



It's Selective, But Is It Effective? Exploring the Predictive Validity of Teacher Selection Tools

Dan Goldhaber

Cyrus Grout

Nick Huntington-Klein

Center for Education Data and Research

Abstract: Evidence suggests teacher hiring in public schools is ad-hoc and often does not result in good selection amongst applicants. Some districts use structured selection instruments in the hiring process, but we know little about the efficacy of such tools. In this paper we evaluate the ability of applicant selection tools used by the Spokane Public Schools (SPS) to predict three outcomes: measures of teachers' value-added contributions to student learning, teacher absence behavior, and attrition rates. We observe all applicants to the district, both those who are and who are not hired. We find that the screening instruments predict teacher value-added in student achievement and teacher attrition, but not teacher absences.

This policy brief is based on [CEDR Working Paper 2014-9](#). We acknowledge support from the Institute of Education Science's Researcher-Practitioner Grant Program (Grant #R305C130030) and by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER) funded through grant #R305A060018 to the American Institutes for Research from the Institutes of Education Sciences, U.S. Department of Education. This research has benefitted from the helpful input of Heather Hill, Angela Jones, Susanna Loeb, Jonah Rockoff, Mary Templeton, and members of Spokane Public Schools' stakeholder group, as well as data management and analysis support from Andrew Katz, Malcolm Wolff, Patricia Martinkova, Roddy Theobald, and Nate Brown. We also thank all data providers, including Spokane Public Schools, the Washington Office of the Superintendent of Public Instruction (OSPI), and the Washington State Information Processing Collective (WSIPC). Any and all errors are solely the responsibility of the study's authors, and the views expressed are those of the authors and should not be attributed to their institutions, the study's funders, or the agencies supplying data.

Suggested citation:

Goldhaber, D., Grout, C., and Huntington-Klein, N. (2014). It's Selective But Is It Effective? Exploring the Predictive Validity of Teacher Selection Tools. CEDR Policy Brief 2014-9. University of Washington, Seattle, WA.

© 2014 by Dan Goldhaber, Cyrus Grout, and Nick Huntington-Klein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission, provided that full credit, including © notice, is given to the source.

You can access other CEDR publications at
<http://www.CEDR.us/publications.html>

1. A Focus on Teacher Applicant Screening

Teachers can have profound effects on student achievement, not only on standardized tests, but on long term outcomes such as going to college and future earnings in the labor market. Not surprisingly, the last decade has seen a considerable amount of research and policy attention directed toward interventions that might improve the quality of the teacher workforce. However, we know very little about how school systems decide which teachers to let into the workforce to begin with. This is surprising because the importance of the hiring process; when a teacher is hired, districts are making what may turn out to be a large long-term financial commitment. Given this, it is important to make sure the recruitment process works well.

This policy brief reports the findings of a recent study analyzing the relationship between teacher selection rubrics used during the hiring process at Spokane Public Schools and three teacher outcomes: value-added measures of effectiveness, teacher absence behavior, and the likelihood of attrition.

Three aspects of our study make it unique. Unlike previous studies of the hiring process, we observe employment outcomes for applicants who are hired into SPS and applicants who are not hired into SPS but are employed in other public school districts in Washington State. This allows for a more comprehensive analysis. In fact, as we describe below, the ability to observe teachers who end up employed outside of SPS affects the interpretation of the value of the screening rubrics given that the relationship between applicant ratings on the rubrics and teacher outcomes varies along the applicant performance distribution and, not surprisingly, Spokane tends to employ teachers who score toward the top of the distribution.

Second, we observe whether a prospective teacher was offered a job, not just whether a teacher is employed in a position. Given this, we are able to distinguish between job-applicant non-matches (i.e. an applicant is not employed in SPS) that result from employee versus employer preferences. We find that teachers in nearly all (95 percent) job searches that include an offer from Spokane accept the offer. The difference between a teacher who ends up employed in Spokane and a teacher who does not is then largely a decision on the district's part.

Finally, we are able to correct for selection bias that may arise from not being able to observe the outcomes of teachers who are not employed in public schools in Washington. Our selection-corrected estimates exploit the fact that a non-trivial proportion of the scores teachers receive on the selection instruments are incorrectly computed due to procedural oversight or arithmetic mistakes, as well as the differing amounts of competition faced by applicants when applying for SPS jobs; these factors are assumed to influence the likelihood of being hired but should not otherwise be related to teacher quality.

We find that teachers hired by Spokane are more effective (as measured by value-added) than applicants who end up employed by a different school district in Washington. Hired applicants also tend to be absent more often and are less likely to leave their district. The scores on the two selection rubrics used by Spokane are associated with these differences. Screening scores have strong relationships with teacher effectiveness in student achievement as well as teacher attrition, and the magnitudes of these relationships are educationally meaningful: A one standard deviation increase in screening scores is associated with an increase of about 0.07 standard deviations of student math achievement, a marginally significant increase of 0.03–0.05 standard deviations of student reading achievement, and a decrease in teacher attrition of 2.5

percentage points. Correcting for selection into a SPS job does not significantly change the findings, suggesting that teachers who do not end up employed in Washington's public schools do not significantly bias the estimates.

These findings are evidence that public schools can improve the quality of the teacher workforce through the use of well-designed applicant selection tools. However, our analyses of the screening rubrics show much stronger relationships for some subcomponents than others, implying that the teacher outcomes we assess could be further improved by weighting certain subcomponents, such as Classroom Management, more strongly than others.

2. Background

The potential for improving workforce quality through effective hiring practices is broadly supported by research in the field of personnel economics and industrial psychology. Studies show that screening based on experience and training improve the worker selection. However, these broad meta analysis studies are limited only to workers that were hired and do not observe those job candidates who were not selected in the screening process and literature about the selection and hiring process for public school teachers is relatively thin.

Under the current approach to hiring in public schools, there is mixed evidence about whether the best applicants are hired. A few studies look at whether schools identify the best applicants in terms of a direct measure of teacher effectiveness, value-added, as opposed to the data available to the school system at the time of hiring, but here too the evidence is mixed. Hanushek et al. (2005) analyze whether schools that offer higher levels of compensation (in the form of both salary and work environment) tend to hire more effective teachers. They find that these schools (generally suburban schools) are more likely to hire teachers with advanced degrees, but find no relationship between hiring and value-added measures of effectiveness.

To our knowledge there are only two studies that directly connect the information available to employers about teacher applicants to the productivity of those applicants once employed, and both are limited to assessments of those applicants who are hired. Dobbie (2011) investigates the link between information used to select Teach for America (TFA) members and the future impact that TFA teachers have on student achievement during the first year of teaching. He finds that a one standard deviation change in an index that averages the standardized measures used to select amongst TFA applicants is predicted to increase student achievement by about 0.15 standard deviations on a math exam (the findings for reading achievement were smaller and not statistically significant).

Rockoff et al. (2011) examine the extent to which traditional (e.g. degree, licensure) and non-traditional (e.g. extraversion, personal efficacy) information about teacher applicants are related to teacher value-added scores, subjective teacher ratings, teacher absences, and retention. The authors find that few individual metrics are significant predictors of teacher effectiveness, but a one-standard deviation increase in distilled measures of cognitive and non-cognitive skills are each associated with significant increases in student math achievement. The variation of predicted value-added using both traditional and non-traditional information explains about 12 percent of the expected variance in teacher effectiveness, compared to about four percent using only traditional information. These findings suggest the quality of the workforce could be improved by collecting and using additional information about prospective teachers.

In sum, the relatively thin literature on the teacher hiring process generally supports the notion that it is possible to improve the quality of the workforce through better hiring. In particular, it is suggestive that one way to do this is by using screening instruments, such as the rubrics Spokane Public Schools use.

The Hiring Process in Spokane Public Schools

The hiring process, outlined in Figure 1, includes four stages following a school job posting:

- 1) Acceptance of applications
- 2) 21-point pre-screening of applicants by Human Resources (HR) hiring officials
- 3) 60-point screening of applicants by school-level hiring officials
- 4) In-person interview and hiring decision

Job applicants must progress through two stages of screening before they become eligible to interview for a position. The first stage, “pre-screening”, is conducted under the direction of HR, using a 21-point scoring rubric with three subcomponents: Experience, Depth of Skills, and Recommendations. An applicant’s pre-screening score is not directly associated with his or her application to any particular job in the district. The applicant is scored on each criterion on a scale of 1 to 6, and the Recommendations score is then multiplied by 1.5. Table 1 describes what the screener looks for in an applicant’s profile to score each criterion. Scores of 1–2, 3–4, and 5–6 indicate the finding of, respectively, “some”, “satisfactory”, and “strong” evidence that the criterion is an area of strength for the applicant.

Table 1. Evaluation Criteria on 21-Point Pre-Screening Rubric

Criterion	Screener should look for...
Experience related to position	Years of experience, type of experience, type of school/district, gaps in teaching experience
Depth of Skills	Evidence of strong content knowledge, strong classroom management, differentiates instruction, engages parents/families, strong rapport with students and colleagues, commitment to the school as a community, socially just practices, experience with diverse student populations, makes learning relevant, engages students in active learning, elementary level currently seeking those with experience using Fosnot, Calkins, GLAD strategies, response to intervention
Quality of recommendations	All items noted in above categories, does writer recommend/strongly recommend, personal or professional recommendation, does the writer regularly evaluate teachers (preference of letter from principal, asst. principal, instructional coach, supt.)

The 21-point score narrows the applicant pool to a manageable size before the second school-level stage of screening, which is led by principals. In the second stage a principal requests a list of applicants from HR for consideration. These requests typically specify a cutoff score on the 21-point screening. For example, a principal hiring a second grade teacher could request a list of all applicants with an endorsement in Elementary Education and a 21-point score of greater than 17. Principals and HR follow these procedures closely.

The third stage selects which candidates receive in-person interviews using a 60-point rubric with 10 evaluation criteria and the same 1–6 scoring scale as the 21-point rubric. The screening criteria are outlined in Table 2. Each applicant may be evaluated by one or more screeners, and principals use the screening scores as the basis for selecting which applicants to interview.

Table 2. Evaluation Criteria on 60-Point Screening Rubric

Criterion	Screener should look for...
Certificate and Education	Note completion of course of study; certificate held (current or pending); education.
Training	Look for quality, depth, and level of candidate’s additional training related to position.
Experience	Note degree to which experience supports the prediction of success, not just the number of years. A beginning candidate could be rated highly.
Classroom Management	Look for specific references to successful strategies. This may not mean quiet and orderly, but planned and directed. Effectively handles large/small or ethnically/sociologically diverse groups; develops routines and procedures to increase learning, establishes clear parameters, and responds appropriately.
Flexibility	Note multiple endorsements, activity, coaching interests, student, building or district, or community support. Willing to learn new concepts and procedures, successfully teaches a variety of assignments, effectively uses various teaching styles.
Instructional Skills	Look for specific references in support of skill in this area – plans, implements, evaluates, relates to students, creative, multiple approaches, monitors and adjusts, uses culturally responsive strategies appropriate to age, background and intended learning of students.
Interpersonal Skills	Develops and maintains effective working relationships with diverse staff, students, parents/guardians, and community.
Cultural Competency	Look for specific references to successful strategies for building and maintaining a relationship with each student and their family. This may not be explicitly mentioned, but the following strategies offer some evidence of cultural competency: specific instructional strategies providing each student access to a rigorous curriculum, inclusive/respectful language about students and families, a belief that all children can achieve at high levels, mention of conflict resolution/restorative practices, specific instructional strategies for integrating culturally responsive materials which are also rigorous, and appropriate statements about their work with diverse populations. Note relevant training, course work, authors/book titles listed.
Preferred Qualifications	Applicant has preferred qualifications as indicated in the job posting.
Letters of Recommendation	Look for current letters of recommendation from the most recent supervisor(s). Your score should reflect the quality and recentness of the recommendation, as well as the author of the letter.

The final stage of hiring consists of in-person interviews. A team assembled by the principal interviews applicants selected through earlier screening stages. The principal has discretion over

the content of the interview and there is no set interview evaluation criteria. Following the interviews, the principal submits a “request to hire” form, along with copies of the interview questions and scoring sheets. After background checks and with final approval from the district’s HR department, a job offer can be made.

3 Data and Methods

We link administrative data at the applicant, job, teacher, student, and school levels. These data allow us to analyze the relationship between teacher screening scores and both teacher value-added, absences, and attrition, controlling for school, teacher, and student characteristics.

Student data come from the statewide Core Student Record System (CSRS). CSRS includes information on student demographics and test scores as well as students’ teacher assignments, which can be used to link students to teachers and schools. Teacher and applicant data come from multiple sources. Information on each applicant is provided by SPS, including records of which jobs each applicant applied to, data on applicant characteristics, the scoring of applicants on screening instruments, the progression of applicants through the hiring process (whether the applicant is screened, interviewed, offered a job, and ultimately hired), and information on teacher absences for those who are hired into Spokane.

We link data on Spokane applicants to statewide teacher data sets using unique teacher certification numbers. These data include teacher licensure test scores and areas of endorsement collected by the Professional Education and Standards Board (PESB), and teacher absence data for teachers who do not work in Spokane collected by the Washington School Information Processing Cooperative (WSPIC). We also link applicants to the S-275 personnel report, which provides a record of all certificated employees of public school districts in Washington State. Data on school characteristics come from Public School Universe data generated by the National Center for Education Statistics (NCES).

Three measures are used to evaluate teacher outcomes: grade 3-8 student performance on Washington State’s annual assessments of student learning for math and reading, teacher absences for each day of the week in 2012 and 2013, and teacher retention. Teacher outcome data are linked to the most recent screening scores. So, for example, consider a teacher employed continuously in Washington State from 2009-2012. If that teacher applied to Spokane in both 2009 and 2010, then the 2011 and 2012 teacher performance outcomes are linked to the 2010 application, the 2010 outcomes are linked to the 2009 application, and the 2009 outcomes are not used. In this manner, we are able to match 274 applicant-year observations to student test score data, and 502 applicant-year observations to absence data. Teacher retention in the district is determined by matching applicants to the S-275 personnel records for the school years ending between 2010 and 2013. For the 2009 applicants, we are able to observe whether a teacher returns after up to four years of service, and for the 2012 applicants we can identify who returns after up to one year of service. We are able to match 736 applicant-year observations to certificated employment records in the S-275 data.

4 Findings

Teacher Effectiveness

We found the applicant scores on the 21-point rubric to have a positive but insignificant relationship with teacher effectiveness in math and reading. The relationship between the 60-

point rubric and teacher effectiveness is larger than the 21-point rubric for both subjects and is statistically significant for math. We believe the results for both rubrics—including the non-significant results—are educationally meaningful. For instance, students assigned to teachers who score one standard deviation higher on the 60-point rubric are predicted to have achievement results that are similar to the difference in achievement associated with being assigned to a novice teacher versus a second or third year teacher.

Of the subcomponents on the 60-point rubric and their relationship to teacher effectiveness—Classroom Management is relatively large for both math and reading. Training, Flexibility and Instructional Skills are also significant and large for math. On the other hand, there is lack of significant impact for Certificate & Education. This is noteworthy given that many schools highly value certification when evaluating teacher applicants. The fact that the significance of the subcomponents varies substantially suggests that re-weighting these subcomponents could increase the ability of the rubric to predict teacher effectiveness.

Teacher Absences

For teacher absences, the 21-point rubric score was a positive predictor of teacher absences whether individually or in tandem with the 60-point screening. The point estimate suggests that a one standard deviation increase in screening score is predicted to increase teacher absences by about a half of a day. The total 60-point screening score is insignificant in each specification and few of the subcomponents are significant.

Previous research found a strong positive relationship between experience and teacher absences. We test whether the magnitude of the relationship between the rubric scores and teacher absences are related to the fact that rubrics value experience by including experience in the model. When we do this, the magnitudes of both summative ratings and specific components decrease substantially, with the 21-point score becoming insignificant and the 60-point score becoming slightly more negative. This suggests the relationship between the rubric scores and teacher absences is indeed related to the fact that the rubrics reward experience.

Teacher Attrition

Applicants scoring higher on each of the rubrics are less likely to leave the district. These results are driven by a few of the subcomponents. On the 21-point rubric, the Depth of Skills component is most strongly predictive for each type of attrition. The 60-point components, like in mathematics achievement, have non-equal coefficients, and predictive power is centralized in a subset of the components: Experience, Classroom Management, Flexibility, Instructional Skills, and Interpersonal Skills and Preferred Qualifications all significantly predict less of each type of attrition.

5 Policy Implications

Our findings show that the two screening rubrics used by Spokane Public Schools predict teacher effectiveness and teacher attrition in expected ways. For some perspective on this, a one-standard deviation increase in the 60-point screening score is associated with about a 0.07 standard deviation increase in math achievement, a marginally significant 0.03–0.05 increase in reading achievement, and a decrease in attrition by about 2.5 percentage points. Given that the turnover of a single teacher can cost a district in the region of \$10,000 (Barnes, Crowe, &

Schaefer, 2007), improved hiring practices have the ability to both save money and improve effectiveness.

The screening scores used by SPS represent the value of guided human interpretation of somewhat subjective information, such as that contained in letters of recommendation, and our findings validate the notion that this type of guidance on how to interpret applications is an improvement on the ad-hoc hiring processes typically seen in public schools (e.g., Ebmeier & Ng, 2006; Oyer & Schaefer, 2011).

While the hiring rubrics appear to be effective, analysis of the subcomponents suggests that Spokane could increase the predictive validity of the summative rubric ratings by re-weighting the components. To illustrate this, we estimate the canonical correlation between the various subcomponents of the 60-point rubric and each of the teacher outcomes in order to derive the weights that optimize the correlation between the rubric rating and the outcomes.

The weights that maximize the outcomes are different for each outcome. The only component which is given more than a 0.1 weight in every outcome in Certificate & Education; while Certificate & Education was not a strong predictor of teacher effectiveness by itself, it is not strongly correlated with the other subcomponents and so enters positively here because it provides some measure of information beyond the other factors. There are, however, several subcomponents that could be significantly and consistently down-weighted across all different outcomes, suggesting that some aspects of the hiring rubric, or the training on how to implement it, could be improved. These results also suggest the capacity to improve teacher training. If teacher training is able to focus on improving the skills found here to relate strongly to teacher effectiveness, then incoming teachers will be more capable of improving student learning.

There are, however, at least two cautions that should be applied to interpreting the above findings. First, we do not know whether some of the subcomponents that do a poor job of predicting the outcomes studied here might predict other teacher behaviors or student outcomes that are valued by school districts but not measured.

Second, and more generally, our findings for Spokane may not generalize across all school districts. Spokane is seen as a desirable place to work in Eastern Washington and may not face the same hiring problems as other districts, who may face more competition for teacher labor. Spokane also hires a high percentage of its workforce from people who already have experience there. Just over 70 percent of jobs were filled with applicants who either already worked or student taught in Spokane. It is possible that the predictive abilities of the screening rubrics are aided by the fact that the screeners may be familiar with those who are writing the letters of recommendation. Having said that, these results are consistent with the wider literature on screening at the hiring stage in other industries and in teaching.

The idea of improving the quality of the teacher workforce through more effective hiring is appealing given the political and financial costs of dismissing ineffective teachers who are in-service (Treu, 2014), and empirical evidence that finds that other teacher performance interventions, such as professional development or performance incentives, tend to have marginal impacts on productivity. The evidence we present here shows a strong relationship between the performance on selection instruments and some measures of in-service teacher quality. This relationship likely overstates what is possible in terms of improving the teacher workforce as a whole since school systems compete with one another in the market for teacher

labor. Nevertheless, given that many school districts rely on much more informal processes for selecting teachers, and likely lose some potentially talented teachers to other occupations at the hiring stage, there appears to be substantial room for improving the quality of the teacher workforce through greater use and refinement of teacher hiring tools and practices.

References

- Barnes, G., Crowe, E., & Schaefer, B. (2007). The Cost of Teacher Turnover in Five School Districts: A Pilot Study. National Commission on Teaching and America's Future.
- Dobbie, W. (2011). *Teacher Characteristics and Student Achievement: Evidence from Teach For America*. Harvard University. Cambridge, MA.
- Ebmeier, H., & Ng, J. (2006). Development and Field Test of an Employment Selection Instrument for Teachers in Urban School Districts. *Journal of Personnel Evaluation in Education*, 18(3), 201–218.
- Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). The Market for Teacher Quality *NBER Working Paper No. 11145*.
- Oyer, P., & Schaefer, S. (2011). Personnel Economics: Hiring and Incentives. In D. Car & O. Ashenfelter (Eds.), *Handbook of Labor Economics* (Volume 4., pp. 1769–1823). Elsevier.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can You Recognize an Effective Teacher When You Recruit One? *Education Finance and Policy*, 6(1), 43–74.
- Treu, R. M. Vergara vs. State of California Tentative Decision (2014).
- U.S. Department of Education. (1997). *Credentials and Tests in Teacher Hiring: What Do Districts Require ?*

The Pipeline to the Classroom at Spokane Public Schools

