STETSON UNIVERSITY

Voices of Reform: Educational Research to Inform and Reform

Volume 5 • Issue 1 • Article 2

Nina B. Hollis Institute for Educational Reform

December 2022

What Makes Classroom Observation Feedback Useful? The Perceptions of Secondary Math and English Teachers

Amanda Slaten Frasier
East Tennessee State University

Follow this and additional works at: http://www.voicesofreform.com

Recommended Citation

Frasier, A. S. (2022). What makes classroom observation feedback useful? The perceptions of secondary math and English teachers. *Voices of Reform, 5*(1), 40-58. Retrieved from https://www.voicesofreform.com/article/70421-what-makes-classroom-observation-feedback-useful-the-perceptions-of-secondary-math-and-english-teachers doi: 10.3623/5.00004

http://dx.doi.org/10.32623/5.00004

Revisions

Submission date: July 21st, 2022
1st Revisions: October 31st, 2022
Acceptance: November 5th, 2022
Publication date: December 28th, 2022

What Makes Classroom Observation Feedback Useful? The Perceptions of Secondary Math and English Teachers

Amanda Slaten Frasier[1]

[1]University School
East Tennessee State University, United States
frasiera@etsu.edu

Abstract

Teachers (n=14) at four high schools in North Carolina were interviewed about their perspectives of evaluation policy at two time points during the 2016-2017 school year. This study specifically examined statements teachers made about feedback from observations using de-identified interview transcripts. Teachers discussed feedback from formal observations along with other sources of informal observational feedback (e.g., coaches, peers). Overall, teachers described useful feedback as that which provided actionable recommendations informed by the observer's knowledge of three domains: the classroom context (as aided by the frequency and timing of observations), subject area, and pedagogy. Teachers also identified two aspects of formal evaluation that interfered with feedback: breadth of the observational standards and use as a growth measure.

Keywords

classroom observation, feedback, teacher perceptions, case study research

Acknowledgements

This work was funded in part by a grant from the East Tennessee State University Research Development Committee Small Grants Program. Initial data collection for this project was supported by the Michigan State University Graduate School and the Department of Educational Administration.

Introduction

Over the last several decades, the way teachers are evaluated has shifted from informal administrator led check-ins to formalized state regulated policies. This shift has been accompanied by attempts to standardize the measures used in evaluation, primarily through the use of

psychometric student growth models and standardized observation rubrics and protocols (Grissom & Youngs, 2016). At the time of this study, North Carolina's teacher evaluation tool included multiple measures including a standardized observational rubric and student growth measures and was intended to serve dually for formative and summative purposes (North Carolina Department of Public Instruction, 2015). The dual purpose of the evaluation policy parallels the ways in which evaluation can motivate changes in teaching practice. Firestone (2014) identified two theories of motivation that guide thinking about teacher evaluation: an economics-based theory focused on external rewards and summative evaluations along with a psychological theory focused on intrinsic rewards and teachers formatively improving practice through assessment, feedback, training, and professional development. The second psychological theory is what is examined in this case study, which identifies not only whether teachers perceive classroom observations as a useful source of feedback, but also determines the characteristics and conditions that teachers feel are necessary for feedback to be useful in influencing classroom practices. As such, I utilize de-identified interview data collected from high school teachers (n=14) from a district in western North Carolina to address the following research question: What criteria do teachers identify as necessary for feedback from classroom observation to be useful in changing classroom practices?

Theoretical Framework

Observational Feedback as a Lever for Improvement

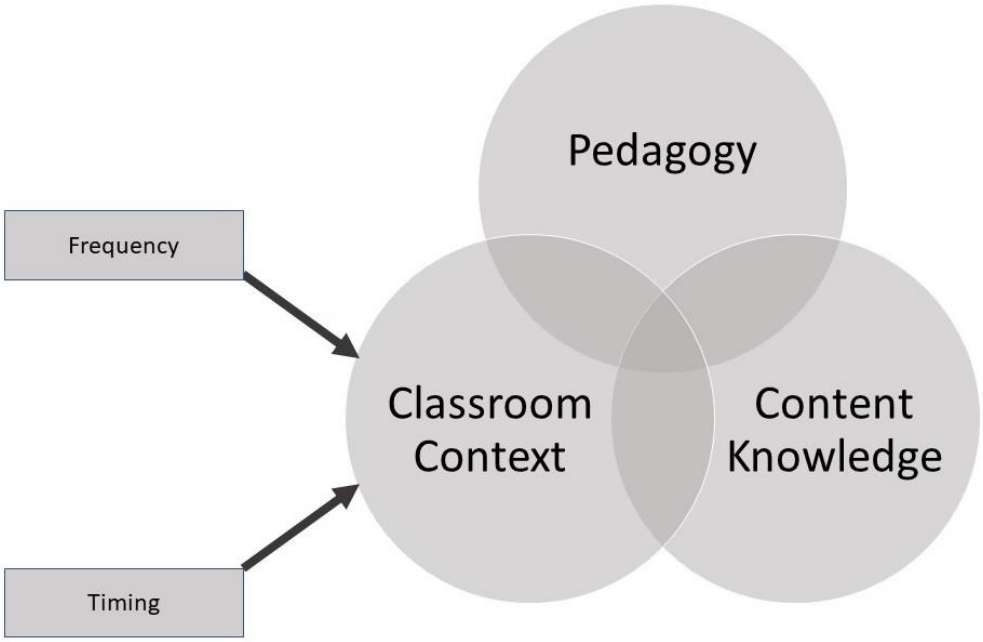
While many policymakers view teacher evaluation as means to summatively assess educators and regulate the workforce, evaluation systems have the potential to influence the work of teachers, particularly those who are new to the field (Kraft & Papay, 2014; Kraft et al., 2020). Moreover, giving frequent and specific feedback on classroom practice leads to improvement in teacher performance (Garet et al., 2017; Kraft et al., 2020; Steinberg & Sartain, 2015; Taylor & Tyler, 2012). While the most immediate feedback at a teacher's disposal comes directly from students, evaluation policies utilizing structured observation rubrics now serve as an additional potential source of formative feedback for teachers (Black & William, 2009; Grissom & Youngs, 2016). Unlike the summative assessment produced with student achievement data, observation feedback occurs practically in real time and often is accompanied by post-conference reflection between the observer and the observed. As such, observational feedback provides an opportunity for teachers to not only have their teaching apprised, but to reflect on their practice with another professional.

It is important to understand how observers may provide useful feedback, as without it, most teachers' performance plateaus by their third or fourth year on the job (Boyd et al., 2006). Moreover, teachers who have less experience may rely more on feedback and the guidelines of the evaluation when forming their instructional practices. Because they have had less exposure to other types of guidance, and perhaps less confidence in identifying good practice, newer teachers may use the components of evaluation as guiding principles in developing practice (Sun et al., 2016). A teacher's tenure status may also interact with how a teacher alters practice due to evaluations (Frasier, in press). Additionally, when teachers are provided scores and feedback from standardized protocols, they improve their practice, regardless of experience level (Allen et al., 2011; Taylor & Tyler, 2012). This large body of previous work on observational feedback demonstrates its potential impact as a lever for improving teacher performance.

In this study, teachers described useful feedback as that which provided actionable recommendations informed by the observer's knowledge of three domains: the classroom context (as aided by the frequency and timing of observations), subject area, and pedagogy (Figure 1). These domains constitute a theory of action for creating useful observational feedback and are described below.

Figure 1

Theory of Action for Analysis



Classroom Context

One of the most important aspects of effective evaluation is an observer's understanding of the classroom context, which requires frequent well-timed interactions between the observer and the observed. Other studies have demonstrated the importance of frequent, targeted feedback. For instance, Garet et al. (2017) employed an experimental design where the treatment group received more frequent targeted feedback and subsequently showed significant gains in performance on the *CLASS* observational rubric and in student performance in mathematics. Others have found that teacher reflection on practice can be enhanced by looping interactions continuously and tying feedback to specific evidence (Bambrick-Santoyo, 2012; Danielson & McGreal, 2000). Additionally, there is evidence that feedback has more utility when it comes quickly, offers opportunities for dialogue, and is accompanied by additional walk-thru observations (Donahue & Vogel, 2018). This knowledge of the context helps develop trust and creates authentic feedback that a teacher is more likely to act upon (Costa et al., 2014). So, observer feedback is more likely to be perceived as effective if the observer provides well-timed and frequent feedback, demonstrating an understanding of the classroom context that moves beyond a single snapshot in time.

Pedagogical Knowledge

Additionally, teachers need to feel confident that their administrator understands what good teaching looks like (Donahue & Vogel, 2018; Frasier, 2021; Kimball, 2022; Stiggins & Duke, 1988). For teachers, this can manifest as the observer providing specific, evidence-based feedback which requires a knowledge of effective instructional practices and management techniques (Derrington & Kirtk, 2017; Smith et al., 2020). Teachers may have a negative reaction to feedback focused on superficial aspects of teaching, such as whether or not the teacher has an objective listed on the board (Donahue & Vogel, 2018). So, it is important that the feedback provided demonstrates a consideration of pedagogical knowledge rather than creating a checklist of teacher actions.

Subject Area Knowledge

Finally, it has been established that when teachers teach across subject areas, such as in an elementary context, the subject they are teaching at time of observation can impact the ratings and feedback received (Bell et al., 2015a; Cohen et al., 2018; Hill et al., 2012). However, less work has been done to understand the mechanisms at play behind this. Bell et al. (2015b) identified that observers sometimes demonstrated uncertainty using both general and subject specific rubrics when scoring observations due to some aspects of the subject area. Additionally, Bell et al. (2015a) found differences in the rank ordering of teachers when general versus subject specific protocols were used and this rank ordering also differed based on the subject area taught by the teacher compared to the observer's subject area background. Furthermore, note-taking and feedback patterns differed depending on the subject matter background of the observer and whether there was alignment between an observer's background and the subject being taught (Bell et al., 2015a). The results of these studies suggest complexity in the ways that protocol, subject matter, and observer background intersect, and that further investigation is warranted.

Methods

This paper utilizes de-identified data collected for a larger mixed method study (Frasier, 2017). In the larger study, I used state level data (Teacher Effectiveness data and results from the Teaching Working Conditions Survey) to deliberately select four high schools in a single district that exhibited varying combinations of teacher evaluation ratings and teacher attitudes toward teacher evaluations. I then conducted a survey of English and Math teachers to measure teacher perceptions of evaluation using a Likert-type scale and used demographic data to select fourteen teachers who were interviewed in-person twice in the 2016-2017 school year and asked to describe evaluation policy implementation in their school. The fourteen teachers' interviews for the original study serve as the data source for this paper (Table 1).

Table 1

Interview Participants

Pseudonym	School	Subject Area Certifications	Grade Area Certifications	Years of Experience
Mr. Donaldson	Riley	English	9-12	17
Ms. McMinn	Riley	English	9-12	13
Mr. Forest	Phoenix	Math, Science	9-12	4
Ms. Street	Phoenix	English	9-12	15
Mr. Brown	Phoenix	English	9-12	2
Mr. Robbins	Charles	Math	9-12	10
Ms. Ranier	Charles	English	9-12	22
Mr. Silver	Charles	Math	9-12, 6-8	4
Mr. Eagle	Charles	English, SS	9-12, 6-8	4
Ms. Proffit	Central	Math, Science	9-12	22
Ms. Williams	Central	English	9-12	22
Mr. Augustus	Central	Math	9-12	7
Mr. Allen	Central	English	9-12, 6-8	5
Ms. Hoard	Central	English	9-12	10

The interviews had been transcribed and previously coded in qualitative software to answer other research questions around teacher perceptions of formal evaluation (Frasier, 2017; Frasier, 2021; Frasier, in press). While the initial study was targeted at gauging perceptions of formal evaluations, which included a standardized observation protocol, all the teachers interviewed were also prompted to discuss their experiences with informal observations. Teachers described a variety of informal observations, primarily from their administrators, peers, or curriculum coaches. So, the data also reveal teachers' perceptions of what makes for useful observation feedback from across both formal and informal observation types.

One code from prior analysis was *feedback* which included disaggregate child codes for both state testing and classroom observation. For the current analysis, the excerpts of the interviews that have been coded as *feedback* and *observation* were revisited to determine what teachers perceived as being the characteristics of useful feedback from classroom observation. In the first round of coding for this analysis, excerpts were viewed in context of the entire transcript and the source of observation feedback was coded (principal, curriculum coach, peer, etc.) as well as the sentiment the teacher was displaying towards the feedback. For the purposes of this study, *positive sentiment* meant the teacher expressed that the observation conditions described produced feedback that was useful or had utility to change classroom practice, *negative sentiment* meant the feedback was not useful, did not have utility, or the utility was criticized, and *neutral sentiment* as there being no definitive stance. In this first round of coding, memos were used to track emerging themes. These iteratively developed themes became the basis for the second round of coding and included *frequency*, *timing*, *subject area*, and *pedagogy*. The second round of codes were not mutually exclusive.

On examining the excerpts from both rounds of coding, I found that three specific observer domains had been identified. Each domain impacted teacher perceptions of whether they received effective feedback from the observer. Domain 1 was understanding of classroom context and the themes of *frequency* and *timing* merged into this domain. Domain 2 was understanding of *subject matter*. Domain 3 was understanding of *pedagogy*. Codes for domain 2 and 3 often overlapped in excerpts and are discussed together. Additionally, I coded excerpts that offered other explanations for why feedback had or lacked utility. The three domains and other explanations will be discussed subsequently.

Results

When teachers discussed observation feedback in the interviews, they usually were mentioning formal observations by their building administrators (principal or vice principal). The frequency was likely biased due to the focus of the initial project of formal evaluation policy, for which this data was initially collected. However, each teacher was also prompted to compare their experience with formal observations to informal observations, such as administrator walk-throughs, curriculum coach visits, or peer observers. At first glance, teachers seemed to overwhelmingly doubt the utility of feedback from their administrators with 67.88% of mentions having *negative sentiment* as opposed to other sources of observation, such as curriculum coach, which had a 75.78% *positive sentiment*. However, most of the administrator observations were tied to teacher descriptions of formal standardized observations. When positive or negative sentiment was

examined in relationship to whether the teacher was discussing an informal or formal observation (Table 2), teachers overwhelmingly viewed informal observational feedback with *positive sentiment* (71.42%), regardless of the observer type, as opposed to *negative sentiment* (16.67%). So, sentiment appeared to be more tied to whether the feedback was from a formal or informal observation as opposed to the role of the observer.

Table 2

Sentiment by Observation Type

Observer Type	Total*	f Positive	% Positive	f Negative	% Negative
Formal	140	41	29.29%	93	66.43%
Administrator	128	31	24.21%	93	72.66%
Curriculum Coach	4	0	0	3	75%
Peer	2	0	0	0	0
Other	3	0	0	1	33.33%
Informal	42	30	71.42%	7	16.67%
Administrator	9	9	100%	0	0
Curriculum Coach	27	25	92.59%	2	7.41%
Peer	8	6	75%	1	12.5%
Other	6	6	100%	0	0

*Includes neutral

Overall, teachers reported that the frequency and timing of observations, which contributed to a better understanding of classroom context, as well as an observer's understanding of teaching and subject area knowledge were factors in creating beneficial observation experiences. The three domains of classroom context, subject area, and pedagogy will subsequently be discussed. As evidenced by the distribution of positive versus negative sentiment frequencies, informal evaluations tended to better operate under the conditions teachers specified as yielding better feedback as opposed to formal evaluation. So, the additional obstacles to effective feedback in formal observation will also be discussed.

Frequency and Timing Contribute to Understanding Classroom Content

The first observer domain that contributed to effective feedback was the understanding of classroom context. All the teachers interviewed agreed with the sentiment that feedback was most effective when it was specific and situated in the context of their classroom. Teachers spoke positively of situations where frequent, informal observations occurred whether the observer was a peer, curriculum coach, or administrator. Not every teacher utilized a curriculum coach for informal observations, but with one exception (Mr. Augustus, who will be discussed later), those who did spoke positively of the experience. Moreover, teachers who spoke positively of coaching expressed wanting more frequent observations. Ms. Street, who herself was a former curriculum coach, described the need for even more frequent and intentional coaching, "We have coaches that are spread so thin, it's like they're scattering fairy dust everywhere." Similarly, Mr. Silver, who found utility in observation feedback from both his curriculum coach and peers, summed up the importance of frequent engagement with other adults providing feedback in the classroom, "[T]eachers sometimes feel like they're on an island surrounded by teenagers or children... You feel so alone in those moments… So, the more you can be a community and be friends and be colleagues, [then the more observation] is really, really helpful."

Teachers also spoke of how having an observer visit frequently and informally improved their relationship with formal observations. Mr. Silver expanded:

> I'm comfortable having people in my room. And I think having observations at a decently consistent pace is really [beneficial]... When someone comes into my classroom, I feel like I am the same teacher as when they're not in here. My students have even said, "Oh wow! You don't change." And I'm like, "That's a compliment." I don't need to change. I want them to see what I'm doing because if I'm doing something wrong, I want them to tell me. And if I'm doing something right, I want them to see it.

Similarly, Ms. Ranier stressed the importance of administrator presence in the classroom aside from formal observations:

> I really enjoy working at a school where the administrators are present and they're often in your class, because then when I sit down with them and they're talking about my teaching, they can say, "Oh, but you did this on the other day when I was just walking through."

> And also, then their presence, in and of itself wouldn't be so disarming when they come here the two or three times they come to do a [formal] observation.

The quotes from Mr. Silver and Ms. Ranier highlight how important it is to teachers to have observers in their room who are there for the express purpose of formatively improving teaching rather than formally evaluating it.

Furthermore, teachers often stressed that administrators who do not frequently visit classrooms may not get an accurate representation of an individual's teaching. Ms. Hoard summed how infrequent observation could not gauge a sufficient sample of what a teacher does in the classroom, "Compared to the 180 days [in the school year]. Statistically, it's not even... That's not a valid sample." Similarly, Ms. Proffitt compared the first year of her career at a middle school where the principal frequently informally observed classrooms opposed to her experience at a high school with formal observations only:

> So, when I received my first evaluations, they were good. They weren't perfect… but [the principal] was very good about giving me constructive feedback. She kind of took me under her wing, which was very helpful. [T]hrough my subsequent evaluations as the year progressed, they got better and better. I ended up having a really successful first year at that school, even though as any teacher knows, first years are very stressful and trying. In the high school… it was just the three evaluations where they just popped in. I never received any bad evaluations, so I can't speak from that end of it. It was always the nitpicky little details that they would ask about and you would be like, "But wait a minute, that doesn't really affect what I do in the classroom at all, that doesn't help student learning." 'Cause that's always what I've been striving to do. What I do in my classroom is to promote student learning.

Thus, observers who appear in classrooms only to conduct formal observation may miss opportunities to engage with teacher practice and offer constructive feedback.

The timing of observations also mattered to teachers. Six teachers gave examples of observations that occurred all within a period of a few days (not allowing teachers time to institute changes based on feedback), or occurred all within the same class, or at a bad time of the year. For instance, Ms. Williams shared, "We only had to do one evaluation, I was literally evaluated the last day of school, the last class and I was like, 'Really? Are you serious?' And at first, it kind of makes your back go up and then you think, 'Well now, wait a minute, if he's showing up when it doesn't matter anymore then that means it doesn't matter.'" Another teacher, Mr. Donaldson described how a new administrator had formally observed him 3 times within two weeks. While he expressed appreciation of the feedback, he stressed that it was difficult to act on that feedback in such a short time frame. So, while a frequent presence in the classroom was necessary for an observer to be

able to best understand the classroom context and offer feedback, teachers also wanted that presence to occur at times where they could actually implement changes based on observer input.

Understanding of Subject Area and Pedagogy

Aside from expressing a need for observers to be frequently in classrooms to understand the context in which teachers are working, 10 of the 14 teachers also expressed a desire to be observed by individuals who understood their subject area and/or good pedagogy in their subject area. Thus, understanding of subject area and pedagogy represent domains 2 and 3, respectively.

Most teachers saw benefit to observations conducted by their same-subject peers and curriculum coaches, underscoring that having an observer who understands the subject matter of the observed classroom influences how teachers perceive the feedback received. The math teachers at two of the four schools in this study had an observing administrator who had a math background. Ms. Proffit described how her principal's math background allowed her to have "a true conversation about what's going on" in her math classroom. She described how her principal's knowledge of math content allowed for conversations that are more specific to how her pedagogy relates to the classroom, "When we meet… I say, 'We were covering completing the square for quadratics before this,' and he knows exactly what I'm talking about. So, the feedback revolves around… the lessons and how they are… written, presented, and modes of instruction."

Another math teacher, Mr. Augustus, shared similar sentiments about his principal who also had a math background and who had in the past offered him specific math related advice. Furthermore, he questioned how an administrator could apply a formal observation protocol fairly across all contexts:

> [H]ow do you compare teachers that are teaching different subjects, teaching at different schools? How do you make those comparisons level and equal? And then you could even go so far down to the classroom level of two different teachers teaching the same course, but one is an inclusion class… [o]r one class may be really large, and one class may be really small? Having those classes at different parts of the day... And you wouldn't believe the climate change of the exact same class because something happened, or something didn't happen.

Interestingly, Mr. Augustus was the only teacher who questioned the benefit of feedback from his curriculum coach. He stressed there was a clear difference between teaching theory and practice, and that his coach only presented him with theoretical approaches that would not work in his real-life classroom context. He stated, "Most of my hang-ups are students that are unwilling to work, and so, they'll give me ideas for that" and he describes that those ideas are things he already tried or that do not work, "There's certain students that are just not going to learn in this environment." Mr. Augustus viewed the coach as having solid knowledge of both subject area and pedagogical theory, but felt the coach lacked the knowledge of his individual classroom issues. In the case of Mr. Augustus, he felt the feedback from his principal, which was focused more on improving what

he perceived to be his "hang-ups" or challenges with managing behavior in his specific classroom, was more helpful.

In contrast, Mr. Forest compared experiences with his curriculum coach, who observed all of his classes at least weekly, to his principal who was also frequently informally in the classroom, but did not have a math background:

> [The coach] made a lot of the curriculum materials that I use, and so, in talking to her afterwards… she gives me feedback… on the things that she feels is the most relevant for my teaching and my class and my kids. [T]hat's the feedback that I need. So, my principal comes in and he checks boxes, He does not have a math background… he doesn't know English either, he doesn't know social studies. And so, that would be okay, if he would see that [he] needs to be able to give people feedback. So [he] needs to go learn, needs to go to the math coach and see what is she looking for. He needs to figure out, what does a good math teacher look like? And that's what truly I think a principal's job is with this evaluation stuff.

Mr. Forest goes on to describe how he wrote a grant with other math teachers in his district to try to promote what he termed "public teaching" in an attempt to help improve awareness across his district's administrators on what good math teaching should look like. Mr. Forest further stressed the importance of providing actionable feedback, "[W]e actually wrote the grant to include principals, so that principals could come start learning the same things us as teachers have to learn, which is, what does a good math teacher do?"

Aside from subject areas, teachers expressed that they needed to feel confident that their observer was an effective appraiser of pedagogy. Across two schools, there were examples of principals asking teachers what to put on their formal observation rubrics in post conferencing. Teachers described how this practice cast doubt on the validity of their formal teacher evaluation results, as Mr. Forest posed, "Why would it be up to me to decide what comment to put, if you're an effective evaluator?" At one school, there was an extreme case which was described by all of the teachers at the school. Ms. Williams shared:

> The person that does our evaluations knows little to nothing about what we do and isn't particularly interested. We literally write our own evaluations. [H]e will make you look like… an idiot. So, we finally... just say, "Give me the computer," and we start typing, because if we don't do that, we'll be misrepresented, simply because he's incompetent when it comes to [feedback].

In these cases, teachers questioned the pedagogical expertise of their observer.

When administrators lack content or pedagogical knowledge, their observation feedback may be limited to assessments of classroom management. Six teachers mentioned feedback focused only on management. For instance, Mr. Eagle had experienced an observer with sound English

pedagogical content knowledge, and several who had not, "People don't understand the content areas that are being taught, or the lessons that are being taught, or the units specifically that we're teaching that day. And the best they can really do is be able to get an idea of how the classroom is run." So, teachers perceive that observer knowledge can limit the scope of feedback that can be offered.

Other Obstacles to Producing Actionable Feedback

The production of actionable feedback was important for teachers in this study. Actionable feedback is often directly linked to the concerns about understanding classroom context, subject matter, and/or pedagogy mentioned previously. However, teachers also specifically took issue with two characteristics of formal observation instruments used for their state teacher evaluations: overly broad standards and the tool's use as a growth model.

Teachers felt that the broad rubric forced observing principals to focus on too many topics or topics irrelevant to the teacher's instruction. Specifically, teachers mentioned standards in the rubric that focused on factors like *technology*, *global awareness*, or things that cannot be observed in a classroom, such as *parent relationships*. For instance, three of the math teachers discussed the difficulty of meeting the *global awareness* standard on the formal rubric in a math class. Sometimes, teachers were told that their observer was only focusing on specific actions of teaching during an observation, and after being told to make sure he was meeting *global awareness* during a pre-conference, Mr. Silver stated how he had to figure out a way to inauthentically work it into a lesson, "I kind of felt like I was plugging, like an advertisement." Two math teachers from different schools also acknowledged that their administrators often gave high marks on technology, because of the use of calculators, which the teachers did not consider to be technology use as intended by the rubric. Six teachers made comments that it would be impossible to obtain high marks universally on the rubric, because the breadth involved so much devotion outside of the classroom and school hours. Ms. Ranier stated anyone who could achieve the highest rating would need to take place in a "parade of martyrs" due to the sacrifices one would need to make in their personal lives to achieve mastery on some of the standards.

Additionally, six teachers across all four schools mentioned the use of the formal observation tool as a growth instrument and gave examples of administrators telling them they had been marked low on certain standards so they would have an opportunity to show growth later by receiving a higher score. Because these examples came from teachers who had different observing administrators, the common story suggests that principals were trained to use the tool in this way, at least at the district level. "Being told that I was marked low on something just so I could be higher later does nothing to help me know how to improve as a teacher," Mr. Allen surmised. This common interpretation of a growth model prevented useful feedback for teachers, and those who mentioned the practice described it using words like "inauthentic" and "forced."

Discussion

Implications for Theory

Teachers found that their experiences with feedback were more useful when the observer understood the classroom context, something which involved taking the time to frequently visit and follow-up on such visits, a finding mirrored in other studies (Bambrick-Santoyo, 2012; Danielson & McGreal, 2000; Donahue & Vogel, 2018). Teachers also stressed that it was important to view the observer as understanding pedagogy. Many teachers also expressed the desire to have observations conducted by someone who understood their subject area, and curriculum coaches were often cited as a useful source of feedback. Additionally, the results of this study suggest that formalizing feedback may interfere with the utility of feedback, a finding that replicates those found elsewhere (e.g., Donahue & Vogel, 2018). The critiques teachers offered that did not fit the theoretical framework, the breadth of criteria and the use of the tool as a growth measure, highlight the difficulty of using observation for both summative and formative purposes, an issue that others have also examined (e.g., Firestone et al., 2014). Overall, this paper supports the notion that observational feedback and its use by teachers is a complicated process impacted by not only the actors involved and their expertise and backgrounds, but also the way in which observation is framed, particularly if it is associated with formal structures and policy requirements.

Implications for Policy and Practice

This study is a case study with a small sample size, so the results should be interpreted cautiously. However, this study raises important questions about why formal observation protocols may fail to produce feedback that teachers find to be useful. Policymakers and those who care about using classroom observation as a lever for improving teacher practice should consider how formal observation practices can be adjusted. Currently, formal observations usually have an observer who makes infrequent visits to classrooms to evaluate a broad number of categories in a short time frame. Essentially, these few observations require a deep dive on a broad range of topics, all of which may not be particularly applicable to the teacher in the moment of observation. Additionally, some teachers report the tool being used as growth instrument, which had caused some administrators to resort to gaming scores by using lower ratings initially to show eventual growth. Teachers who reported this felt this approach to growth impacted the validity of the feedback they were receiving.

While three domains have been identified as important in this study, this result should not be interpreted that every observer needs to be proficient in all three domains. Instead, having a combination of feedback from people who understand each domain may best help a teacher craft their practice. For instance, the Garet et al., 2017 study employed observers who would not normally observe the teacher for the treatment group and showed statistically significant gains in some measures. Moreover, teachers in the treatment group reported being more interested in utilizing feedback to improve their practice as well as having a greater perception of their own effectiveness as a teacher (Garet et al., 2017). Having a variety of observers who have a clear understanding of at least one of the domains identified here could help teachers gain better

feedback. In order for a teacher to receive feedback from all three domains, observations from an administrator who understands the classroom context and good pedagogy could be combined with observations from a coach who has expertise in the subject area.

Opportunities for Future Studies

The policy and practice trade-offs of adjusting current policy to allow for more frequent and personalized observations should be examined in future work. Additionally, if an observer does not have a strong understanding of classroom context, subject area, and/or pedagogy, then the feedback given by the observer may not have much utility for the teacher. More work is needed to better understand which observer characteristics are linked to producing effective feedback. The relationship between observer characteristics and teacher characteristics (such as shared subject area expertise) is also warranted. It is likely that having feedback from multiple sources adept across multiple domains may be more useful than a handful of single observer visits a year. Work on such as model, including one where instructional or subject-area expert coaches can play a role in providing observational feedback, would also be beneficial in informing future policy.

Conclusion

Overall, teachers in this study overwhelmingly reported that observational feedback from informal sources was more useful to changing teaching practice than that from formal sources. Teachers reported that the breadth of the formal rubric and its use as a growth tool interfered with the utility of the feedback received. Teachers also reported that the most useful observational feedback came when they were given actionable suggestions from individuals who had a robust understanding of their classroom context (due in large part to the frequency and timing of classroom visits), subject area, and/or pedagogy. When teachers had administrators who fit these conditions, they expressed that they could find utility in the feedback given by their administrators, even using formal protocols, though this hypothesis should be tested further in other studies.

References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science, 333*(6045), 1034–1037.

Bambrick-Santoyo, P. (2012). Beyond the scoreboard. *Educational Leadership, 70*(3), 26-30.

Bell, C., Drake, C., Wilson, M., Frasier, A., Qi, Y., McCaffrey, D., Lockwood, J. R., & Kim, J. (2015a). *Subject specific and general observation protocols as tools for the evaluation and improvement of teaching* [Paper presentation]. American Education Research Association Meeting, Chicago, IL, United States.

Bell, C. A., Qi, Y., Croft, A. J., Leusner, D., McCaffrey, D. F., Gitomer, D. H., & Pianta, R. C. (2015b). Improving observational score quality: Challenges in observer thinking. In T. J. Kang, K. A. Kerr & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the measures of effective teaching project* (pp. 50-97). Jossey-Bass. https://doi.org/10.1002/9781119210856.ch3

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*(1), 5-31.

Boyd, D. J., Grossman, P., Lankford, H., Loeb, S., Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1*(2), 176–216.

Cohen, J., Ruzek, E., & Sandilos, L. (2018). Does teaching quality cross subjects? Exploring consistency in elementary teacher practice across subjects. *AERA Open, 4*(3).

Costa, A. L., Garmston, R. J., & Zimmerman, D. P. (2014). *Cognitive capital: Investing in teacher quality*. Teachers College Press.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Educational Testing Services.

Derrington, M. L., & Kirk, J. (2017). Linking job-embedded professional development and mandated teacher evaluation: Teacher as learner. *Professional Development in Education*, *43*(4), 630-644.

Donahue, E., & Vogel, L. R. (2018). Teacher perceptions of the impact of an evaluation system on classroom instructional practices. *Journal of School Leadership*, *28*(1), 31-55.

Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher, 43*(2), 100-107.

Frasier, A. S. (2017). *Examining teacher perceptions of the relationship between evaluation policy and teacher practice in a North Carolina school system* [Doctoral dissertation, Michigan State University]. ProQuest One Academic; Social Science Premium Collection.

Frasier, A. S. (2021). Principal-led state evaluation policy: Teacher perceptions at four high schools. *The Qualitative Report, 26*(9), 2769-2796.

Frasier, A. S. (2023). Do high school teachers alter classroom practice due to evaluation? Evidence from North Carolina's career status and teacher evaluation policies. *Journal of Education Human Resources.* https://doi.org.10.3138/jegr-2022-0007.

Garet, M. S., Wayne, A. J., Brown, S., Rickles, J., Song, M., Manzeske, D., & Ali, M. (2017). *The impact of providing performance feedback to teachers and principals*. American Institutes for Research. https://www.air.org/resource/report/impact-providing-performance-feedback-teachers-and-principals

Grissom, J. A., & Youngs, P. (2016). *Improving teacher evaluation systems: Making the most of multiple measures.* Teachers College Press.

Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., Humez, A., Litke, E., & Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment, 17*(2–3), 88–106.

Kimball, S. M. (2002) Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal of Personnel Evaluation in Education, 16*(4), 241-268.

Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis 36*(4), 476-500.

Kraft, M. A., Papay, J. P., & Chi, O. L. (2020). Teacher skill development: Evidence from performance ratings by principals. *Journal of Policy Analysis and Management*, *39*(2), 315-347.

North Carolina Department of Public Instruction. (2015). NC Educator Evaluation System. http://www.dpi.state.nc.us/effectiveness-model/ncees/

Smith, E. C., Starratt, G. K., McCrink, C. L., & Whitford, H. (2020). Teacher evaluation feedback and instructional practice self-efficacy in secondary school teachers. *Educational Administration Quarterly*, *56*(4), 671-701.

Steinberg, M. P., & Sartain, L. (2015). Does teacher evaluation improve school performance? Experimental evidence from Chicago's Excellence in Teaching project. *Education Finance and Policy*, *10*(4), 535-572.

Stiggins, R. J., & Duke, D. L. (Eds.). (1988). *The case for commitment to teacher growth: Research on teacher evaluation*. SUNY Press.

Sun, M., Mutcheson, R. B., & Kim, J. (2016). Teachers' use of evaluation for instructional improvement and school supports for such use. In J. A. Grissom & P. Youngs (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 102-115). New York, NY: Teachers College Press.

Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, *102*(7), 3628-51.