

6-2021

A Regional Case Study: Effective Administrator Feedback to Inform Teacher Performance

Richard M. Vandermolen
Grand Valley State University, vanderri@gvsu.edu

Catherine L. Meyer-Looze
Grand Valley State University, meyerlca@gvsu.edu

Follow this and additional works at: https://scholarworks.gvsu.edu/coe_articles



Part of the [Education Commons](#)

ScholarWorks Citation

Vandermolen, Richard M. and Meyer-Looze, Catherine L., "A Regional Case Study: Effective Administrator Feedback to Inform Teacher Performance" (2021). *Peer Reviewed Articles*. 19.
https://scholarworks.gvsu.edu/coe_articles/19

This Article is brought to you for free and open access by the Education at ScholarWorks@GVSU. It has been accepted for inclusion in Peer Reviewed Articles by an authorized administrator of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

A Regional Case Study: Effective Administrator Feedback to Inform Teacher Performance

By Richard M. Vandermolten and Catherine L. Meyer-Looze

In 2013, the Michigan Department of Education required all public school districts to adopt a state-approved evaluation system. The Department's rationale for this change was to increase teacher accountability and thus to improve student achievement. Because school-level administrators are responsible for implementing district policy and acting as the primary evaluators of teacher performance, the authors conducted a study to identify the type of administrator feedback principals give in observations and evaluations. In this 3-year study, evaluation data were collected from random districts within a rural Michigan Intermediate School District and analyzed for average total score and manner of feedback. Three primary results emerged. First, the rating count and average score per element rated varied widely; second, evaluators frequently provided comments with each rating; and finally, these evaluators' comments were not actionable for teacher performance improvement. The results of this work suggest a need for clarity around the definition and purpose of feedback in teacher performance evaluations.

Several studies have demonstrated that teacher quality significantly affects student achievement more than any other improvement strategy (Center for Public Education, 2005; Strong et al., 2001). In the last decade, a movement has galvanized around improving teacher quality through rigorous evaluation systems.

Wiliam (2016) reviewed several studies on the impact of educational factors and found that teacher quality affects student learning the most. Students with the most effective teachers learn in 6 months what students with average teachers learn in 1 year or students with poor teachers learn in 2 years. Students show a 90% learning return by 3 years with an effective teacher, but a student who has spent 3 years with an ineffective teacher may show only a 37% rate of learning return.

However, the qualities that make one teacher effective and another ineffective remain elusive. In fact, Wiliam stated, "We know that teachers make a difference, but we don't know what makes the difference in teachers" (2016, p. 38). His research supported the Michigan Department of Education (MDE) data that indicated teachers remain at the same effectiveness rating year after year (MDE, 2019). With Michigan's current shortage of teachers, it is critical that all teachers, including those identified as ineffective, are provided with the tools to deliver effective instruction that will increase student achievement.

With the goal of increasing student achievement and becoming a top-10 performing state in 10 years, MDE identified a key statewide strategy to improve educator effectiveness: increase the rigor of educator evaluations. This strategy was enacted into law when the Michigan Legislature passed Public Act 173 of 2015, which required all public school districts in the state to adopt "rigorous, transparent, and fair evaluations for teachers and administrators." The law further requires that (a) teachers be evaluated at least annually and be provided with "constructive feedback," and (b) teacher evaluation ratings be based in part on student growth (Public Act 173 of 2015). This bold legislative move was based on the premise that

stricter accountability systems and attention to the technical core of teaching and learning would yield increased results for students (Elmore, 2000).

The 2015 legislation further required the MDE to maintain a list of state-approved evaluation tools. Following an extensive vetting process, the Department formally adopted four evaluation tools, all of which have similar components: a growth plan, a specific research-based framework, a technology observation tool to provide feedback, a professional learning sequence to better understand the evaluation tool, a process to include multiple types of observations (informal, formal, or walkthrough), and a means of assigning a definitive score to the final evaluation (MDE, 2020).

Purpose Statement

The rationale for raising the bar of Michigan’s evaluation system was to improve student achievement through increased teacher accountability and effectiveness. Yet state statistics continue to show a disconnect between the high number of effective teachers and the lack of a significant increase in student achievement (MDE, 2020). The purpose of this research was to identify the types of feedback principals give to teachers in the state-mandated evaluations.

Research Question

This study addressed the following research questions: What type of feedback comments do building-level administrators provide in teacher evaluations? How does the feedback align with what research says is the type most likely to improve teacher performance? We, the authors, hypothesized that the majority of feedback comments would not be constructive or actionable for teacher improvement.

Subject Population

Data were collected from five northern Michigan counties served by the Traverse Bay Area Intermediate School District (TBAISD), which is an educational support organization connecting the MDE to 16 traditional public school systems, 6 charter schools, and 14 nonpublic school systems. All of the schools are classified as rural. The TBAISD provides services to its constituents such as special education, professional development, and assessment and data management.

All districts within the TBAISD region adopted one of the state-approved evaluation tools: the Marzano’s (Focused) Evaluation Model (Marzano, 2013). The TBAISD professional staff began rolling out implementation practices to its local districts in 2013. The TBAISD houses all the evaluation data points for these districts in Learning Sciences International’s platform, iObservation (Marzano & Toth, 2013). The growth plans for each teacher and leader, as well as their observation touch points for each of the framework’s five domains, are collected within iObservation.

Literature Review

Conceptual Framework

Parsons and Thompson (2006) created a conceptual framework to depict a Model of Organizational Levels in which the technical cores of teaching and learning are at the center of educational organizations. Parsons and Thompson’s premise is that this is where student success is most affected; however, factors outside the educator’s control—such as limited resources, socioeconomic status, demographics, school culture, and a global pandemic—can interfere with teaching and learning. Historically, school administrators have been taught to manage their buildings and



Dr. Rick Vanderمولen is Assistant Professor and Graduate Program Director for the educational leadership master's degree program at Grand Valley State University. An administrator of all levels in K-12 public education during his 25+-year career, Vanderمولen has always had a passion to understand and know how people learn. His recent transition to higher education has afforded him more time to make this an area of emphasis in his research and teaching. Vanderمولen received his doctoral degree in educational leadership from Eastern Michigan University; attended Harvard's Graduate School of Education to study the topic of Instructional Rounds; and, most recently, has become a certified trainer in the Process Communication Model. He is also engaged in his community as a co-facilitator and member of Leadership Grand Traverse Steering Committee for Traverse Connect in Traverse City, MI.

vanderri@gvsu.edu

keep these organizational and cultural factors at bay (Elmore, 2000). The goal of legislative accountability measures was to change the historical role by moving administrators who evaluate teaching back into the classroom, believing it would improve both teacher effectiveness and student achievement. However, student achievement results have not provided clear evidence of significant improvement.

What is Feedback?

Feedback helps teachers improve their performance and, in turn, improve student performance. Holland (2014) defined feedback as providing information about past behavior and presenting it in the present to potentially influence future behavior. Providing effective feedback should move learners forward. Wiggins (2012) posited that feedback is only effective if a goal is in view alongside evidence of the goal's result. The feedback given (and received) must be specific and aligned with the goal. For example, when an evaluator provides feedback around grouping strategies to improve student collaboration and, thus, improve sustainable learning around a concept through dialogue and discussion, the teacher is able to interpret not only the strategy but also the goal behind the strategy given. When this criterion of feedback is given, feedback has been proven to produce positive results (Wiggins, 2012). Getting feedback right is complex and critical, requiring a specific skill set as well as an understanding of a growth mindset.

For the purpose of this study, the authors defined feedback from the theoretical model for performance feedback in teacher-evaluation systems created by Cherasaro et al. (2016). In their correlational study, Cherasaro et al. studied teacher responses to four different characteristics: the usefulness of the feedback, the accuracy of the feedback, the evaluator's credibility, and the teacher's access to resources. These four characteristics are themes that exist in a number of the definitional studies already noted (Cherasaro et al., 2016).

Teacher Evaluation

The TBAISD chose to implement Marzano's Teacher Evaluation Model and to support its use throughout the entire region of school districts. The model includes four domains: (a) Classroom Strategies and Behaviors, (b) Planning and Preparing, (c) Reflecting on Teaching and Collegiality, and (d) Professionalism (Marzano, 2014). Within each domain, instructional strategies, as well as planning and reflective practices, are broken down into design questions and 60 different elements for which to observe and evaluate.

When an administrator observes a teacher, he or she captures that observation in iObservation, Learning Sciences International's software that "collects, manages and reports longitudinal data from classroom walkthroughs, teacher observations and teacher evaluations" (Learning Sciences International, n.d.). The observation must first be logged as an Informal, Formal, or Walkthrough Observation. The difference between an Informal and Walkthrough Observation is one of time. A Walkthrough is a short snippet of teaching and learning that does not last longer than 10 minutes, and an Informal Observation exceeds 10 minutes.

While observing instruction, the administrator makes a judgement call as to which elements are appropriate to score on the evaluation tool based on what is observed. The five choices of scores include *Not Using*, *Beginning*, *Developing*, *Applying*, and *Innovating*. The administrator determines the score based on a rubric that he or she must interpret.

Experimental control studies were conducted (Cherasaro et al., 2016) by looking at teacher practice and identifying the effectiveness of specific classroom strategies. These studies were designed to establish a direct causal link between elements of the model and student achievement. As of 2013, 300 experimental control studies had been conducted, involving more than 14 districts across the United States, including 38 schools, 300 teachers, and more than 14,000 students (Cherasaro et al., 2016). On average, when teachers used the classroom strategies and behaviors in the Marzano Evaluation Model, student achievement increased by 16 percentile points, with greater gains realized if specific strategies were used in specific ways (Cherasaro et al., 2016).

Wiliam (2016) found that teacher-evaluation frameworks must be comprehensive in order to have an impact on teaching effectiveness because teaching is dynamic and complex. Therefore, an observation tool and teaching framework should include all aspects of the teaching dynamic that will influence student learning. However, when teachers are under undue pressure to improve their effectiveness, they might become so focused on their score that they are unable to pay attention to student learning. This focus on the evaluation process itself can result in no improvement of student achievement. Some of these high-stakes practices are causing more stress, anxiety, fear, and distrust among the very educators who should be working and collaborating together. Thus, Wiliam (2016) believed that leaders should focus on *improvement* frameworks, as opposed to *evaluation* frameworks, to encourage feedback on the teacher's next level of learning.

**“
When teachers are under undue
pressure to improve their
effectiveness, they might become
so focused on their score that
they are unable to pay attention to
student learning.
”**

Teacher Use of Feedback

The Regional Educational Laboratory Central and Education Effectiveness Research Alliance (the Alliance) developed the Examining Evaluator Feedback Survey to help leaders better understand how teachers view feedback from their evaluator (Cherasaro et al., 2016). Survey questions included reflections on the feedback characteristics of usefulness, accuracy, credibility, and access to resources and responsiveness, as well as the importance of those characteristics and the belief around instructional improvement. “Usefulness of the feedback” speaks to the specificity of the feedback as well as to its timeliness and frequency. “Accuracy of the feedback” measures how the receiver believes in the feedback. “Credibility” refers to the receiver's perception of the competence of the evaluator, which may determine how the feedback is taken. Last, because the goal of effective feedback is to improve instruction, the survey asks about the “availability of resources” as it relates to improved teaching and learning (Cherasaro et al., 2016).

The Alliance survey revealed that what matters most when it comes to putting feedback into action are actionable steps, the teachers' perceptions of the feedback, and access to effective practices to improve instruction, such as collaborative models (Cherasaro et al., 2016). A correlational study found that a teacher's perception of feedback was positively correlated to an evaluator's giving specific



Dr. Cathy Meyer-Looze

is Associate Professor and Department Chair for Educational Leadership and Counseling at Grand Valley State University. A member of Phi Chapter in Michigan State Organization, she has been an administrator of all levels in K-12 public education, including serving as a professional development director. Meyer-Looze received her doctoral degree and master's degree in educational leadership from Eastern Michigan University and Central Michigan University, respectively; her bachelor's degree is in elementary education from Saginaw Valley State University. She is a graduate of both Leadership Grand Traverse and Leadership Little Traverse in Petoskey, MI.

meyerlca@gvsu.edu

recommendations for next levels of work along with providing the feedback in a timely manner (Cherasaro et al., 2016). In addition, a teacher utilized the feedback if he or she found the evaluator to be credible or knowledgeable about teaching and learning, regardless of content area (Cherasaro et al., 2016). Last, it made a difference if an evaluator not only gave feedback that was relatable but also provided resources for the next levels of work (Cherasaro et al., 2016).

Current Leadership Practices

Current leadership practices have not been ideal as they relate to improving teacher effectiveness. Although the intent of the 2015 legislation was to improve student achievement through improving teacher practice, a strong majority of teachers have received an “effective” or “highly effective” rating while the student achievement scores remained static.

Although it has been established that a teacher is the most important factor in student achievement, the second most significant element is building leadership (Grissom et al., 2015). To have a significant increase in student achievement, a principal must be an instructional leader who provides quality, effective feedback to teachers (Hattie, 2009). But providing the building leader an evaluative tool steeped in effective instructional practice does not mean that the leader knows how to use the tool or give effective feedback. The evaluator not only needs to understand the tool; he or she needs to understand the importance and qualities of effective feedback.

Methodology

Participants and Setting

The participants in this study were teachers and administrators of local districts within the TBAISD region in Michigan who participated in state-mandated evaluations conducted from 2015 to 2018. Because more than 100,000 data points existed in a 3-year time period in just one of the four domains of the evaluation tool, we decided to analyze a sample group of schools within the TBAISD using a stratified sampling technique (Boudah, 2020).

The names of the teacher and the administrator in each evaluation were removed. Evaluation data years included 2015–2016 (Year 1), 2016–2017 (Year 2), and 2017–2018 (Year 3). The four domains of the Marzano tool adopted by TBAISD provided categories for evaluation: (a) Classroom Strategies and Behaviors, (b) Planning and Preparing, (c) Reflecting on Teaching, and (d) Collegiality and Professionalism. The authors focused on collecting and analyzing data from Domain 1, Classroom Strategies and Behaviors. Within this domain, Marzano identified 13 “elements of rigor” or essential strategies to achieve rigor in the classroom (Marzano & Toth, 2013). These 13 Essential Strategies of Rigor included the following:

- Identifying Critical Content;
- Previewing New Content;
- Organizing Students to Interact with Content;
- Helping Students Process Content;
- Helping Students Elaborate on Content;
- Helping Students Record and Represent Knowledge;
- Managing Response Rates with Tiered Questioning Techniques;
- Reviewing Content;
- Helping Students Practice Skills, Strategies, and Processes;
- Helping Students Examine Similarities and Differences;

- Helping Students Examine Their Reasoning;
- Helping Students Revise Knowledge; and
- Helping Students Engage in Cognitively Complex Tasks.

This choice to focus on rigor was important, as the strategies within this segment of the model assisted teachers with creating conditions for maximum student learning. Thus, we analyzed data collected regarding the 13 elements of rigor identified within the evaluation model used by all locals within the ISD.

Research Procedures and Data Analysis

We used the following procedure to prepare the raw data for analysis:

Step 1: Organize and prepare data for analysis. We coordinated with the Director of Instructional Services at the TBAISD, as all evaluations were housed in the ISD server once they were completed by administrators in local districts. The data sets for each academic year from the randomly selected schools provided (a) comments by administrators to all evaluated teachers, (b) ratings of teacher performance on elements of rigor, (c) score count for each element of rigor, (d) score distribution for each element of rigor, and (e) count of building observation type (informal, formal, or walkthrough).

Step 2: Read and analyze all the data. Once these data sets were provided, a duplicate copy was made so that each of us could review the data separately first before collaboratively discussing patterns, themes, or results. We followed Tesch’s eight guiding steps for coding data (Tesch, 2008, as cited in Creswell and Creswell, 2018). Steps 1 and 2 were completed by each of us individually, and Steps 3–8 were done collaboratively:

1. Get a sense of the whole. Read the data carefully. Jot ideas as you analyze.
2. Pick one document to surmise the underlying meaning. Write thoughts in the margin.
3. Make a list of all topics. Cluster similar topics.
4. Take this list and go back to the data. Abbreviate the topics as codes and identify potential categories.
5. Look for ways to reduce the list of categories by grouping related topics.
6. Make a final decision on the abbreviation for each category and alphabetize.
7. Assemble the data material for each category and perform a preliminary analysis.
8. If necessary, recode existing data (p. 196).

Step 3: Collaboratively compare notes.

Step 4: Collaboratively share and interrogate the data to identify emerging patterns and evidence for those patterns.

Step 5: Cluster the data points to identify regional patterns and trends.

Step 6: Summarize findings with a few synthesis statements backed by the data.

By using these steps, we followed a process that vetted the descriptions, findings, and themes described in the results and implications for practice.

Results and Findings

What Were the Lessons Learned?

Three prominent findings emerged from the analysis of the data. First, the overall rating count and average score per the 13 elements varied widely within the 3 years of the study. The overall rating count was the number of times an administrator

would “score” an element in a teacher observation. The average score was a number somewhere between 0 and 4.0 that was an average of all the elements scored. Second, as Table 1 indicates, authors found the frequency of comments provided with ratings was high throughout the duration of the study, and the percentage of comments provided with each rating varied across each of the 3 years. The authors determined the percentage of comments made by administrators by reviewing all of the ratings made by evaluating administrators, then calculating the total number of times a comment was made when a rating was given to a teacher in an element. Year 1 was the first year of implementation and use of the Marzano model in the TBAISD region, and therefore the authors considered that year’s data as providing the baseline percentage for comments made with ratings by administrators in this study (Table 1). Finally, the comments made by administrators in the evaluation reports consistently indicated a lack of actionable and essential feedback that would move teachers to the next level of improvement.

Table 1

Frequency of Comments with Ratings

Rating Type	Year 1 2015–2016	Year 2 2016–2017	Year 3 2017–2018
Not Using	85%	100%	98.75%
Beginning	100%	96.16%	88.14%
Developing	95%	91.33%	90%
Applying	87.28%	81.44%	76.63%
Innovating	N/A	90%	79.85%

The first finding underscored that the overall rating count and average score per element varied widely across all elements within the 3 years of the study. The total number of ratings for each element year-to-year varied, while the average score of the ratings year-to-year remained relatively constant. Results also indicated that, of the schools sampled, all elements of rigor over the 3 years averaged an “effective” rating or higher. Very few elements were scored lower than “effective.” Teachers were doing a good job based on observation scores, which means student achievement should have been improving. However, this conflicts with national data of student performance. Many states across the country have seen a drop in student test scores since adopting more rigorous standards (Learning Sciences, 2017) even though teachers remain “effective” or “highly effective” (MDE, 2020).

We reviewed 3,441 rating data points over the 3 years of the study. Overall, 34.6% of ratings conducted by administrators in the study were for elements associated with lecture, practice, and review. However, elements associated with developing critical thinking and cognitively complex skills made up only 7.8% of ratings. With the exception of Elements One (*Learning Goals and Scale*) and Six (*Identifying Critical Content*), all remaining elements measured in the study increased total rating count over the course of the study.

Each administrator evaluator had a choice in elements scored. He or she also had a

choice in how many elements to score based on the individual observation. However, administrator evaluators overwhelmingly chose to observe the elements *Learning Goals and Scales* and *Identifying Critical Content* in Year 1. This was likely due to the professional learning received on the evaluation tool when administrators were told that these two elements were foundational: If teachers identified the critical content and shared learning goals and scales, students would demonstrate more learning. *Learning Goals and Scales* and *Identifying Critical Content* were also the lowest-scoring elements all 3 years, with only *Identifying Critical Content* breaking the 3.0 rating in the Year 3. In addition, the evaluating administrators, over the course of the 3-year study, rated and provided feedback most heavily on elements leading up to cognitively complex tasks designed to elicit higher-order thinking in students. For example, the element *Helping Students Record and Represent Knowledge* improved in the rating count each year of the study (in Year 1, 35 ratings; in Year 2, 62 ratings; and in Year 3, 82 ratings). Similarly, *Previewing New Content* rose each year (in Year 1, 19 ratings; in Year 2, 39 ratings; and in Year 3, 52 ratings). Yet elements or strategies requiring students to demonstrate application and higher-order thinking were not rated nearly as frequently over the 3 years of the study. For example, the element *Helping Students Revise Knowledge* received 7 ratings in Year 1, 13 ratings in Year 2, and 23 ratings in Year 3. The element *Organizing Students for Cognitively Complex Tasks* received 2 ratings in Year 1, 9 ratings in Year 2, and 25 ratings in Year 3. This is significant, as the premise for adopting the state-wide strategy was based on the belief that stricter accountability systems and attention to the technical core of teaching and learning would yield increased results for students (Elmore, 2000). Yet administrators had not begun to rate teachers on strategies or elements within the model that would most improve student learning.

The frequency of comments (i.e., feedback) within the ratings given by administrators yielded a high percentage over the 3 years. The administrators made comments associated with teacher action and performance in the observations conducted. As specified in Table 1, the administrator evaluator identified the level of use of the element as *Not Using*, *Beginning*, *Developing*, *Applying*, or *Innovating*. These levels of use were put into each teacher’s evaluation account and aggregated for a final instructional rating (Table 2). The values indicated in Table 2 were determined by Learning Sciences International and adopted by the TBAISD and local districts when this evaluation model was approved for use in the region. Learning Sciences International is the proprietor of the Marzano evaluation model.

Table 2

Instructional Practice: Aggregated Ratings by Administrators

Rating	Score
Ineffective	0.0 – 1.49
Minimally Effective	1.50 – 2.49
Effective	2.50 – 3.49
Highly Effective	3.50 – 4.0

Administrators provided feedback across all rating levels; however, observations with the *Applying* designation had the fewest comments (with 87% in Year 1, 81% in Year 2, and 77% in Year 3). iObservation software embedded within the Marzano evaluation tool calculates the percentage of comments associated with ratings conducted by evaluators for each school. We tallied the percentages from each

“
When we examined the quality and substance of the feedback given, we did not find actionable feedback that the teacher could use for the next level of work on instructional strategies or pedagogy.
”

randomly selected school in this study and calculated a mean or “average” percent of comments provided with ratings given by administrators. Observations with ratings in the *Not Using*, *Beginning*, *Developing*, and *Innovating* categories had comments at least 90% of the time a rating was provided. However, when we examined the **quality** and **substance** of the feedback given, we did not find actionable feedback that the teacher could use for the next level of work on instructional strategies or pedagogy. The quality and substance were evaluated through the use of TBAISD’s feedback

rubric based on a 4-point scale. If the feedback given included questions or ideas for next levels of work on given elements, we identified it as a “4.” If the feedback given was verbatim from iObservation’s suggested feedback, it was identified as a “3.” When feedback included some specific language around the elements observed but used general terms such as “good,” “great,” or “I liked,” it was considered a level “2” feedback. If the feedback was merely a regurgitation of what occurred in the observation, it was given a “1” as it did not push any thinking or improvement of instruction.

Overwhelmingly, feedback was given via comments describing what was witnessed in the observation. For example:

- “Students spend the majority of the time comparing and contrasting the quotes from *The Lightning Thief* and *A Hero’s Journey*.”
- “The majority of students are paying attention during the demonstration.”
- “The class came in, got out their chrome books and opened their textbook on the chrome book.”
- “A warm up was displayed. Before having the students solve the problem, she had them look at page 177 and walk through which of the units of length in the book gave them the information they needed to solve the problem. She took volunteers and then solved the problem as a class.”

We observed questions, suggestions, or directives most frequently in elements rated as *Developing*, *Not Using*, or *Beginning*:

- “It is worth taking the time to make sure they (students) know what it is before revealing the rest of the scale.”
- “Ideally, learning targets or objectives should be posted and communicated in a way that uses a verb for students to implement the learning.”
- “It would be more efficient to have the notes page already set up on your device, along with all practice problems and items to demonstrate ready to go.”

We found that a significant percentage of the administrators’ comments to teachers were copied from the scale provided within the observation tool. Similarly,

administrators used some of the reflection questions provided within the tool for each element to “stimulate or begin” a conversation with teachers.

A majority of comments provided in the *Innovating*, *Applying*, *Developing*, *Beginning*, or *Not Using* categories showed a distinct pattern: a statement about what was observed; a follow-up question, suggestion, or judgement from the observer; and then a closing comment with an expectation for the future. As a researcher scanning the comments, one could predict the questions and suggestions for improvement would occur near the end of the feedback.

Beginning and *Developing* ratings contained a question, suggestion, or judgement regarding performance a majority of the time. *Applying* and *Innovating* ratings did not use questions, suggestions, and judgement to support growth and greater mastery of the element. The vast majority of these ratings simply restated what was observed and often provided complimentary feedback to the teacher.

In summary, based on the comments provided by regional evaluators in this study, we surmise that principal feedback does not impact teacher performance. The majority of the ratings were either *Applying* or *Innovating*, and *Applying* or *Innovating* comments were descriptive, merely stating what the principal had observed in the classroom.

Implications for Practice

Feedback is effective with multiple data points over time. Timeliness and accuracy are important to improving instruction. Based on anecdotal conversations, building leaders and evaluators find that one-on-one conversations with teachers around instruction often result in constructive conversations. Teachers want to improve their practice through meaningful and specific feedback. However, it seems that the moment a score is placed on the evaluation, much like a grade on a student test, meaningful, growth-minded conversations are stunted. We would like to understand better the connection between feedback and teacher effectiveness so that building and district leaders can have meaningful conversations that promote growth.

Teachers will most likely listen to feedback if they believe that feedback accurately represents the teaching and learning that has occurred (Cherasaro et al., 2016). Feedback provided in real time or in short feedback cycles is more beneficial. It is also best to provide feedback when teachers are acquiring or practicing new strategies to minimize incorrect practices of teaching a specific strategy. However, when teachers are applying, refining, or extending instructional practices, it is more beneficial for the teachers to discover errors and nuances themselves in the use of the instructional strategy before providing the feedback (Cherasaro et al., 2016).

**“
Changing legislation to a high-stakes, stressful situation will not enhance teaching and learning on its own; teachers and leaders need tools and resources to facilitate growth in all.
”**

Implications for Future Study

The following three questions might be considered next steps of this particular research:

- What knowledge and skills do the evaluators have as it relates to providing feedback?
- What changed for teachers and students as a result of teacher-evaluation processes?
- How can a strategic plan that is designed around student growth and achievement create an efficient improvement cycle in staff talent and instructional infrastructure?

Feedback should strengthen a teacher's efficacy to create a classroom that gives all students high-quality instruction and classroom management. Feedback should provoke teachers to think deeply and abstractly about their profession. This is accomplished by providing feedback that creates a safe enough climate for teachers to self-appraise, reflect, and adapt practices to the changing teaching environments and the needs of their students (Kee et al., 2017). Changing legislation to a high-stakes, stressful situation will not enhance teaching and learning on its own; teachers and leaders need tools and resources to facilitate growth in all.

References

- Boudah, D. J. (2020). *Conducting educational research: Guide to completing a thesis, dissertation, or action research project* (2nd ed.). SAGE.
- Center for Public Education. (2005). *Teacher quality and student achievement: Research review*. Author.
- Cherasaro, T., Broderson, R., Reale, M., & Yanoski, D. (2016). *Teachers' responses to feedback from evaluators: What feedback characteristics matter?* Institutional Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. <http://ies.ed.gov/ncee/edlabs>.
- Creswell, J., & Creswell, J. (2018). *Research design: Qualitative, quantitative and mixed methods approaches*. SAGE.
- Elmore, R. (2000). *Building a new structure for school leadership*. Albert Shanker Institute.
- Grissom, J. A., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. *Educational Evaluation and Policy Analysis*, 37(1), 3–28. <https://doi.org/10.3102/0162373714523831>
- Holland, E. (2014). *Making feedback work*. Executive Development.
- Kee, K. (2017). *RESULTS coaching next steps: Leading for growth and change*. Corwin.
- Learning Sciences International. (n.d.). *iObservation*. <https://www.iobservation.com/>
- Learning Sciences. (2017, June). *Learning Sciences essentials for achieving rigor*. Author. <https://www.learningsciences.com/wp/wp-content/uploads/2017/06/Essentials-Service-Index.pdf>.
- Marzano, R. (2013). *The Marzano teacher evaluation model*. Marzano Center. Learning Sciences International. <https://www.sjschools.org/images/District/TeacherEvaluation.pdf>
- Marzano, R. (2014). *Marzano teacher evaluation model: Learning map*. Marzano Center, Learning Sciences International. https://learn.k20center.ou.edu/api/pds/1c2bb46ffdf0fed14bcbaaf4909b3ea/2014_Marzano_Learning_Map.pdf
- Marzano, R., & Toth, M. (2013, July). *Deliberate practice for deliberate growth: Teacher evaluation systems for continuous instructional improvement*. In *iObservation Software*. <https://www.learningsciences.com/tech/iobservation/>

- Michigan Department of Education. (May 2019). *Michigan educator evaluations at-a glance*. https://www.michigan.gov/documents/mde/Educator_Evaluations_At-A-Glance_522133_7.pdf
- Michigan Department of Education. (2020, February 18). Educator effectiveness snapshot. *MI School Data*. <https://www.mischooldata.org>.
- Public Act 173 of 2015. Revised School Code, 380.1249. (2015). *Michigan Legislature, Michigan Compiled Laws*. <https://www.legislature.mi.gov/documents/2015-2016/publicact/pdf/2015-PA-0173.pdf>
- Strong, R., Silver, H., & Perini, M. (2001). *Teaching what matters most: Standards and strategies for raising student achievement*. Association for Supervision and Curriculum Development.
- Thompson, J. (2006). *Organizations in action*. Transaction Publishers.
- Wiggins, G. (September, 2012). Seven keys to effective feedback. *Educational Leadership*, 70(1), 10–16.
- William, D. (2016). *Leadership for teacher learning: Creating a culture where all teachers improve so that all students succeed*. Learning Sciences International.