a follow-up study aiming to identify aspects of student teaching that contributed to its quality, Ronfeldt, Reiningger, and Kwok (2013) found that PSTs who reported better quality CTs and more autonomy over instructional decisions felt better prepared and had stronger teacher self-efficacy. Given results from this study suggested CTs contribute to the quality of student teaching experiences, Matsko, Ronfeldt, Greene Nolan et al. (conditionally accepted) drew upon surveys of more than 1,000 CTs in the Chicago area to better understand whether and how their roles as models and coaches predicted PSTs' feelings of preparedness in different instructional areas. The study found that PSTs felt better prepared when their CTs were more instructionally effective (as measured by a district evaluation system and by the PSTs themselves) and when they rated the quality of the mentoring they received from their CTs as stronger and more frequent. Finally, using nationally representative data, Ronfeldt, Schwartz and Jacob (2014) found that teachers who reported completing more methods-related courses and more weeks of student teaching felt better prepared during their first year of teaching and were more likely to stay in teaching.

Value-Added to Student Achievement Measures (VAMs). As it became more viable in some labor markets to link preservice programs to administrative and evaluation data on program graduates, a number of studies began to use graduates' VAM scores as a way to examine the effects of preparation. As with research using survey-based measures of self-perceived preparedness, early studies using VAMs focused on comparing pathways (e.g., alternative versus traditional) or programs

in terms of average differences but, over time, studies began to consider which features of preparation, varying within and across programs and pathways, predicted better VAMs.

Among studies comparing pathways and programs in terms of average VAMs, results were mixed. Most studies suggested that there were few, if any, significant or meaningful differences between programs and pathways in terms of graduates' average VAMs (Constantine et al., 2009; Goldhaber et al., 2013; Koedel et al., 2015, von Hippel et al., 2014). However, other studies indicated that some programs and pathways had significant and meaningful effects on graduates' VAMs (Boyd et al., 2009; Darling-Hammond et al., 2005; Glazerman, Mayer, & Decker, 2006; Henry et al., 2014). In a new analysis, von Hippel and Bellows (2017) re-analyzed estimates from prior studies based in six states, including many cited above, and concluded that, in some cases, previous research may have used statistical practices that exaggerated differences, including underestimating standard errors and not accounting for multiple comparisons. The authors concluded that typically only one program per state differed significantly from the state average.

Other studies that have linked features of preparation to graduates' VAMs have found that certain features are associated with better instructional effectiveness. In their groundbreaking study, Boyd et al. (2009) used graduate surveys and review of program documents to gather information on features of preparation of all major providers of teachers to New York City. They found that graduates from programs with more oversight of clinical experiences and with more practice-based opportunities to teach had better VAMs.

Focusing specifically on characteristics of the field placement schools used during student teaching experiences, Ronfeldt (2012) found that NYC teachers who student taught in schools with better average teacher retention – a signal for better teacher working conditions – had better VAMs and were more likely to remain teaching in NYC schools. In a follow up study in a different urban district, Ronfeldt (2015) found similar positive relationships between field placement school average

retention rates and graduates' VAMs. Additionally, he found that graduates had better VAMs when they learned to teach in field placement schools with better average teacher collaboration and a history of better achievement gains.

Other studies have examined effects of coursework. In Monk's (1994) study using the nationally-representative Longitudinal Study of American Youth data, he found that teachers who completed more pedagogy and content coursework during preservice preparation had better achievement gains. By contrast, Harris and Sass (2006) found preservice coursework and degrees to be unrelated to math and reading VAMs of teachers in Florida; they argued that prior work, including that by Monk (1994), did not adequately account for various forms of selection. Constantine et al. (2009) also found the amount and kinds of preservice coursework to mostly be unrelated to later achievement gains. Additionally, they found that alternatively certified teachers who completed coursework while teaching had lower math achievement gains than traditionally certified counterparts, suggesting that completing coursework while teaching may be detrimental to student achievement.

Studies based upon graduates' VAMs have provided important initial evidence of relationships between teacher preparation and graduates' instructional effectiveness. However, there are a number of limitations of using VAMs for these purposes, including that VAMs are available for only a minority of graduates in tested grades and subjects and VAMs are indirect measures of teachers' instructional quality. By contrast, observational evaluations, which we consider next, are available for most or all teachers in some states and are direct measures of instructional quality.

Observational Ratings. Despite being available for most teachers, we are aware of only two studies that have linked preservice preparation to graduates' observational ratings. Ronfeldt and Campbell (2016) found that about 20 percent of teacher preparation providers in Tennessee could be distinguished statistically from the state mean; additionally, highest-quartile providers graduated

teachers who performed as though they had an additional year of teaching experience, on average, as compared with lowest-quartile providers. Similarly, researchers in North Carolina found that some programs differed significantly from the state mean in terms of graduates' average observational ratings (Bastian, Patterson, & Pan, 2017). To our knowledge, no existing research has linked features of teacher preparation to observational ratings of graduates – an important contribution of the present study.

Intending to address a number of gaps in the existing literature, this study asks:

1. Are PSTs who, at the end of preparation, (a) report feeling better prepared or (b) are rated as more instructionally effective by their CTs actually more instructionally effective during their first year of teaching (as measured by district observational ratings)?
2. What features of preparation predict how instructionally effective PSTs (a) feel at the end of preparation, (b) are evaluated by their CTs, and (c) are rated by district evaluators during their first year of teaching?
3. Do the same features of preparation that predict one of these outcomes also predict the others?

METHODS

Setting

This study takes place in the context of Chicago Public Schools (CPS), which serves about 400,000 predominantly Latino and African American students each year.¹ CPS is also the site of clinical or student teaching placements for nearly 50 university-based teacher preparation programs. As a result, hundreds of PSTs receive their clinical training in CPS each year through their work in CTs' classrooms. PSTs register to student teach in CPS through a centralized registration process

¹ CPS At a glance (website) http://cps.edu/About_CPS/At-a-glance/Pages/Stats_and_facts.aspx

that maintains information on PSTs, CTs, and preparation programs. Recently, CPS began evaluating the instructional effectiveness of all teachers. This study makes use of information on the instructional effectiveness of CPS teachers, specifically the CTs that mentored PSTs in their student teaching placements and the PSTs who student-taught and then were hired as first-year teachers.

Data

We administered surveys to PSTs and CTs during student teaching in the 2014-15 school year. All registered PSTs received pre- and post-student teaching surveys; CTs were invited to complete surveys about their PSTs. Survey administration timelines and response rates are listed in Appendix Table 1.² Surveys were sent to PSTs via email prior to the start of the fall and spring terms. Our research team sent post-student teaching surveys by email to all registered PSTs toward the end of each term.³ Survey completers were offered a \$25 gift card.

Using registration data and additional CT data collected by CPS, we identified the CTs for all registered PSTs and sent them individualized online surveys at the end of the fall and spring terms.⁴ Survey completers were offered a \$50 gift card. We linked CTs and their survey responses to CPS personnel and evaluation data, as well as to data about their schools.⁵ For those hired in CPS in the subsequent year, we could also link PSTs to administrative data on the schools in which they were employed in 2015-16 and to their district evaluation data for their first year of teaching.

Sample

Of our initial population of 1,122 PSTs who completed student teaching in CPS during the 2014-15 school year, 305 subsequently gained employment in CPS during the 2015-16 school year and could be linked to their first-year observation ratings; this constituted our main analytic sample

² To accommodate multiple start dates and placement lengths for PSTs, which vary by program, CPS administered pre-student teaching surveys to PSTs twice during the year to incoming student teachers.

³ PSTs who had placements that were more than a term in length and with two different CTs received the survey twice.

⁴ Mentors who worked with multiple student teachers were sent a separate survey for each student teacher.

⁵ Where PSTs could not be linked to CTs (and their schools) we used registration information to identify their field placement schools; thus, we were able to link some PSTs to CPS data on their field placement schools even where CT information was missing.

for this study.⁶ Of these 305 PSTs who can be linked to first-year observation ratings, 225 could be linked to measures of PSTs' self-perceived preparedness based on post-student teaching surveys and 226 could be linked to measures of CTs' perceptions of how well prepared they felt their PSTs were based upon CT surveys.⁷

Pre-Service Teacher Characteristics. Table 1 (top) summarizes characteristics of PSTs. PSTs in our analytic sample were mostly female (72%) and White (57%), with another 22% Latino, 10% Black, 6% Asian, and 6% Other. About 29% graduated from a CPS high school themselves, and the average undergraduate GPA was the equivalent of an A-. About one-third (31%) had prior teaching experience, such as substitute teaching. During student teaching, PSTs were an average age of 26.2 years old, and 12% were parents. Appendix Table 2 (top) compares PSTs in our analytic sample (individuals who student taught in CPS, were employed in CPS the following year, and could be linked to first-year observational ratings) with PSTs in the non-analytic sample (those who student taught in CPS but were not employed in CPS the following year and/or could not be linked to first-year observational ratings). Generally, the analytic sample had more Latino PSTs, more CPS graduates, PSTs with slightly higher undergraduate GPAs, and fewer PSTs who identified as 'other' race/ethnicity. This suggests that the analytic sample may not be representative of all PSTs who

⁶ Actually, n=321 of the original sample (n=1122) could be linked to observation ratings; however, we had no information about preservice preparation (including surveys) for 16 of these individuals, so they would have dropped from all analyses; thus, we focus on n=305 individuals for whom we have information about their preparation. We considered using as our analytic sample the 178 PSTs who had complete data on all three of our focal outcome measures. However, doing so would have cut the number of PSTs with observation ratings by almost one-half. Since a primary contribution of this paper is that it links features of preparation to observational ratings of graduates, we wanted to maximize our sample for these analyses, especially given concerns over statistical power. A limitation of this approach is that analyses using other outcomes then are based on a subsample of our 321 PSTs in our main analytic sample. Where we have concerns that using different samples might influence differences we observe across the outcomes we use, we report on results using the constrained sample of PSTs who have data on all outcomes.

⁷ In some cases, PSTs had multiple CTs. Of the 305 PSTs in our analytic sample, 27 PSTs matched to 2 CTs, and the remaining 278 (91%) of PSTs matched to just one CT. For those PSTs with multiple CTs, we kept all CT information available by averaging or summing values. For continuous CT measures, we averaged the values; for example, if a PST was linked to one CT with an observation rating of 4.2 and another with a score of 2.2, then we assigned that PST's CT an average observation rating of 3.2. For dichotomous indicators, such as race and gender, we counted all responses; for example, a PST with a male CT one term and a female CT the next was counted as having a male CT *and* a female CT. We treated CTs' school characteristics in the same ways. We treated individual survey items differently.

student teach in CPS. This is, to some degree, expected. Given our analytic sample is conditional on being employed in CPS, one might expect the district to employ the best prepared and most promising candidates, as difference in GPA may reflect.⁸ Additionally, PSTs in our analytic sample had to be interested and willing to live in the Chicago area, which may partially explain why the analytic sample included a larger proportion of CPS graduates. Among those in our analytic sample, we also compared those PSTs who did (n=225) and did not (n=80) complete surveys (see Appendix Table 3). We found no differences between groups in terms of PST characteristics (gender, race, GPA, likelihood of being a CPS graduate).

Cooperating Teacher Characteristics. CT characteristics are described in Table 1 (bottom). About 76% of PSTs had at least one female CT, while 24% of PSTs had at least one male CT. Nearly two-thirds (66%) of PSTs had at least one White CT, while 23% had at least one Latino CT, 11% had at least one Black CT, and just 3% had at least one Asian CT. 29% of PSTs had at least one CT who graduated from a CPS high school themselves. Almost all PSTs were paired with a CT with tenure (92%), and most worked with a CT who had a post-baccalaureate degree (78%). Fewer PSTs (19%) worked with a CT who had National Board certification. On average, PSTs worked with CTs with about 12 years of experience as CPS teachers and with an overall observation score of 3.5 (out of 4). Appendix Table 2 (middle) compares CTs in our analytic sample to those in our non-analytic sample. CTs in our analytic sample did not significantly differ from the rest of CTs in terms of their background, experience, or professional qualifications. Among CTs in our analytic sample, we compared those who did (n=226) and did not (n=79) complete surveys (see Appendix Table 3). Though both groups were statistically similar in terms of gender, race, likelihood of graduating from CPS, tenure, advanced degree completion, years of experience, and evaluation

⁸ Consistent with this explanation, PSTs who were employed in CPS were rated as significantly better prepared by their CTs than PSTs who were not employed in CPS.

scores (VAM, district observation ratings, PST ratings), CTs who responded to the survey were significantly more likely to be National Board certified.

Field Placement School Characteristics. PSTs' field placement school ("FPS") characteristics are shown in Table 2 (top). About 3 in 10 PSTs completed student teaching in high school (grades 9-12); the rest completed student teaching in mostly K-8 primary settings. While almost half of PSTs (46%) completed student teaching in a FPS without a Black or Latino student majority,⁹ 41% were in a majority Latino school and 14% were in a majority Black school. The average poverty concentration in FPSs was slightly lower than average (-0.01 standard deviation units), and the average school-level prior-year achievement was slightly above average (0.15 standard deviation units). On average, PSTs perceived their working conditions at their FPS positively, with nearly two-thirds of PSTs (63%) rating their FPS working conditions at an average of 3.5 out of 4 or higher.¹⁰ Five Essential¹¹ values (Bryk et al., 2010) for FPSs were in the range of 53 to 67 (out of 100), with Ambitious Instruction being the highest and Effective Leader being the lowest. Appendix Table 2 (bottom) compares average field placement school characteristics for those PSTs who could be linked to district observational data (our analytic sample) to average field placement school characteristics of other PSTs from the same cohort who student taught in CPS but could not be linked to district observation evaluations (mostly individuals who were not employed in CPS, though a small minority were employed in CPS but could not be linked to observational evaluation data). Overall, our analytic sample was more likely to have completed student teaching in schools with higher average poverty concentration.

⁹ We define "majority" Black (or Latino) schools as having at least 70 percent of students who are Black (or Latino).

¹⁰ As described in Methods, our working conditions measure is the mean of four survey items about how strongly PSTs agreed with statements about the working conditions of their field placement schools; these items were on a 4-point scale where 1=strongly disagree, 2=disagree, 3=agree, and 4=strongly agree. A score of 3.5 means that the PSTs' ratings fell, on average, between agree and strongly agree.

¹¹ Through an extensive research program, Tony Bryk and colleagues at the Consortium on Chicago School Research identified the "Five Essentials" which are characteristics of schools that they found to be important for supporting school improvement.

Employed School Characteristics. Table 2 (bottom) summarizes the characteristics of the schools in which PSTs gained employment during school year 2015-16. First, 53 (17.4%) PSTs student-taught in the school in which they became employed. Employed PSTs worked mostly in K-8 primary settings (78%), and 22% worked in high schools. Nearly 4 in 10 (39%) PSTs were employed in schools with a majority Latino student population, and 1 in 4 (26%) were employed in schools with a majority Black student population.

Measures

In this section, we describe the focal outcome measures—PSTs’ and CTs’ perceptions of PSTs’ preparedness, and employed PSTs’ school year 2015-16 observation ratings. We also describe focal Rasch¹² measures used as predictors in various parts of our analyses.

Perceptions of Preparedness. We asked PSTs and CTs a series of similar survey questions about PSTs’ preparedness to take on the responsibilities of teaching in four domains of instruction aligned with CPS’s teacher evaluation system:¹³ planning and preparation, instruction, classroom environment, and professional responsibilities. See Appendix Table 6 for information on preparedness survey items. We submitted these survey items to Rasch analysis to create domain-level measures. To create measures for this study, we first standardized the Rasch domain-level measures for ease of interpretation. We then used the standard errors associated with each domain to create precision-weighted mean measures in each domain.¹⁴ To create an overall measure of preparedness across domains, we divided the sum of the four precision-weighted domain-level measures by the sum of the four weights. In the end, we had one precision-weighted measure for PST self-perceptions of preparedness and another for CTs’ perceptions of their PSTs’ preparedness.

¹² Rasch IRT theory posits that questions of varying degrees of difficulty differentiate people’s placement along a developmental scale: Endorsing more difficult questions means that respondents have higher levels (or more positive beliefs) on the underlying construct (Bond and Fox, 2015). Both item difficulties and respondent abilities are placed on the same scale and expressed in logits. Most measures used in our study met minimum thresholds for reliability (0.8).

¹³ See <http://www.cps.edu/ReachStudents/Pages/AtaGlance.aspx> for more information for CPS “REACH” system.

¹⁴ We did so by multiplying each standardized domain-specific measure by $1/\text{standard error}_{\text{domain}}$, or its weight.

Observation Ratings. REACH Students (Recognizing Educators Advancing Chicago Students) is Chicago Public Schools’ system of educator evaluation and support. A significant component of the REACH evaluation system is its observations of practice scores, which are based on the Danielson-inspired CPS Framework for Teaching. All beginning teachers are rated by their principals or assistant principals a total of 4 times (based on three formal and one informal observation). Evaluators score teachers on 19 different components on a rating from 1-4, and each domain score consists of the ratings on 4-5 components. For the overall observation rating measures in this study, we computed a simple average of the four domains for each employed PST and for each CT;¹⁵ we refer to these measures as “observational ratings” as shorthand.

Predictor Rasch Measures. We used five PST Rasch measures and three CT Rasch measures as predictors. Appendix Table 7 shows detailed information about the items and Rasch reliabilities. PST Rasch predictors included measures of PST-perceived: (i) CT teaching effectiveness (drawing on questions about the four instructional domains plus teaching in an urban context); (ii) field instructor helpfulness; (iii) instructional domain-specific conversations; (iv) mentoring relationship and feedback received; and (v) assistance with the job search. For CTs, we used Rasch measures of CT-perceived (i) instructional domain-specific mentoring provided; (ii) frequency of feedback given; and (iii) assistance with the job search.

Analytic Method

Research Question 1: To estimate employed PSTs’ observational ratings in school year 2015-16 as a function of PST and CT perceptions of PST preparedness, we used 2-level hierarchical linear models with PSTs at level 1 and employment school at level 2. The general form of the model is summarized in Equation 1:

¹⁵ Seven PSTs were missing observational ratings on the planning and preparation and professional responsibilities domains, so we imputed mean values in these rare cases.

$$RATING_{ij} = \gamma_{00} + \gamma_{10}Prep_{ij} + \gamma ES_{0j} + \gamma Z_{ij} + u_{0j} + r_{ij} \text{ (Equation 1)}$$

where the school year 2015-16 observational rating of PST i in school j is a function of an intercept (γ_{00}), focal predictor $Prep_{ij}$ (PSTs' and CTs' perceptions of PST preparedness), ES_j , a vector of employed school characteristics (school level, poverty concentration, racial composition, school-level prior achievement, and whether the PST student-taught there), a school random effect u_{0j} , and a PST-level residual, r_{ij} . We enter focal predictors ($Prep_{ij}$) independently in separate regression models. In a second model specification, we also control for Z_{ij} , a vector of characteristics of the PST (race/ethnicity and gender) and FPS (school level, school-level prior achievement¹⁶).

Research Questions 2-3: To estimate PSTs' observational ratings in school year 2015-16 as a function of features of preparation, we used 2-level hierarchical linear models with PSTs at level 1 and employment school at level 2. The general form of the model is summarized in Equation 2:

$$RATING_{ij} = \gamma_{00} + \gamma_{10}Feat_{ij} + \gamma ES_{0j} + \gamma Z_{ij} + u_{0j} + r_{ij} \text{ (Equation 2)}$$

where the school year 2015-16 observational rating of PST i in school j is a function of an intercept (γ_{00}), focal predictor $Feat_{ij}$ (features of preparation), ES_{0j} , a vector of employed school characteristics (described above), a school random effect u_{0j} , and a PST-level residual, r_{ij} . We enter focal predictors ($Feat_{ij}$) independently in separate regression models, as well as in groups.¹⁷ When entered as groups, we also control for Z_{ij} , a vector of characteristics of the PST and FPS (described above). To estimate PSTs' and CTs' perceptions of PST preparedness as a function of features of preparation, we use Equation 2 but substitute perceptions of preparedness for observational ratings, drop ES from the

¹⁶ We retained this final set of covariates after testing other potential covariates (e.g. GPA, FPS race composition) that either reduced the sample, did not contribute to model fit, and/or were conceptually redundant of the covariates we retained.

¹⁷ We include models with predictors entered independently because, in some cases, estimates changed substantially when included with conceptually related predictors. For example, CT qualifications such as observational ratings, years of service, and tenure status are related to one another; we are interested in the effects of CTs' observational ratings both with and without adjusting for years of experience and tenure status.

model, and, since all data collection for these analyses occurred prior to completing initial preparation, nest PSTs in teacher education programs instead of employment schools.

RESULTS

In this section, we investigate the relationships between PSTs' and CTs' perceptions of PST preparedness and PSTs' effectiveness as first-year teachers. We then consider whether PSTs who feel better prepared or who are rated as better prepared by CTs at the end of student teaching also receive better observational ratings during their first year of teaching. Next, we describe which features of teacher preparation—CT qualifications, CT coaching, field placement schools, and other aspects—predicted PSTs' feelings of preparedness, their CTs' ratings of their preparedness, and their observational ratings based on district evaluations during their first year of teaching. We then examine whether certain features of preparation predict better instructional readiness regardless of how it is measured or certain features of preparation predict some but not other measures of instructional readiness.

Are preservice teachers who, at the end of preparation, (a) report feeling better prepared or (b) are rated as being more prepared by their cooperating teachers more instructionally effective based upon district observational evaluations during their first year of teaching?

Correlation and regression analyses both indicated that PSTs' self-perceptions of preparedness at the end of student teaching were unrelated to PSTs' first-year observation evaluations. However, CTs' evaluations of their PSTs' preparedness positively and significantly predicted PSTs' first-year observational evaluations. In other words, CTs seemed able to discriminate which PSTs would perform well in their first-year observational ratings whereas PSTs self-evaluations did not predict their future evaluations. We elaborate below.

We began by considering simple correlations between our three measures: PSTs' own perceptions of preparedness, CTs' perceptions of PSTs' preparedness, and observational ratings of PSTs by district evaluators during their first year of teaching. As summarized in Table 4, PSTs' perceptions of their own preparedness were weakly correlated with CTs' perceptions of their PSTs' preparedness (0.06) and with their first-year observational ratings (0.03). However, CTs' perceptions of PSTs' preparedness were more strongly correlated (0.24) with PSTs' first-year observational ratings.

Table 5 summarizes results from regression models estimating PSTs' first-year observational ratings as a function of PSTs' self-perceptions of preparedness and CTs' perceptions of PST preparedness at the end of student teaching. Results suggest that PSTs were significantly more effective as first-year teachers when their CTs rated them as more prepared, both overall and in individual instructional domains. Results were similar when we included PST and field placement school characteristics as covariates (Model B). When PSTs themselves felt better prepared, they were generally no more or less effective as first-year teachers; this was true across model specifications.

Having sought evidence for convergent validity among the three measures for PSTs' instructional readiness, in the next section we investigated which features of preparation predicted these three measures.

What features of preparation predict (a) how prepared PSTs feel at the end of preparation, (b) how prepared their cooperating teachers believe PSTs are at the end of preparation, and (c) how effective they are according to district evaluators during their first year of teaching?

In this section, we examined whether features of PSTs' preparation predicted their feelings of preparedness, their CTs' ratings of their preparedness, and their first-year teaching effectiveness as measured by district observational ratings. The features of preparation on which we focused

included CTs' professional qualifications, PSTs' and CTs' perceptions of CTs' coaching, student teaching field placement school characteristics, and other aspects of preparation.

CT Qualifications. Table 6 summarizes results from models estimating our three outcomes as a function of CTs' professional qualifications. We begin by summarizing results with PST observational ratings as the outcome (far right). A consistent finding across model specifications was that PSTs who learned to teach with CTs who had stronger observational ratings themselves had significantly stronger observational ratings during their first year. Every additional point (on a scale of 1-4) in CTs' overall observational ratings was associated with a 0.16 point gain for PSTs' overall observational ratings. Additionally, these differences appear to be meaningful. Based on a recent study in Chicago (Jiang & Sporte, 2016), the average difference ($B=0.17$) between a first-year teacher and a teacher with between two and five years of experience¹⁸ is roughly equivalent to the average difference we observe ($B=0.16$) between a newly hired teacher whose CT had received a district rating of 3.0 and a newly hired teacher whose CT had received a district rating of 4.0. CTs' value-added scores in reading and math also trended positive but were not statistically significant.

Other aspects of CTs' professional qualifications were either unrelated to PSTs' first-year effectiveness or predicted worse performance. Having a CT with tenure or more years of teaching experience appeared unrelated to PSTs' first-year teaching effectiveness. Perhaps counterintuitively, having a CT with National Board certification was associated with lower first-year observational ratings for PSTs. Similarly, when PSTs rated their CTs as more effective teachers, their own first-year performance was significantly lower.

Turning from first-year observational ratings as an outcome, we next considered PSTs' and CTs' perceptions of how prepared PSTs were at the end of student teaching (Table 6, left and

¹⁸ See Jiang and Sporte (2016) Table 3, page 18. These findings are also consistent with Ronfeldt and Campbell (2016) who find the average difference between first and second year teachers in Tennessee, using a different observational rubric, is 33 percent of a standard deviation unit. Since the standard deviation of observation ratings of PSTs in our sample is 0.40, our observed difference of 0.16 corresponds with 40 percent of a standard deviation unit.

middle). PSTs felt somewhat better prepared when their CTs were more experienced as CPS teachers, though results were not statistically significant. Additionally, when PSTs rated their CTs as more instructionally effective in observation-aligned domains, they also felt significantly better prepared; on the other hand, though they felt better prepared, these PSTs actually received worse first-year observational ratings (see above).

CT Coaching. Table 7 summarizes results from models investigating various CT coaching measures as predictors, including PSTs' perceptions of CTs' coaching (top), and CTs' perceptions of their own coaching, as well as their coaching experience and training (bottom). In terms of PSTs' first-year effectiveness as measured by observational ratings on the district rubric (Table 7, right), only CTs' domain-specific support (e.g. mentoring in instruction, environment, planning) was a positive predictor, suggesting that PSTs were more effective in their first year when their CTs reported providing stronger support in specific domains of instruction, including those assessed on the district rubric. Otherwise, perceptions of CTs' coaching and their training and experience as mentors were mostly unrelated to PSTs' first-year effectiveness. One possible exception is that PSTs received lower observational ratings when they reported that their CTs had provided better job support; however, estimates were significant only after adjusting for mentoring in other areas.

Regarding CTs' perceptions of PSTs' preparedness (middle columns), most coaching measures, whether from the perspective of PSTs or CTs, were positively and significantly predictive in at least one model specification. In other words, CTs rated their PSTs as better prepared when either they or their PSTs reported stronger coaching. By contrast, only PSTs' perceptions of coaching predicted their own levels of self-perceived preparedness; CTs' perceptions of coaching

were unrelated to PSTs' self-perceptions of preparedness. Lastly, having training or more experience in mentoring was unrelated to both PSTs' or CTs' perceptions of preparedness.¹⁹

Field Placement School Characteristics. Next, we analyzed how field placement school characteristics explained PSTs' first-year observational ratings and perceived preparedness; Table 8 summarizes these results. First, we found that PSTs were significantly more effective as first-year teachers when they were hired into a school in which they student-taught. Being a first-year teacher in a field placement school was associated with a 0.15 point increase in first-year observational ratings. Though prior research has indicated field placement school-level achievement and working conditions positively predicts graduates' VAM, we found PSTs i) received lower first-year observational ratings when they learned to teach in field placements with better average achievement and (ii) received no higher or lower first-year observational ratings based on their perceptions of the FPS having better working conditions. Field placement school level, socioeconomic status, and racial composition were not significantly associated with first-year teaching effectiveness.

The majority of field placement school characteristics we tested was unrelated to PSTs' and CTs' perceptions of PST preparedness. One notable exception was PSTs' perceptions of working conditions. Although perceiving better working conditions was unassociated with first-year observational ratings, PSTs who thought their field placement schools were positive places to work felt better prepared. Additionally, being in a field placement school with better average achievement was associated with significantly lower perceptions of PST preparedness by CTs, though not for PSTs themselves.

Other Features of Preparation. Table 9 summarizes results from models examining the associations between other features of preparation and the three focal outcomes. In terms of

¹⁹ In separate analyses not shown, we also explored the question of whether instructionally stronger CTs, as measured by their observation ratings, make better coaches, as measured by PSTs' and CTs' perceptions. We found that when CTs had higher observation ratings, PSTs perceived coaching significantly more favorably; conversely, CTs with higher observation ratings did not perceive their own coaching more or less positively.

predictors of first-year observational ratings, only being an elementary or all-subject student teacher was associated with higher first-year observational ratings across model specifications. This finding is consistent with prior research that shows elementary teachers are generally rated more favorably than secondary teachers. In one model specification, PSTs also received higher first-year observational ratings when their programs were primarily responsible for selecting their field placement schools. Finally, PSTs who reported taking more courses prior to student teaching had slightly lower first-year observational ratings.

On the whole, CTs did not feel PSTs were better or worse prepared as a function of these other features of preparation (Table 9, middle). However, several features of preparation significantly explained PSTs' own feelings of preparedness. Taking more courses prior to student teaching, spending more hours in the placement, feeling that student teaching was instructive, primarily being a lead teacher during student teaching, and finding the field instructor to be helpful were all associated with PSTs feeling better prepared.

Do the same features of preparation that predict one outcome also predict the others?

By investigating features of preparation that positively predicted outcomes for PSTs, this study aimed to identify promising features of preparation. In the ideal case, the most promising features would not only make PSTs feel better prepared but would also predict better CT ratings and first-year observational evaluations. However, we found no features of preparation to positively predict all three outcomes used in this study.

Only one predictor positively predicted PSTs' first-year observational ratings and either the PST or CT feelings of preparedness outcome measures. Namely, when CTs felt they provided stronger mentoring in specific instructional domains (including those evaluated on the district

rubric) they rated their PSTs as being better instructionally prepared and their PSTs received stronger first-year evaluations.²⁰

Only two predictors were positively related to both of our survey-based feelings of preparedness outcome measures (PST and CT). PSTs who reported better mentoring relationships and better job support from their CTs felt better prepared and were evaluated as better prepared by their CTs.²¹

Feeling more prepared, but being less effective. A few features of preparation were associated with PSTs feeling significantly better prepared after student teaching but receiving significantly lower observational evaluations as first-year teachers. This was the case for PSTs' perceptions of their CTs' teaching effectiveness. PSTs felt better prepared when they thought their CTs were more effective teachers, but they turned out to be less effective themselves the following year. The same was true for the number of courses taken by PSTs prior to student teaching. Taking more courses prior to student teaching predicted PSTs feeling more prepared by the end of student teaching but was associated with significantly lower observational ratings as a first-year teacher.

Seeming more prepared, but being less effective. In one case, a feature was associated with higher CT ratings of preparedness but lower first-year teaching effectiveness. When CTs reported providing more job search assistance, they thought their PSTs were better prepared, but their PSTs had lower first-year observational ratings.

DISCUSSION

²⁰ We also find some evidence that CT observational ratings are positively related to both PST first-year evaluation scores and PST self-perceived preparedness, though the latter is non-significant in models using our analytic sample. In alternative specifications including all PSTs employed in CPS (whether or not they have first-year observational ratings) we find the relationship to be statistically significant; moreover, in prior work with the full population of survey respondents, we also find this relationship to be significant in some models (Matsko et al., conditionally accepted).

²¹ Completing more hours of student teaching positively predicted PSTs' self-perceived preparedness. This predictor also trended positive on CTs' perceptions of PST preparedness, though was not statistically significant. However, in alternative specifications that included all PSTs employed in CPS (whether or not they had first-year evaluation scores) and in prior work (Matsko et al., conditionally accepted), results were positive and significant at the $p < 0.05$ level.

This study makes a number of contributions to the existing literature. It provides the first analysis of which we are aware, of whether PSTs who report feeling better prepared are actually more effective teachers in their first year. Specifically, it finds that PSTs' self-perceptions of preparedness are unrelated to first-year observational ratings. Though teacher education programs often use PSTs' perceptions of preparedness to assess program quality and teacher education scholars use these same measures to identify promising features of preparation, our results suggest programs and scholars should not use PSTs self-perceptions of preparedness at the time of program completion as a signal for instructional effectiveness in the first year. On the other hand, we find that CTs' perceptions of their PSTs' preparedness do have predictive validity.

This study also contributes to the literature aiming to identify promising features of preparation by examining which ones predict better PST outcomes. Importantly, this is the first study of which we are aware to link features of preparation to graduates' first-year instructional effectiveness as measured by observational evaluations. In doing so, it identifies a number of features positively associated with this outcome: (i) CTs' instructional effectiveness as measured by observational ratings on the district rubric, (ii) CTs' self-perceptions of mentoring they provided in specific instructional areas (including those evaluated on the district rubric), (iii) gaining employment in the same school in which PSTs completed their student teaching experience, and (iv) completing an elementary, self-contained student teaching placement.

It is important to bear in mind that results of this study are correlational in nature and we cannot conclude that the relationships we observe are causal in nature. A particular threat to casual interpretations is that there are many forms of PST selection likely at work, including selection into certain kinds of field placement and current schools and to certain kinds of CTs. For example, it might be that PSTs who are more instructionally promising prior to student teaching are more likely to seek out or get assigned to more instructionally effective CTs. Thus, the positive relationship

between CT instructional effectiveness and PST instructional effectiveness could be due to this initial selection rather than a causal story.

In some cases, findings from this study based on first-year observational ratings were somewhat inconsistent with prior research using graduates' VAMs as outcomes. While Ronfeldt (2012) and Ronfeldt (2015) found that graduates had better VAM when they learned to teach in field placement schools with better working conditions and a history of prior school-level achievement gains, we found field placement school working conditions as measured by the Five Essentials to be unrelated to graduates' first-year observational ratings. Moreover, field placement school-level achievement was negatively related to first-year observational ratings, while PSTs' ratings of their field placement working conditions were unrelated. It is possible that these mixed results are due to the fact that these studies focused on different measures of instructional quality – VAM scores versus observational evaluations. Another possibility is that these features of preparedness function differently in different labor markets. More research is needed to interrogate these mixed findings.

This study also extends prior research that uses PSTs' self-perceived preparedness in order to identify promising features of preparation. Consistent with prior work, it finds that PSTs felt better prepared when they rated their CTs as being more instructionally effective and as providing stronger coaching (Matsko et al., conditionally accepted). Also consistent with prior work, PSTs felt better prepared when they reported more student teaching, more coursework prior to student teaching, better field instructor support, better student teaching experiences generally, and better field placement working conditions specifically (Ronfeldt & Reiningger, 2012; Ronfeldt, Reiningger, & Kwok, 2013; Ronfeldt, Schwartz, & Jacob, 2014;). None of these features of preparation, though, predicted stronger first-year instructional effectiveness as measured by observational ratings. To the degree that program leaders and policymakers intend to use outcomes-based research to identify promising levers for program improvement, an implication is that they may first determine which

outcome they want to impact. Rather than there being certain features that positively predicted all outcomes, different features predicted different outcomes.

Student teachers' perceptions of preparedness measures remain the most commonly used indicators of program and graduate quality. There is some evidence to support their use, as PSTs' feelings of preparedness are related to teacher retention (Ronfeldt et al., 2014) and teachers' self-efficacy (Darling Hammond et al., 2002). However, our findings suggest that the use of this common quality indicator might need to be re-considered. For example, teacher educators and scholars who tend to rely on measures such as PST perceptions of preparedness to assess graduate readiness to teach may gain more insights from the perceptions of the CTs who spend the most time with preservice candidates during student teaching. Teacher educators interested in improving program design and content may want to consider relationships between features of preparation and graduates' first-year observational ratings as well.

REFERENCES

- Bastian, K. C., Patterson, K. M., & Pan, Y. (2017). Evaluating Teacher Preparation Programs With Teacher Evaluation Ratings: Implications for Program Accountability and Improvement. *Journal of Teacher Education*, 0022487117718182.
- Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, 31(4), 416-440.
doi:10.3102/016237370935312
- Boyd, H., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2012). Recruiting effective math teachers: Evidence from New York City. *American Educational Research Journal*, 49(6), 1008-1047.
- Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). An Evaluation of Teachers Trained through Different Routes to Certification. Final Report. NCEE 2009-4043. *National Center for Education Evaluation and Regional Assistance*.
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach?. *Journal of teacher education*, 53(4), 286-302.
- Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, 13(42).
- Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes. *Journal of Policy Analysis and Management*, 25(1), 75-96.
- Goldhaber, D., Liddle, S., & Theobald, R. (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review*, 34, 29-44.

- Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, 95(7), 798-812.
- Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2014). The effects of teacher entry portals on student achievement. *Journal of Teacher Education*, 65(1), 7-23.
- Jiang, J. Y., & Spörte, S. (2016). Teacher evaluation in Chicago: Differences in observation and value-added scores by teacher, student, and school characteristics. Retrieved from: <https://consortium.uchicago.edu/sites/default/files/publications/Teacher%20Evaluation%20in%20Chicago-Jan2016-Consortium.pdf>
- Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2015). Teacher preparation programs and teacher quality: Are there real differences across programs?. *Education Finance and Policy*, 10(4), 508-534.
- Matsko, K., Ronfeldt, M., Greene, H., Reiningger, M., & Brockman, S., (conditionally accepted). The role of cooperating teachers in preparing pre-service teachers: A district-wide portrait. *Journal of Teacher Education*.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of education review*, 13(2), 125-145.
- Ronfeldt, M. (2012). Where should student teachers learn to teach?: Effects of field placement school characteristics on teacher retention and effectiveness. *Educational Evaluation and Policy Analysis*, 34(1), 3-26.
- Ronfeldt, M. (2015). Field placement schools and instructional effectiveness. *Journal of Teacher Education*, 66(4), 304-320.
- Ronfeldt, M. & Campbell, S. (2016). Evaluating teacher preparation using graduates' observational ratings. *Educational Evaluation and Policy Analysis*, 38(4), 603-625.

- Ronfeldt, M. & Reininger, M. (2012). More or better student teaching? *Teaching and Teacher Education*, 28(8), 1091-1106.
- Ronfeldt, M., Reininger, M., & Kwok, A. (2013). Recruitment or preparation? Investigating the effects of teacher characteristics and student teaching. *Journal of Teacher Education*, 64(4), 319-337.
- Ronfeldt, M., Schwartz, N., & Jacob, B. (2014). Does pre-service preparation matter? Examining an old question in new ways. *Teachers College Record*, 116(10), 1-46.
- von Hippel, Paul T., Osborne, C., Lincove, J.A., Mills, N., & Bellows, L. (2014) Fooled by Randomness? The challenges of seeking exceptional teacher preparation programs among many noisy estimates. Retrieved from <http://ssrn.com/abstract=2506935>
- von Hippel, Paul T. and Bellows, Laura, How Much Does Teacher Quality Vary Across Teacher Preparation Programs? Reanalyzing Estimates from 6 States (June 20, 2017). Available at SSRN: <https://ssrn.com/abstract=2990498>.

TABLES

Table 1: Pre-service and cooperating teacher characteristics

	Percent/Mean (sd)	N
Pre-Service Teachers		
Female	72.0	304
Male	28.0	304
White	57.4	305
Latino	22.0	305
Black	10.2	305
Asian	5.6	305
Other	5.9	305
Graduated from CPS	28.5	305
Undergraduate GPA (100-pt scale)	91.6 (7.5)	305
Had prior teaching experience (e.g. substitute)	30.5	220
Age during student-teaching	26.2 (5.7)	218
Parent during student-teaching	11.9	219
Cooperating Teachers		
Female	75.9%	282
Male	24.1%	282
White	66.0%	282
Latino	23.4%	282
Black	11.0%	282
Asian	3.2%	282
Other	5.3%	282
Graduated from CPS	28.6%	220
Advanced degree (e.g. M.A.)	78.0%	282
Tenure	92.1%	280
National Board certification	18.6%	280
Years of CPS service	12.1 (6.5)	278
Average Observational Rating (min=1, max=4)	3.5 (0.3)	279
Reading VAM	0.2 (0.7)	86
Math VAM	0.1 (0.9)	68
PST Perception of CT Teaching Effectiveness (Standardized Rasch)	-0.02 (0.96)	220

Note: PST gender, race, CPS graduate status, and GPA information came from student teaching registration data and reflect a maximum possible sample of 305 PSTs in our analytic sample. Having prior teaching experience, age, and parent status came from survey items and reflect a maximum of 225 PSTs in our analytic sample who have the survey-based post preparedness outcome. CT information came mostly from CPS personnel data, except for CPS graduate status which came from CT surveys and PST perception of CT teaching effectiveness which came from PST survey measures and, thus, have smaller samples.

Table 2: School characteristics

	Percent/Mean (sd)	N
Field Placement School (2014-15)		
High school (9-12)	33.4%	299
Majority Latino students	40.7%	295
Majority Black students	13.6%	295
Concentration of poverty (scon)	-0.01 (0.52)	295
School-level prior achievement	0.15 (0.57)	292
PST-Perceived working conditions (3.5 or above, out of 4)	62.6%	222
5E: Ambitious Instruction	67.0 (13.7)	285
5E: Involved Families	65.9 (16.2)	282
5E: Collaborative Teachers	61.4 (14.6)	282

5E: Supportive Environment	55.1 (15.9)	285
5E: Effective Leader	53.5 (14.5)	282
School of Employment (2015-16)		
PST student-taught in same school	17.4%	305
High school (9-12)	22.4%	304
Majority Latino students	39.8%	304
Majority Black students	25.7%	304
Concentration of poverty (scon)	0.20 (0.54)	304
School-level prior achievement	-0.01 (0.49)	301

Note: School-level prior achievement is measured in standard deviation units and is based on prior-year NWEA reading scores of current students (standardized within grade within year). A difference of 0.5 sd units reflects approximately the difference between a school with average prior achievement and a school with top-third (or bottom-third) prior achievement. We define “majority” Black (or Latino) schools as having at least 70 percent of students who are Black (or Latino).

Table 3: Characteristics of student teaching and perceptions of mentoring

	Percent/Mean (sd)	N
<i>Student Teaching</i>		
Student-taught all subjects (i.e. elementary)	32.4%	222
Number of courses taken prior to student teaching	5.4 (2.7)	252
Number of total hours spent student teaching	214.2 (69.9)	222
PST strongly agrees s/he learned a lot from student teaching	74.1%	220
Teacher Ed Program (TEP) chose the placement school	56.3%	222
Primarily lead teacher during student teaching	43.7%	222
PST-Perceived field instructor support [^]	-0.00 (1.2)	217
<i>PST Perceptions of CT Coaching</i>		
Domain-specific conversations [^]	-0.06 (1.03)	219
Mentoring relationship and feedback [^]	0.02 (0.98)	220
Assistance with job search [^]	0.06 (0.97)	220
<i>CT Perceptions of Own Coaching</i>		
Domain-specific mentoring [^]	-0.01 (0.99)	227
Frequency of feedback [^]	-0.08 (0.93)	227
Assistance with job search [^]	0.07 (0.97)	226
CT’s teaching strongly aligns with TEP	36.7%	205
CT has had training in mentoring	29.0%	231
CT has had prior mentees	74.7%	233

Note: For more information about the coaching (Rasch) measures see Appendix Table 7. Unlike the other items in this table which are from the post survey only, we asked about number of courses taken prior to student teaching on both the pre and the post surveys which is why the sample is greater for this item. [^]Indicates standardized Rasch measures; since these were standardized based on Rasch measures in the full sample, the means and standard deviations differ slightly from 0 and 1 for the analytic sample, shown here.

Table 4: Correlation Matrix of PSTs’ Overall Observational Rating & Perceptions of Overall Preparedness

	1	2	3
1. Overall Self-Perceived PST Preparedness	1.00		
2. Overall CT-Perceived PST Preparedness	0.06	1.00	
3. Overall PST Observational Rating	0.03	0.24	1.00

Note: We also examined disattenuated correlations and results were very similar.

Table 5: Employed PSTs’ First-Year Observational Rating as a Function of Perceived Preparedness after Student Teaching

	Overall Observational Rating: Model A	Overall Observational Rating: Model B
Overall Self-Perceived PST Preparedness	-0.02	-0.02

	(0.03)	(0.03) N=216
Overall CT-Perceived PST Preparedness	0.07** (0.02)	0.06* (0.02) N=220
Domain-Specific Self-Perceived PST Preparedness:		
Planning & Preparation	-0.00 (0.03)	
Instruction	-0.02 (0.03)	
Environment	0.00 (0.03)	
Prof. Responsibilities	-0.05* (0.02)	
Domain-Specific CT-Perceived Preparedness of PSTs':		
Planning & Preparation	0.05* (0.02)	
Instruction	0.07** (0.02)	
Environment	0.06* (0.02)	
Prof. Responsibilities	0.05* (0.02)	

Notes: In Models A and B, the focal predictors are included independent of one another (i.e., PST measures of preparedness is not included in same models as CT measures of preparedness). Models A and B both include current 2015-16 school characteristics (school grade level, school area poverty concentration, student race composition, school prior achievement, and whether the PST student-taught in the current school) as covariates. Model B *also* includes PST and field placement school covariates (PST race, PST gender, placement school grade level, and placement school prior achievement). Results were very similar for domain-specific observational ratings, so we only report average observational ratings. (+p<0.10; *p<0.05; **p<0.01; ***p<0.001)

Table 6: Perceived Preparedness and First-Year Effectiveness as a Function of CT qualifications

CT Instructional Quality and Qualifications	Self-Perceived PST Preparedness at end of Student Teaching		CT-Perceived PST Preparedness at end of Student Teaching		First-Year PST Observation Rating	
	A	B	A	B	A	B
Tenure	0.06 (0.21)	0.10 (0.24)	-0.02 (0.23)	-0.17 (0.26)	0.01 (0.07)	-0.03 (0.09)
National Board	-0.09 (0.16)	-0.15 (0.16)	0.03 (0.15)	0.05 (0.16)	-0.09+ (0.05)	-0.11* (0.05)
Years of CPS service	0.02* (0.01)	0.02+ (0.01)	0.01 (0.01)	0.01 (0.01)	-0.00 (0.00)	-0.00 (0.00)
CT Observational Rating: Overall	0.20 (0.19)	0.17 (0.21)	-0.07 (0.20)	-0.09 (0.22)	0.13* (0.07)	0.16* (0.07)
CT Observational Rating: Planning	-0.01 (0.16)		-0.13 (0.16)		0.10* (0.05)	
CT Observational Rating: Instruction	0.22 (0.19)		0.08 (0.20)		0.12* (0.06)	
CT Observational Rating: Environment	0.23 (0.18)		-0.10 (0.19)		0.12* (0.06)	
CT Observational Rating: Prof. Responsibilities	0.20 (0.15)		-0.02 (0.15)		0.05 (0.05)	
PST-Perceived CT Teaching Effectiveness (Standardized Rasch)	0.28*** (0.06)		0.10 (0.08)		-0.06* (0.03)	
CT VAM: Reading	0.09 (0.19)		-0.04 (0.18)		0.02 (0.06)	
CT VAM: Math	0.19 (0.13)		0.05 (0.15)		0.05+ (0.03)	
N		203		215		268

Notes: For the outcomes of PST and CT perceived preparedness, Model A includes each focal predictor (row) independently. In the case of models with first-year observation ratings as outcomes, Model A also includes covariates for current 2015-16 school characteristics (school grade level, school area poverty concentration, student race composition, school prior achievement, and whether the PST student-taught in the current school). For all outcomes, Model B includes focal predictors together in the same model and *also* includes PST and field placement school covariates (PST race, PST gender, placement school grade level, and placement school prior achievement). (+p<0.10; *p<0.05; **p<0.01; ***p<0.001)

Table 7: Perceived Preparedness and First-Year Effectiveness as a Function of CT Coaching

	Self-Perceived PST Preparedness at end of Student Teaching		CT-Perceived PST Preparedness at end of Student Teaching		First-Year Observational Rating	
	A	B	A	B	A	B
<i>PST Perceptions of CT Coaching</i>						
Domain-specific conversations (Standardized Rasch)	0.22*** (0.06)	0.04 (0.08)	0.08 (0.07)	-0.12 (0.09)	-0.04+ (0.02)	-0.05 (0.03)
Mentoring relationship and feedback (Standardized Rasch)	0.33*** (0.06)	0.28*** (0.08)	0.21** (0.07)	0.13 (0.10)	-0.02 (0.02)	0.01 (0.03)
Assistance with job search (Standardized Rasch)	0.24*** (0.06)	0.11 (0.07)	0.33*** (0.07)	0.36*** (0.09)	-0.01 (0.03)	0.00 (0.03)
N		211		163		210
<i>CT Perceptions of Own Coaching</i>						
Domain-specific mentoring (Standardized Rasch)	-0.02 (0.07)	-0.14 (0.09)	0.45*** (0.06)	0.34*** (0.07)	0.07** (0.02)	0.11*** (0.03)
Frequency of feedback (Standardized Rasch)	-0.04 (0.07)	-0.01 (0.09)	0.14* (0.07)	0.01 (0.07)	0.01 (0.02)	0.01 (0.03)
Assistance with job search (Standardized Rasch)	-0.00 (0.07)	0.08 (0.09)	0.32*** (0.06)	0.14* (0.07)	-0.02 (0.02)	-0.08** (0.03)
CT's teaching strongly aligns with TEP	0.05 (0.15)	0.12 (0.17)	0.61*** (0.13)	0.31* (0.13)	0.06 (0.05)	-0.03 (0.05)
<i>CT Mentoring Experience</i>						
CT has had training in mentoring	-0.02 (0.15)	-0.08 (0.17)	-0.06 (0.14)	-0.05 (0.14)	0.01 (0.05)	-0.00 (0.05)
CT has had prior mentees	0.08 (0.16)	0.07 (0.18)	0.19 (0.14)	0.09 (0.14)	0.05 (0.05)	0.03 (0.05)
N		149		197		198

Notes: For the outcomes of PST and CT perceived preparedness, Model A includes each focal predictor (row) independently. In the case of models with first-year observation ratings as outcomes, Model A also includes covariates for current 2015-16 school characteristics (school grade level, school area poverty concentration, student race composition, school prior achievement, and whether the PST student-taught in the current school). For all outcomes, Model B includes focal predictors together in the same model and *also* includes PST and field placement school covariates (PST race, PST gender, placement school grade level, and placement school prior achievement). (+p<0.10; *p<0.05; **p<0.01; ***p<0.001)

Table 8: Perceived Preparedness and First-Year Effectiveness as a Function of Field Placement School Characteristics

	Self-Perceived PST Preparedness after Student Teaching		CT-Perceived PST Preparedness after Student Teaching		First-Year Observational Rating	
	A	B	A	B	A	B
Employed in one of their field placement schools					0.15** (0.05)	0.15** (0.05)
Concentration of poverty	-0.13 (0.13)	-0.19 (0.19)	-0.19 (0.12)	-0.35+ (0.18)	-0.06 (0.04)	-0.10+ (0.06)
High school level (9-12)	-0.14 (0.13)	0.04 (0.15)	-0.03 (0.13)	0.06 (0.15)	-0.01 (0.05)	0.02 (0.05)
Majority Black students (vs. mixed)	0.11 (0.20)	0.36 (0.26)	-0.21 (0.19)	-0.10 (0.26)	-0.12+ (0.07)	-0.09 (0.08)
Majority Latino students (vs. mixed)	-0.06 (0.14)	0.08 (0.18)	0.02 (0.14)	-0.00 (0.17)	-0.03 (0.05)	-0.05 (0.05)
School prior achievement	0.13 (0.12)	0.14 (0.17)	-0.11 (0.11)	-0.28* (0.14)	-0.04 (0.04)	-0.12** (0.04)
PST-Perceived working conditions (3.5 or above, out of 4)	0.51*** (0.12)		0.28+ (0.15)		-0.06 (0.05)	
5E: Involved Families	0.01 (0.00)		-0.00 (0.00)		0.00 (0.00)	
5E: Collaborative Teachers	-0.00 (0.00)		-0.01 (0.00)		0.00 (0.00)	

5E: Ambitious Instruction	0.00 (0.00)		-0.01+ (0.00)		0.00 (0.00)	
5E: Supportive Environment	-0.00 (0.00)		-0.01 (0.00)		-0.00 (0.00)	
5E: Effective Leader	0.00 (0.00)		-0.01 (0.00)		0.00 (0.00)	
N		217		220		289

Notes: For the outcomes of PST and CT perceived preparedness, Model A includes each focal predictor (row) independently. In the case of models with first-year observation ratings as outcomes, Model A also includes covariates for current 2015-16 school characteristics (school grade level, school area poverty concentration, student race composition, school prior achievement, and whether the PST student-taught in the current school). For all outcomes, Model B includes focal predictors together in the same model and *also* includes PST and field placement school covariates (PST race, PST gender, placement school grade level, and placement school prior achievement). (+p<0.10; *p<0.05; **p<0.01; ***p<0.001)

Table 9: Perceived Preparedness and First-Year Effectiveness as a Function of Other Features of Preparation

	Self-Perceived PST Preparedness after Student Teaching		CT-Perceived PST Preparedness after Student Teaching		First-Year Observational Rating	
	A	B	A	B	A	B
Student-taught all subjects (i.e. elementary)	0.19 (0.13)	0.01 (0.15)	0.18 (0.16)	0.20 (0.19)	0.12* (0.06)	0.17** (0.06)
Number of courses taken prior to student teaching	0.09*** (0.02)	0.07** (0.02)	0.04 (0.03)	0.02 (0.03)	-0.01 (0.01)	-0.03** (0.01)
Number of total hours spent student teaching	0.003*** (0.001)	0.003** (0.001)	0.002+ (0.001)	0.002 (0.001)	-0.00 (0.00)	0.00 (0.00)
PST strongly agrees s/he learned a lot from student teaching	0.56*** (0.14)	0.57*** (0.14)	0.26 (0.17)	0.26 (0.18)	0.03 (0.05)	0.04 (0.06)
TEP chose the placement school	-0.09 (0.13)	-0.15 (0.12)	0.18 (0.15)	0.07 (0.17)	0.10* (0.05)	0.05 (0.05)
Primarily lead teacher during student teaching	0.25* (0.13)	0.36** (0.13)	-0.01 (0.15)	0.05 (0.18)	-0.05 (0.05)	-0.07 (0.06)
PST-Perceived field instructor support (Standardized Rasch)	0.28*** (0.05)	0.19*** (0.05)	0.02 (0.06)	0.01 (0.07)	-0.01 (0.02)	0.01 (0.02)
N		180		141		179

Notes: For the outcomes of PST and CT perceived preparedness, Model A includes each focal predictor (row) independently. In the case of models with first-year observation ratings as outcomes, Model A also includes covariates for current 2015-16 school characteristics (school grade level, school area poverty concentration, student race composition, school prior achievement, and whether the PST student-taught in the current school). For all outcomes, Model B includes focal predictors together in the same model and *also* includes PST and field placement school covariates (PST race, PST gender, placement school grade level, and placement school prior achievement). (+p<0.10; *p<0.05; **p<0.01; ***p<0.001)

APPENDICES

Appendix Table 1: Summary of Survey Counts and Response Rates

Survey Administration	Timeline	Response Rate
PST Pre-Student Teaching	Fall: Aug-Sep 2014 Spr: Dec-Jan 2015	77%
PST Post-Student Teaching	Fall: Dec 2014-Jan 2015 Spr: May-June 2015	60%
CT Post-Student Teaching	Fall: Dec 2014-Jan 2015 Spr: May-June 2015	73%

Note: Over half (53%) of PSTs in our sample completed both a pre and a post student teaching survey. For more details about the response rates and samples see Matsko et al. (conditionally accepted).

Appendix Table 2: Comparing Analytic vs. Non-Analytic Sample of PSTs, CTs, and FPSs

	<i>Non-Analytic Sample (n=817)</i>		<i>Analytic Sample (n=305)</i>		Chi-Square/ T-Test Difference
	N	Percent/ Mean	N	Percent/ Mean	
<i>PSTs</i>					
Male	723	26.0%	304	28.0%	0.4
White	817	56.7%	305	57.4%	0.0
Latino	817	14.1%	305	22.0%	10.2**
Black	817	7.1%	305	10.2%	2.9
Asian	817	6.6%	305	5.6%	0.4
Other	817	17.3%	305	5.9%	23.5***
Graduated from CPS	753	19.1%	305	28.5%	11.2**
Undergraduate GPA (1-100 scale)	727	89.1	305	91.6	2.5***
Prior teaching experience [^]	424	31.4%	220	30.5%	0.1
Age during student-teaching [^]	422	25.6	218	26.2	0.8
Parent during student-teaching [^]	425	11.3%	219	11.9%	0.0
Self-perceived preparedness [^]	431	0.0	225	-0.0	0.0
CT-perceived preparedness [^]	439	-0.0	226	0.1	0.1
<i>CTs</i>					
Male	563	25.4%	282	24.1%	0.2
White	563	69.6%	282	66.0%	1.2
Latino	563	20.1%	282	23.4%	0.2
Black	563	10.1%	282	11.0%	1.3
Asian	563	6.4%	282	3.2%	3.8
Other	563	3.7%	282	5.3%	1.2
Graduated from CPS [^]	431	26.5%	220	28.6%	0.4
Advanced degree (e.g. M.A.)	563	76.6%	282	78.0%	0.2
Tenure	562	91.8%	280	92.1%	0.1
National Board certification	562	18.7%	280	18.6%	0.0
Years of CPS service	555	11.8	278	12.1	0.2
Average Observational Rating	555	3.5	279	3.5	0.0
Reading VAM	143	0.2	86	0.2	0.0
Math VAM	106	0.2	68	0.1	-0.1
PST Perception of CT Teaching Effectiveness (Standardized Rasch) [^]	425	0.0	220	-0.0	0.0
<i>Field Placement Schools</i>					
Concentration of poverty (scon)	608	-0.1	295	-0.0	0.1**
High school level (9-12)	615	27.6%	299	33.4%	3.3
Majority Black students	608	10.2%	295	13.6%	2.2
Majority Latino students	608	34.9%	295	40.7%	2.9
School prior achievement	600	0.2	292	0.1	-0.0
PST-Perceived working conditions (3.5 or higher, out of 4) [^]	410	61.8%	222	62.6%	0.0
5E: Involved Families	589	67.1	282	65.9	-1.1
5E: Collaborative Teachers	592	62.6	282	61.4	-1.4
5E: Ambitious Instruction	592	67.2	285	67.0	-0.1
5E: Supportive Environment	589	56.9	285	55.1	-1.4
5E: Effective Leader	589	54.6	282	53.5	-1.2

[^] These items draw on PST or CT survey items and therefore have lower sample sizes than other PST items, which draw on registration data or other CT items, which draw on personnel data. (Note: +p<0.10; *p<0.05; **p<0.01; ***p<0.001)

Appendix Table 3: Comparing Within Analytic Sample of PSTs and CTs

	<i>Do Not have Focal Preparedness Measure</i>		<i>Have Focal Preparedness Measure</i>		Chi-Square/ T-Test Difference
	N	Percent/ Mean	N	Percent/ Mean	
<i>PSTs</i>					
Male	80	32.5%	224	26.3%	1.1
White	80	48.8%	225	60.4%	3.3
Latino	80	30.0%	225	19.1%	4.1
Black	80	15.0%	225	8.4%	2.8
Asian	80	2.5%	225	6.7%	1.9
Other	80	3.8%	225	6.7%	0.9
Graduated from CPS	80	35.0%	225	26.2%	2.2
Undergraduate GPA (1-100 scale)	80	90.8	225	91.9	1.1
Prior teaching experience [^]	N/A (These items are based on PST survey items, so no PST has this information and lacks a focal preparedness measure)				
Age during student-teaching [^]					
Parent during student-teaching [^]					
<i>CTs</i>					
Male	57	28.1%	225	23.1%	0.6
White	57	57.9%	225	68.0%	2.1
Latino	57	21.1%	225	24.0%	0.2
Black	57	12.3%	225	10.7%	0.1
Asian	57	1.8%	225	3.6%	0.5
Other	57	7.0%	225	4.9%	0.4
Advanced degree (e.g. M.A.)	57	71.9%	225	80.0%	1.5
Tenure	57	93.0%	223	91.9%	0.1
National Board certification	57	7.0%	223	21.5%	6.3*
Years of CPS service	56	12.6	222	11.9	-0.6
Average Observational Rating	56	3.4	223	3.5	0.1
Reading VAM	15	0.4	71	0.1	-0.3
Math VAM	13	-0.0	55	0.1	0.1
PST Perception of CT Teaching Effectiveness (Standardized Rasch)	50	-0.0	170	-0.0	0.0

Note: +p<0.10; *p<0.05; **p<0.01; ***p<0.001

Appendix Table 4: Correlation Matrix of PSTs' First-Year Observational Rating and Domain-Specific Perceived Preparedness

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Overall Self-Perceived Preparedness	1.00														
2. Overall CT-Perceived Preparedness	0.06	1.00													
3. Overall Observational Rating	0.03	0.24	1.00												
4. Self-Perceived: Planning	0.76	0.02	0.07	1.00											
5. Self-Perceived: Instruction	0.88	0.00	0.02	0.63	1.00										
6. Self-Perceived: Environment	0.82	0.07	0.04	0.45	0.67	1.00									
7. Self-Perceived: Prof. Responsibilities	0.75	0.07	-0.10	0.47	0.54	0.51	1.00								
8. CT-Perceived: Planning	0.05	0.92	0.21	0.05	0.01	0.07	0.05	1.00							
9. CT-Perceived: Instruction	0.04	0.95	0.27	0.03	0.02	0.05	0.01	0.83	1.00						
10. CT-Perceived: Environment	0.03	0.91	0.23	-0.03	-0.05	0.07	0.07	0.76	0.84	1.00					
11. CT-Perceived: Prof. Responsibilities	0.07	0.91	0.19	0.04	-0.01	0.04	0.13	0.76	0.82	0.79	1.00				
12. Observational Rating: Planning	-0.02	0.18	0.91	0.06	-0.03	-0.02	-0.12	0.16	0.21	0.15	0.14	1.00			
13. Observational Rating: Instruction	0.10	0.22	0.89	0.11	0.10	0.13	-0.04	0.20	0.26	0.20	0.16	0.73	1.00		
14. Observational Rating: Environment	0.05	0.27	0.93	0.09	0.05	0.05	-0.07	0.22	0.30	0.24	0.22	0.81	0.87	1.00	
15. Observational Rating: Prof. Responsibilities	-0.03	0.21	0.82	-0.03	-0.03	0.01	-0.12	0.19	0.21	0.24	0.15	0.67	0.59	0.64	1.00

Appendix Table 5: Perceived Preparedness and First-Year Effectiveness as a Function of PST and CT Characteristics

	Self-Perceived PST Preparedness after Student Teaching		CT-Perceived PST Preparedness after Student Teaching		First-Year Observational Rating	
	A	B	A	B	A	B
<i>PST Characteristics</i>						
Male	-0.27+ (0.14)	-0.30* (0.15)	-0.21 (0.14)	-0.23 (0.21)	-0.06 (0.04)	-0.05 (0.06)
Asian (vs. White)	-0.24 (0.25)	-0.22 (0.29)	0.03 (0.27)	0.03 (0.34)	0.12 (0.09)	0.18 (0.12)
Black (vs. White)	0.15 (0.22)	0.06 (0.23)	-0.09 (0.20)	-0.17 (0.27)	0.01 (0.07)	-0.06 (0.10)
Latino (vs. White)	0.10 (0.16)	-0.02 (0.18)	0.03 (0.16)	-0.06 (0.22)	0.01 (0.05)	0.04 (0.08)
Other (vs. White)	-0.32 (0.25)	-0.42 (0.25)	-0.12 (0.26)	-0.30 (0.32)	-0.00 (0.08)	-0.07 (0.10)
Prior teaching experience	0.08 (0.13)	0.08 (0.14)	0.01 (0.16)	0.04 (0.17)	-0.02 (0.05)	-0.01 (0.05)
Undergraduate GPA	-0.01 (0.01)	-0.01 (0.01)	0.00 (0.00)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)
CPS graduate	0.19 (0.14)	0.20 (0.16)	0.06 (0.14)	-0.00 (0.20)	-0.02 (0.05)	-0.08 (0.06)
Parent during student teaching	0.18 (0.20)		-0.10 (0.24)		-0.10 (0.07)	
Age during student teaching	0.00 (0.01)		-0.02 (0.01)		-0.01+ (0.00)	
N		217		220		289
<i>CT Characteristics</i>						
Male	-0.16 (0.15)	-0.23 (0.19)	-0.04 (0.15)	0.06 (0.16)	0.01 (0.05)	-0.01 (0.06)
Asian (vs. White)	-0.00 (0.00)	-0.01 (0.00)	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)
Black (vs. White)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00+ (0.00)	-0.00 (0.00)
Latino (vs. White)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Other (vs. White)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Advanced degree	0.04 (0.15)	0.02 (0.18)	0.07 (0.15)	0.15 (0.16)	-0.03 (0.05)	-0.10+ (0.06)
CPS graduate	-0.19 (0.16)	-0.15 (0.17)	0.07 (0.14)	0.06 (0.15)	0.04 (0.05)	0.06 (0.06)
White (vs. Non-White)	-0.00 (0.00)		-0.00 (0.00)		0.001* (0.000)	
N		161		212		213

Notes: For the outcomes of PST and CT perceived preparedness, Model A includes each focal predictor (row) independently. In the case of models with first-year observation ratings as outcomes, Model A also includes covariates for current 15-16 school characteristics (school grade level, school area poverty concentration, student race composition, school prior achievement, and whether the PST student-taught in the current school). For all outcomes, Model B includes focal predictors together in the same model and *also* includes PST and field placement school covariates (PST race, PST gender, placement school grade level, and placement school prior achievement). (+ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Appendix Table 6: PST and CT Perceptions of Preparedness Rasch Measures

<p>Planning and Preparation (PST reliability=.90; CT reliability = .89) Planning Lessons Designing student assessments Selecting instructional Outcomes Using results From Assessments to Improve Teaching Anticipating student misconceptions about content when planning for class</p>	<p>Classroom Environment (PST reliability=.90; CT reliability = .89) Developing Relationships With Students Managing Students' Behaviors Implementing Classroom Routines & Procedures Developing Classroom Communities for Learning</p>
<p>Instruction (PST Reliability = .91; CT reliability = .92) Using developmentally appropriate instructional language Posing variety of questions to probe student understanding Facilitating Discussions Maintaining student interest Using variety of instructional methods Adapting curricula to fit students' needs Teaching subject matter</p>	<p>Professional Responsibilities (PST reliability = .89; CT reliability = .85) Maintaining Accurate Grades and Student Data Performing Administrative tasks Interacting with school administrators Communicating With Families Reflecting on teaching (CT only)</p>

Appendix Table 7: Coaching Rasch Measures

<p>PST: CT Teaching Effectiveness (Reliability = 0.88) <i>How effectively did your mentor teacher:</i> Plan lessons Deliver instruction Create & maintain a positive classroom environment Model professionalism Teach in culturally responsive ways</p>	<p>CT: Domain-Specific Mentoring (Reliability = 0.72) <i>How effective do you feel you were in mentoring your student teacher in each of the following areas:</i> Planning lessons Delivering instruction Creating & maintaining a positive environment Modeling professionalism Teaching in culturally responsive ways</p>
<p>PST: Field Instructor Helpfulness (Reliability = 0.89) <i>How often did your field instructor/supervisor conduct the following activities:</i> Observed PST teach Provided PST with feedback on teaching Observed PST teach frequently enough Provided PST feedback frequently enough Provided feedback that helped PST learn to teach Would recommend field instructor to future PSTs</p>	<p>CT: Frequency of Feedback (Reliability = 0.76) <i>Think about the times you provided feedback to your student teacher about her/his instruction. How often did you:</i> Offer concrete suggestions Ask reflective questions Offer general observations Refer to specific things the PST needed to improve Refer to specific things the PST did well Share specific data when providing feedback</p>
<p>PST: Domain-Specific Conversations (Reliability = 0.86) <i>How much did you learn about the following skills from conversations you had with your mentor:</i> Planning lessons Delivering instruction Creating & maintaining a positive environment Modeling professionalism Teaching in culturally responsive ways</p>	<p>CT: Job Search Assistance (Reliability = 0.89) <i>With your student teacher, how often did you:</i> Offer advice on kinds of jobs to apply for Discuss specific job openings in the FPS Discuss specific job openings elsewhere Offer feedback on PST's resume Help PST prepare for an interview</p>
<p>PST: Mentoring Relationship and Feedback (Reliability = 0.87) <i>To what extent do you agree/disagree with the following:</i> CT observed me teach frequently enough CT provided feedback frequently enough Feedback CT provided helped me learn to teach Feedback CT gave was consistent with field instructor If I struggled with teaching, I could go to CT for help CT's expectations were appropriate for a beginner CT let me make my own instructional decisions Felt comfortable taking instructional risks in front of CT</p>	
<p>PST: Job Search Assistance (Reliability = 0.85) <i>How often did your mentor teacher:</i> Offer advice on kinds of jobs to apply for Discuss specific job openings in the FPS Discuss specific job openings elsewhere Offer feedback on PST's resume Help PST prepare for an interview</p>	