Does quality principal feedback help improve teaching practices? Perspectives from the teachers and students in an authentic evaluation system

Xintong Li
University of Missouri

Jihyun Kim
Lehigh University
Introduction

In the past, measurement of teacher effectiveness was not rigorous and typically did not discriminate between more and less effective teachers (Weisberg et al., 2009). To address this problem, federal policy, through Race to the Top and ESEA waivers, influenced many states to establish new, more rigorous teacher evaluation systems with multiple measures of teaching quality in recent years (Ballou & Springer, 2015; Harris, Ingle, & Rutledge, 2014; Herlihy et al., 2014; Pogodzinski, Umpstead, & Witt, 2015; Steinberg, Sartain, & Policy, 2015). Most states now require school districts to implement teacher evaluation systems to ensure in-service teachers are effective and to provide accurate feedback to improve teaching quality (Doherty & Jacobs, 2013). There is widespread agreement that high-quality evaluation systems use multiple measures from different sources for better conceptual coverage and stability (Grissom & Youngs, 2016; Popham, 2013).

For a time, these new, rigorous systems emphasized measuring teacher effectiveness using students’ achievement gains, or value-added measures. However, while seemingly reasonable on the surface, these measures garnered significant concerns and criticism. Among these are technical issues of how estimates are derived, the non-random assignment of students to teachers, the limited availability of tests in all subjects and grades, and the inability of this measure to provide useful feedback for teachers to guide instructional improvement (Briggs & Domingue, 2011; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Hallinger, Heck, & Murphy, 2014; Hill, Kapitula, & Umland, 2011; Rothstein, 2010). Organizations such as the National Association of Secondary School Principals and the American Statistical Association have taken a stand against value-added measures.
In contrast to value-added measures, the use of classroom observation for measuring teacher effectiveness has grown and is now the most widely used measure to evaluate teachers (Cohen & Goldhaber, 2016; Steinberg et al., 2015). One of the most prevailing assumption of using classroom observation is that it affords ample opportunities for teachers to receive meaningful feedback so that they can improve their instructional practices while VAM at best merely provides information about where the teachers are at in the spectrum of teaching quality. However, we know little about how the quality of feedback is related to the improvement of teaching quality, and more importantly, what makes feedback helpful for teachers.

As a first step to address this urgent issue, we ask two questions as follow.

1. How is teachers’ perceived quality of feedback related to their improvement of teaching quality?

2. Which aspect of principal feedback is associated with the improvement of teaching quality the most?

Another unique contribution that this study makes is that we use student survey as a measure of teaching quality rather than student achievement. Student ratings of teaching effectiveness are reliable with much more data points than classroom observation and predictive of student learning (Downer, Stuhlman, Schweig, Martínez, & Ruzek, 2014; Feldlaufer, Midgley, & Eccles, 1988; Fraser & O’Brien, 1985; Little, Goe, & Bell, 2009). Students spend much more time with teachers than any other external observers, and they are the ones that have first-hand experience with teachers’ instructional practices. Indeed, the Measures of Effective Teaching study found that student ratings of teachers predicted the student’s growth on achievement tests and that they were more reliable than achievement data or classroom observation (MET; Kane &
Staiger, 2012b). Given the various limitations of using student test scores to measure teaching quality, using student survey renders itself as a promising alternative.

**Quality Feedback for the Improvement of Teaching Practices**

Most teacher evaluation systems in the U.S. and other counties include pre- and/or post-observation conference between evaluators and teachers as a requirement. Specific practices of feedback sessions vary, but observers generally debrief what they observed and provide suggestions for improvement based on observation rubric. In most cases, the format of such feedback sessions, such as the number of meetings between evaluator and teachers, when the sessions have to happen, is generally determined by district- or state-level policies (Kim & Sun, under review). Critically, however, the substantive aspects of such sessions (i.e., where to focus on during the feedback session) largely depend on evaluators, which suggests variation in efficacy of such feedback sessions.

Recent studies show that providing focused and frequent feedback can induce positive changes in teaching quality and student achievement (Garet et al., 2017; Steinberg et al., 2015). Researchers generally agreed on the characteristics of high-quality feedback; based on descriptive and observable data (Danielson & McGreal, 2000); providing characteristics of effective teaching (Marzano, Pickering, & Pollock, 2001); promoting reflective inquiry and self-directedness to foster improvements in teaching supported by evidence (Glickman, 2002); frequent, specific and actionable (Curtis & Wiener, 2012).

However, the body of research that unravels the relationship between the quality of evaluators’ feedback and teacher effectiveness remains scant. Previously, feedback was regarded as a part of a rigorous evaluation system where high-quality feedback is simply assumed without
empirical examinations. This leads to the lack of an empirically backed definition of high-quality feedback and the consequent difficulties in evaluator training.

There are few well-established feedback quality measures that are incorporated in a widely employed teacher evaluation system. One is the teacher survey developed by the University of Chicago Consortium on School Research. This is called the 5Essentials Teacher survey and it has been administered to all teachers working in Chicago Public Schools every year since 2011 (Jiang, Sporte, & Luppescu, 2015). Their survey includes multiple items that asked teachers’ perceptions on their teacher evaluation system, REACH (Recognizing Educators Advancing CHicago). The items asked general questions about whether the feedback that they received was accurate and useful.

Another large-scale teacher survey that includes items related to teachers’ evaluation of feedback quality was developed by the University of Missouri Network for Educator Effectiveness (NEE). Their survey is called Principal Effectiveness Teacher Survey (PETS), which was developed based on the 10 Professional Standards for Educational leaders (PSEL; National Policy Board for Educational Administration, 2015). The purpose of the survey is to evaluate principals and it is deployed to teachers at the end of each school year. Some items on the survey specifically address teachers’ perceptions of principals’ feedback to them following classroom observations, which were developed based on the previous literature on the characteristics of effective feedback and corresponding suggestions (Feeney, 2007; Reinhorn, Johnson, & Simon, 2017). Those items are included to address the professional capacity of school personnel of PSEL.
The Outcome Measures for Feedback Quality

Previous studies used student achievement or classroom observation to investigate the effect of feedback quality (Garet et al., 2017; Steinberg et al., 2015). However, solely relying on teachers’ value-added measures has well-documented technical issues, such as the bias due to non-random assignment of students to teachers and limited test availability across subjects and grades. Moreover, student achievement is also a distal outcome that may compromise the power to detect the effect of feedback quality on teaching practices. Using principal classroom observation may bring in confounding effects, such as the teacher-principal relationship, as teachers and principals would be evaluating each other.

Student surveys, in contrast, would be a solid alternative outcome measure to bring in a third-party perspective on teacher effectiveness. They are not as widely used, but are emerging as an important tool for evaluating teachers. In teacher evaluation systems, student surveys are typically administered at the end of the class. Students are asked to rate their teachers on several specific behaviors (e.g., coming to class well-prepared, treating students fairly). Student surveys have been used to evaluate teachers at the post-secondary level for many years, but only recently in K-12 schools. Despite some concerns about children’s ability to rate teachers, ratings of teaching effectiveness by students above upper elementary grades are generally reliable (Kane & Staiger, 2012a; Marsh, Dicke, & Pfeiffer, 2019; Polikoff, 2015).

Moreover, student survey has some other advantages as an outcome measure of feedback quality. They are cost-efficient and based on many hours of interaction (whereas classroom observations are necessarily limited to brief snapshots of teaching), which means students can provide information not available to other observers, such as student-teacher relationship. Measurement error can also be attenuated because ratings are aggregated across many students.
(Feldlaufer et al., 1988; Fraser & O’Brien, 1985; Little et al., 2009). Furthermore, student survey responses predict student achievement (Downer et al., 2014; Fraser & McRobbie, 1995; Kane & Staiger, 2012a; Roeser, Eccles, & Sameroff, 2000). Indeed, the Measures of Effective Teaching study found that student ratings of teachers predicted the student’s growth on achievement tests and that they were more reliable than achievement data or classroom observation (Kane & Staiger, 2012a).

Many districts currently include a student surveys component in their teacher evaluation systems (e.g., Memphis, Chicago, Pittsburgh, and the New Teacher Project). Three surveys – the Tripod, My Student Survey and the Teacher Effectiveness Student Survey – are widely used for measuring teaching effectiveness, and there are also less commonly used student surveys available in the market. We decide to use Teacher Effectiveness Student Survey (TESS) which was also developed by NEE, as the major purpose of the current study is to collect authentic data from an existing teacher evaluation system and examine whether principal feedback on teacher evaluation results can predict better student rated teaching practices.

Therefore, the current paper drew on data from the Network for Educator Effectiveness (NEE), an authentic educator evaluation system that provides services to over 283 school districts throughout Missouri. TESS and PETS are two of the surveys that NEE deploys at the end of each school year. Both surveys were created by the researchers at the University of Missouri in partnership with school districts.

Method

Measures

Teacher Survey of Feedback Quality (TSFQ). 5 items in the Principal Effectiveness Teacher Survey specifically address the feedback quality that principal provide to teachers, and
we call them the teacher survey of feedback quality, or TSFQ in the current paper. The five items were developed by the Network for Educator Effectiveness at the University of Missouri (NEE; for more information, see www.neeadvantage.com). For example, teachers respond to the statement “This principal provides specific feedback to me regarding areas of strength in my teaching (i.e., focused, detailed, concrete).” TSFQ asks teachers to rate their principals on a four-point scale from 0 (strongly disagree) to 3 (strongly agree). Higher scores indicate more effective and useful classroom observation feedback.

Teacher Effectiveness Student Survey (TESS). TESS was developed by researchers at the University of Missouri based on the nationally recognized InTASC standards framework as moderately adapted by the Missouri Department of Elementary and Secondary Education (Interstate Teacher Assessment and Support Consortium, 2011). The Interstate New Teachers Assessment and Support Consortium (INTASC) was established by the Council of Chief State School Officers (CCSSO) to provide support to new teachers and raise the levels of learning in U. S. classrooms. It identifies 10 standards of the knowledge, dispositions, and skills expected of effective teachers. The Council of Chief State School Officers is a nonpartisan, nationwide, nonprofit organization that provides leadership, advocacy, and technical assistance on major educational issues. The standards apply to all subject areas and grade levels. Accredited institutions of higher education with teacher-education programs use the InTASC standards to define expectations for new teachers.

TESS is a modular survey and includes 39 indicators of teaching practices that school districts can choose from according to their priority. For the purposes of this study, to minimize missing data and response burden of participants, we focus on the two teaching practices most
often prioritized by network school districts and that have substantive effect sizes on student learning, i.e. cognitive engagement (CE) and problem-solving and critical thinking (PC).

Cognitive engagement (CE) examines the degree to which a teacher cognitively engages students in the content in their teaching practices (e.g., “This teacher expects us to think a lot and concentrate in this class”). 4 items are included, and higher scores indicate more perceived practices in cognitively engaging students. Problem-solving and critical thinking (PC) measures the extent to which a teacher uses instructional strategies that lead students to problem-solving and critical thinking in teaching practices (e.g., This teacher asks “how?” and “why?” questions to make us think more.”). 4 items are included, and higher scores indicate more noticeable teaching practices that induce problem-solving and critical thinking activities among students. Both measures are mean aggregated student ratings at the teacher level.

**Covariates.** We included two other measures in the PETS as covariates which may also play roles in improving teachers instruction practices and help adjust the teacher ratings of principal feedback quality. The first one is Principal Support for Professional Development (PSPD) developed by NEE researchers to jointly address the professional capacity of school personnel of PSEL. It measures how well the principal supports comprehensive professional development of staff. 4 items are included in this measure, e.g., teachers respond to the statement “This principal provides me with valid and meaningful professional development opportunities.” PSPD asks teachers to rate their principals on a four-point scale from 0 (strongly disagree) to 3 (strongly agree). Higher scores indicate a stronger principal support for teachers’ professional development. The other one is Principal Support for Professional culture (PSPC) which was developed to address the professional community for teachers and staff in PSEL and it measures how well principals promotes a positive and professional culture among teachers and staff. 5
items are included, e.g., “This principal develops positive relationships with staff. “This principal develops positive relationships with staff.” PSPC also asks teachers to rate their principals on a four-point scale from 0 (strongly disagree) to 3 (strongly agree), and higher scores indicate a stronger principal support for a professional culture in the school.

Teachers experience (Tch_exp) and student grades (Grade) are also included in the study. Teacher experience is self-reported years of teaching and student grades are student reported mean grades aggregated at the teacher level.

Participants

As TESS is a modular based survey and different school districts may prioritize different teaching practices to evaluate their teachers, two samples are collected for each of the TESS teaching practices where teachers took PETS and students took either TESS-CE or TESS-PC across school year 2017-2018 and 2018-2019. The CE sample includes 248 school principals and 2229 teachers, with the teachers’ CE scores based on the mean aggregation of 88822 student ratings. The PC sample involves 285 principals and 2203 teachers rated by 80209 student surveys. The focal measures show satisfactory psychometric properties and latent factor scores were used in the current study.

Since students in TESS and teachers in PETS are strictly anonymous, accurate description of the samples are impossible. However, what is known is that Missouri is both geographically and educationally at the center of the nation (e.g., national average 2017 NAEP 8th grade reading and math scores were 267 and 283, compared to Missouri’s 266 and 281). Participating districts in Missouri are diverse, serving both high- and low-income students in urban, suburban, and rural areas. Many serve very low-income, rural White students in the nations’ poorest counties. Overall, students in NEE school districts are 80% White, non-Hispanic
and 49% are eligible for free or reduced-price meals. The state averages are 73% and 50% respectively, suggesting NEE members are fairly representative of the state.

**Procedures**

Teachers complete the annual PETS on a web-based platform. The teachers have access to the survey when they sign in the NEE data tool website and click the survey button after the full name of a particular principal. Teachers will be given detailed instructions when their district join NEE and NEE user support personnel are available for any potential technical issues during working hours. Teachers will be notified at the beginning of the survey that their responses are “completely anonymous and no one in the district can view individual responses from any teacher”.

TESS was delivered online at the end of the school term within an accessible time period specified by principals. Students entered an access code unique to each teacher to ensure they were evaluating the right teacher and at the same time prevent unauthorized access to the evaluation. Students remained anonymous throughout the entire evaluation procedures and their responses are confidential. An adult other than the evaluated teacher must administer the survey using standard administration scripts provided by NEE. The proctor read instructions to the students, informed them of the purpose of the survey, the anonymity of their responses, the voluntary nature of the evaluation, and how this evaluation is important for school improvement. Students were encouraged to ask questions, or request the explanation of difficult words, but the proctors were instructed not to interpret any survey items to avoid possible influences on students’ responses. Moreover, though the online survey interface was designed to be intuitive and easy to use for students, the proctors also received simple training for potential technical issues.
TESS includes 3 screening items that are evenly distributed among the survey items (e.g., “I am being totally honest on this survey.”), which may help to improve survey validity and identify inattentive responses (Cornell, Klein, Konold, & Huang, 2012). Surveys that fail (score < 2) two of the three screening items are flagged for manual review. To minimize the threats to validity due to high turnover in some districts where students move during a school year, a screening question is also included asking if they have been in the class for at least a month. If they answer “no,” their survey is eliminated. Further, responses that are finished within unrealistic time (three standard deviations from the mean) are also flagged for manual review. The data involved in the current study were manually downloaded from the NEE data engine in de-identified form by a NEE staff with IRB approval at the end of 2019, and all the participating districts consent to use their data in the evaluation system for research purposes.

Results

The teacher survey of feedback quality (TSFQ) shows strong ceiling effects in both the CE sample (37.8%) and the PC sample (37.5%), which can be seen in Figure 1. Possible explanations include, 1) some teachers are concerned about the anonymity and/or potential consequences on themselves and/or on their principals, 2) some teachers are very lenient in evaluating their principals as they may have close personal relation, 3) some of the teachers are not attentive enough to the survey as they may not think they can benefit from the survey, 4) principals who received ceiling scores did a great job in providing quality feedback to the teachers, and 5) the scale is not well designed. The strong ceiling effect may also be a result of multiple factors.
Table 1. Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Sample size</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PC Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC_19</td>
<td>2203</td>
<td>-4.27</td>
<td>2.57</td>
<td>0.07</td>
<td>0.89</td>
</tr>
<tr>
<td>PC_18</td>
<td>2203</td>
<td>-4.89</td>
<td>2.60</td>
<td>0.08</td>
<td>0.90</td>
</tr>
<tr>
<td>TSFQ_18</td>
<td>2203</td>
<td>-3.94</td>
<td>0.93</td>
<td>0.05</td>
<td>0.94</td>
</tr>
<tr>
<td>Grade</td>
<td>2203</td>
<td>1.00</td>
<td>10.00</td>
<td>5.71</td>
<td>2.51</td>
</tr>
<tr>
<td>PSPC</td>
<td>2088</td>
<td>-3.50</td>
<td>3.00</td>
<td>0.22</td>
<td>1.00</td>
</tr>
<tr>
<td>PSPD</td>
<td>2153</td>
<td>-3.72</td>
<td>3.00</td>
<td>0.10</td>
<td>0.96</td>
</tr>
<tr>
<td>Tch_exp</td>
<td>1880</td>
<td>0.00</td>
<td>45.00</td>
<td>10.46</td>
<td>7.15</td>
</tr>
<tr>
<td><strong>CE Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE_19</td>
<td>2229</td>
<td>-6.36</td>
<td>2.34</td>
<td>0.14</td>
<td>0.95</td>
</tr>
<tr>
<td>CE_18</td>
<td>2229</td>
<td>-3.92</td>
<td>2.60</td>
<td>0.08</td>
<td>0.94</td>
</tr>
<tr>
<td>TSFQ_18</td>
<td>2229</td>
<td>-3.94</td>
<td>0.93</td>
<td>0.03</td>
<td>0.97</td>
</tr>
<tr>
<td>Grade</td>
<td>2229</td>
<td>0.75</td>
<td>10.00</td>
<td>5.89</td>
<td>2.47</td>
</tr>
<tr>
<td>PSPC</td>
<td>2117</td>
<td>-3.50</td>
<td>0.96</td>
<td>0.02</td>
<td>0.88</td>
</tr>
<tr>
<td>PSPD</td>
<td>2161</td>
<td>-3.72</td>
<td>1.04</td>
<td>-0.04</td>
<td>0.93</td>
</tr>
<tr>
<td>Tch_exp</td>
<td>1267</td>
<td>0.00</td>
<td>45.00</td>
<td>10.22</td>
<td>7.82</td>
</tr>
</tbody>
</table>

Table 1 shows the descriptive statistics of PC and CE samples. Since some of the schools did not follow the NEE suggestion to avoid using student survey in lower elementary classes, teachers that were rated by students below 4th grade are excluded from the following analysis (488 cases in the CE sample and 521 cases in the PC sample). To examine whether feedback quality may affect student-rated instructional quality and to further probe into the potential explanations, we recoded the TSFQ into a binary grouping variable CEIL (1 = the ceiling score; 0 = non-ceiling scores), and see whether the two groups have the same distribution and whether the ceiling group receives higher student rating in instructional quality. Non-parametric tests are used as it has no distribution assumptions and the results are shown in Table 2. To control for teaching practice scores in the previous year, we use the residulized teaching practice scores in the analysis. The distributions of teachers’ PC scores seem different between the ceiling and non-ceiling group and teachers’ median PC scores are higher when their principals receive ceiling scores in feedback quality. In contrast, the distributions of CE scores are not significantly
different between the two groups. Therefore, it is likely that explanation 4) may be true for PC scores but not CE scores, i.e., high quality principal feedback may help with teacher’s PC skills but not CE skills. It is equally possible that there is not enough statistical power to detect the group difference in teachers’ CE scores.

Figure 1. Ceiling Effect of TSFQ

![Histogram of TSFQ scores for PC and CE samples](image)

Although, technically, it is also possible that teachers who are concerned about their anonymity, lenient to their principals, and/or inattentive to the survey tend to have better PC skills but not better CE skills, there is a lack of direct theoretical support of this explanation. To rule out this possibility, and at the same time address the research question whether quality principal feedback may result in better teacher practices, we removed all the ceiling cases in both samples and conducted regression analysis that includes PSPC to approximate how teachers may be concerned about their anonymity or lenient to their principals and PSPD to approximate how
likely the teachers think they can benefit for their professional development students grades and teacher experience are also included.

**Table 2. Results of Non-parametric Tests**

<table>
<thead>
<tr>
<th>Ceiling TSFQ_18</th>
<th>PC_r a</th>
<th>CE_r a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>N</td>
<td>644</td>
<td>1038</td>
</tr>
<tr>
<td>Median</td>
<td>.129</td>
<td>.046</td>
</tr>
<tr>
<td>P</td>
<td>.024</td>
<td>.509</td>
</tr>
<tr>
<td>Mean Rank</td>
<td>872.657</td>
<td>822.170</td>
</tr>
<tr>
<td>Man-Whitney U</td>
<td>314171</td>
<td>348782</td>
</tr>
<tr>
<td>z-value</td>
<td>2.072</td>
<td>.822</td>
</tr>
<tr>
<td>p-value (Asymptotic)</td>
<td>.038</td>
<td>.403</td>
</tr>
<tr>
<td>p-value (Monte Carlo)b</td>
<td>.034</td>
<td>.413</td>
</tr>
</tbody>
</table>

Note: a. PC_r and CE_r are the CE and PC scores in 2019 controlling their prior year scores.

b. The Monte Carlo simulation is based on 10000 sampled tables.

**Figure 2** shows the distribution of TSFQ scores in both samples in the regression analysis with all ceiling cases removed. When all ceiling cases are removed, similar results can be seen in both samples that teachers received better principal feedback in the previous year tend to have better instruction quality in PC and CE in the following year, controlling their prior year PC and CE scores and covariates (see **Table 3**). Also, as both PSPC and PSPD are not statistically significant in the models, explanation 4) may play a more important role in teachers’ ceiling TSFQ scores than explanation 1) to 3). More importantly, both the non-parametric tests and the regression analysis give the same answer to our first research question, i.e., quality principal feedback predicts better teaching practices. The difference shown between CE and PC in the non-parametric tests may be due to sampling error or some inherent difference between
how CE and PC skills can benefit from quality principal feedback, which we do not know so far. Another interesting finding here is the negative effect of teachers’ experience on improving their instruction while controlling other variables. A possible explanation is that new teachers have more room to improve their instruction practices in general.

**Figure 2. Distribution of TSFQ Without Ceiling Cases**

To further identify the aspect of quality feedback that best predicts better teaching practices, we generated 5 latent factor scores of prior year TSFQ dropping each of the 5 items before the same regression models were conducted for both samples. It is shown that item 4, i.e. providing specific feedback regarding teachers’ strength, seems the most important contributor among the five items predicting both teachers’ PC and CE skill, as when dropping item 4 from the TSFQ factor score the models have the smallest $R^2$ in both PC sample, $b = .083$, $SE = .042$, $t_{(653)} = 1.978$, $p = .048$, $R^2 = .371$, and CE sample, $b = .073$, $SE = .053$, $t_{(501)} = 1.376$, $p = .169$, $R^2 = .375$. 
Table 3. Regression Results with Ceiling Cases Removed

<table>
<thead>
<tr>
<th>Variables</th>
<th>PC sample (R² = 0.374)</th>
<th>CE sample (R² = 0.375)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DV: PC_19</td>
<td>DV: CE_19</td>
</tr>
<tr>
<td>Intercept</td>
<td>-.196 .108 -1.819 .069</td>
<td>-.064 .135 -.474 .636</td>
</tr>
<tr>
<td>PC_18</td>
<td>.516 .031 16.876 .000</td>
<td>.575 .037 15.610 .000</td>
</tr>
<tr>
<td>TSFQ_18</td>
<td>.109 .043 2.513 .012</td>
<td>.119 .056 2.129 .034</td>
</tr>
<tr>
<td>Grade</td>
<td>.053 .015 3.646 .000</td>
<td>.045 .018 2.455 .014</td>
</tr>
<tr>
<td>PSPC</td>
<td>-.067 .044 -1.534 .126</td>
<td>-.037 .055 -.676 .499</td>
</tr>
<tr>
<td>PSPD</td>
<td>.024 .043 .553 .580</td>
<td>-.007 .053 -.129 .897</td>
</tr>
<tr>
<td>Tch_exp</td>
<td>-.008 .004 -2.350 .019</td>
<td>-.009 .004 -2.120 .035</td>
</tr>
</tbody>
</table>

Discussion

This study is one of the first to examine the effect of feedback quality on student-rated teacher effectiveness. The results support the important role of the principals’ feedback quality in teachers’ improvement of instruction practices. The study can also inform future evaluator training to improve teacher evaluation policy implementation. For example, an item-wise analysis suggests that providing specific positive feedback is of unique importance to encourage teacher to improve instruction practices to cognitively engage students and encourage student to solve problems and think critically.

We believe that the current study makes a significant contribution toward a meaningful impact through research, policy and practice.
The major limit of the study is that districts that choose to use both student survey and teacher survey maybe those who stress teacher professional development and students’ perception of instruction quality, which may bring in some selection bias. For future studies, the quality of the TSFQ measures may be improved by using an elongate scale, and the addition of quality control questions may help better adjust the feedback quality measure. It may be also helpful to include other teaching practice measures, such as teacher-student relation, classroom management and motivational strategies.
Reference


doi:https://doi.org/10.1016/j.cedpsych.2019.01.011


