11-2017

Mitigating the Tensions of Teacher Evaluation through Pragmatism

Scott Sheedlo

Ovid-Elsie High School, scott.sheedlo@ovidelsie.org

Follow this and additional works at: <https://scholarworks.gvsu.edu/lajm>

Part of the [Curriculum and Instruction Commons](#), and the [Teacher Education and Professional Development Commons](#)

Recommended Citation

Sheedlo, Scott (2018) "Mitigating the Tensions of Teacher Evaluation through Pragmatism," *Language Arts Journal of Michigan*: Vol. 33: Iss. 1, Article 8.

Available at: <https://doi.org/10.9707/2168-149X.2167>

This Article is brought to you for free and open access by ScholarWorks@GVSU. It has been accepted for inclusion in Language Arts Journal of Michigan by an authorized editor of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

PRACTICE

Mitigating the Tension of Teacher Evaluation Through Pragmatism

SCOTT SHEEDLO

Teacher evaluation is state law in Michigan (Legislative Council, 2015). To satisfy this requirement, many public school districts have adopted the *Five Dimensions of Teaching Learning* evaluation model from the University of Washington's Center for Educational Leadership (CEL). According to the Michigan Association of Secondary Principals website, 187 Michigan districts and intermediate districts use the "5 D+" model. The district in which I teach implemented it approximately three years ago. Due to the popularity of the 5D+ model in Michigan, many K-12 teachers are familiar with its five dimensions (purpose, student engagement, curriculum and pedagogy, assessment for student learning, classroom environment and culture, professional collaboration and communication) and its scoring rubric (distinguished, proficient, basic, and unsatisfactory) (Center for Educational Leadership, 2012). In being evaluated by this model, they may have discovered, as I have, that its implementation raises many questions and tensions that are difficult to understand, especially for veteran teachers. As I will argue, the model presumes a research-based body of expert teaching practice that may significantly differ from what teachers and administrators had previously considered good teaching. From this discrepancy comes the potential for misunderstanding and mistrust between teachers and their evaluators that are difficult to work through when previously proficient or distinguished teachers start to earn ratings of basic or even unsatisfactory on practices they realize are actually unfamiliar. In response, teachers may become resistant to change, which further exacerbates the problem. This is a story of my struggle through this on-going adjustment as a twenty-four year veteran English Languages Arts (ELA) teacher and of how embracing a pragmatic mindset is helping me stay positive and relevant through this period of intense professional change.

Teacher Evaluation and Teacher Resistance

Like Thoreau (1849) in "Civil Disobedience," who met the government in the person of the tax collector to exercise resistance to its intrusion on his life, teachers meet this state mandate in the person of a district administrator responsible for evaluation. When there is agreement between the teacher, the administrator, and the evaluation language on what good teaching looks like, evaluation is an affirming and enjoyable process. However, if there is confusion or disagreement, then tension emerges, as teachers are forced to give additional time and energy to learning the unfamiliar terminology and possibly to changing their practice.

The 5D+ model contains multiple research-based terms that caused confusion when it was introduced to our staff. In this article, I will focus primarily on formative assessment and differentiated instruction. Early on in our district, understanding learning targets and formative assessment required further definition and received professional development time, which did help. Differentiated instruction has not been addressed specifically, but as my research suggests, it may need to be. However, what was most troubling to me were the descriptions in the model of teaching practices that were clearly a departure from older, accepted practices, going well beyond a new name for what good teachers have always done. The example of this that brought me to a professional decision point is the sub-dimension under Assessment labeled A6, "Adjustments." It reads a teacher must use "formative assessment to make in-the-moment instructional adjustments, modify future lessons and give targeted feedback aligned with the learning target to individual students" to be distinguished (Center for Educational Leadership, 2012, p. 5). I asked myself, am I using formative assessments and making instructional adjustments on the fly? I thought in a broad sense I probably was. However, when I asked myself,

am I modifying future lessons and giving targeted feedback to individual students? I was pretty sure I was not, at least not with every lesson and in every strand of ELA teaching. From this analysis, I calculated that meeting this demand was going to require more than just a “tweak” of my teaching practice to accommodate. I was going to have to rethink how I plan, assess, score assessments, and give feedback because my current planning strategies did not allow for this.

My first inclination was to resist making these changes. The findings of Ewald Terhart on teacher resistance indicate this is normal. In studying the resistance of German teachers to educational reform, Terhart (2013) confirmed previous studies in the 1990s of resistant attitudes to reform of US teachers. He found teachers give six common arguments for resisting research-based reforms:

1) “The ‘No time!’—argument:” Given the already full work day, teachers cannot imagine where they will find the time to adjust their practices.

2) “The ‘I am innocent!’—argument:” Teachers cannot see the logic of changing their practices when so many of the problems in education are the fault of others and beyond their control.

3) “The ‘burnt child’—argument:” This is a workplace version of the burnt child syndrome, where children who have been wronged by adults develop an ongoing mistrust that things will ever get better. For teachers, it is the sense that many reforms have been tried in the past and the problems remain the same, so why change?

4) “The ‘two worlds’—argument:” Teachers believe reform ideas are developed by non-teachers (i.e., researchers and politicians) who do not understand the world of the classroom. Therefore, their ideas do not work.

5) “The ‘biographical’—argument:” Veteran teachers have taken their turn experimenting with reforms. They argue it is up to the younger teachers to be the guinea pigs.

6) “The ‘lack of personal benefit’—argument:” Teachers see school reforms as benefitting researchers and administrators. The only thing they get out of them is more work (p. 494).

Some version of all these arguments went through my mind as I considered the demand of changing to comply with the evaluation rubric. Teachers should not feel guilty for thinking this way; however, they should ask themselves, is it worth it to resist reform?

Before sharing how I decided to deal with that question, it is important to explore the depth of this conflict further as it relates to the 5D+ evaluation. Terwald’s (2013) work is insightful here as well. It is not that teachers are uninterested in what is best for students; it is that they have developed personal metrics for what works with students that they trust from experience (p. 493). These personal metrics may not be fully aligned with current research trends on best practice. All teachers have a mental list of students who are examples of their successful teaching practices. Often those exemplar students stand in as archetypes for the range of ‘all students whose needs should be met,’ which is held up as the standard goal of public education and given as the reason for so many educational reforms. Consequently, in our minds our teaching practice is at least proficient and often distinguished based on the time and effort we have put in and this mental review of our student success stories.

This personal metric view is in direct conflict with the philosophy of the developers of the 5D+ model. In *Leading for Instructional Improvement, the professional development guide for the 5D+ model*, authors Stephen Fink and Anneke Markholt (2011) present a very different view of teaching practice from the self-satisfied view of teachers. They argue in the introduction, “The quality of teaching in the vast majority of our schools is inadequate to ensure quality learning for all students...By all students we mean just that—each and every student, including students academically behind, students still developing the English language to make sense of their academic subjects, and students who have adapted some form of behavior to cope but not learn” (pp. xvii-xviii). Their confidence in the model’s ability to address this inadequacy comes from a belief that a body of educational research knowledge exists that can effectively meet the needs of all students. The problem, in their view, is not intractable. The solution lies in developing leaders in local districts who can help teachers become experts in this body of teaching practices, which are embedded in the 5D+ model. They contend, “Without a shared understanding of what we mean by quality instruction, we have no basis from which to mount an improvement effort. This is an issue of expertise or in our case a lack of sufficient expertise necessary to improve the quality of teaching in every school and every classroom...This means the primary role of school and district leaders must be the cultivation of expertise to

improve practice, including both teaching and leadership practice” (p.5). In short, the CEL believes that despite what their personal metrics indicate, teachers do not have the expertise they need in the practices that will effectively ensure learning for all students. They need to change and administrators need to be leaders in this transformation. In this context, the choice confronting me and other teachers under the 5D+ evaluation process can be framed by these questions: Do I try to defend the teaching practice expertise I believe I already have? Or do I accept the promise of expertise from someone outside my classroom, even though the promise of others like them in the past has failed?

Critical Pragmatism Provides a Way Forward

After about a year of discomfort with the model and indecision about how much to adjust my teaching to accommodate the deficiencies it suggested, I tacitly committed to allowing the evaluation process to change my teaching. The decision was consciously pragmatic. By that I mean I was interested in achieving desirable outcomes instead of defending personal beliefs about my philosophy of good teaching. I decided it would be easier to face the “adjustment” sub-dimension head on by choosing it as the focus of my teaching goal for the 2016-17 school year than to allow my discomfort to remain. It was an optimistic choice, yet lacked enthusiasm.

Pragmatism as an educational philosophy was familiar to me from working on my master’s degree in Curriculum and Teaching. In *Reading Pragmatism*, Cleo Cherryholmes (1999) describes the character of this philosophy as “a discourse that attempts to bridge where we are with where we might end up” (p. 3). He takes care to distinguish his meaning from a view of pragmatism as vulgar, a mere expedient to immediate gain, to a view of it as critical, which ties its use to a concern for finding acceptable outcomes in terms of aesthetics and power (p. 7). Pragmatic teachers consider the outcomes for the school, their students, and themselves and seek out acceptable compromises. It requires a wider view that goes beyond personal beliefs and issues.

In this case critical pragmatism was particularly effective in easing tensions over expertise as I have described between a teacher’s personal metric for effectiveness and the research-based expertise of the 5D+ model because it allows teachers to compromise personal beliefs in favor of organi-

zational harmony. If teachers view effective teaching practice as based on a set of principles that cannot be compromised, they find themselves in conflict with the evaluation model if the set of “expert” principles from the model do not agree with their own. Since teachers develop their set of principles from practice, and not from a research paradigm, it is very likely the conflict will exist. If, instead, they focus on desirable outcomes, they can try unfamiliar practices with the outcome of cooperating with the larger goals of the organization so that they may engage in the ongoing discourse on the new practices with both colleagues and administrators. Thus, alienation from the organization, an unfavorable personal outcome, is avoided.

Furthermore, teachers who choose compliance move to the position of learners, which, if embraced, has its own merits when compared with being infallible experts in an occupation where change is the norm and there are few guarantees. They now are able to enter into the inquiry of the new teaching practice, allowing them to better judge its effectiveness so that they may employ it regularly or criticize it from perspective of experienced practitioners. These are highly desirable outcomes when facing an uncertain teaching future. This is what I hoped for in choosing to comply fully with the evaluation process.

These benefits of pragmatism are especially useful for public school teachers because public schools reflect the political interplay of the democracy they support. Just as pragmatism allows citizens to resolve the problems that arise from conflicts of aesthetics, knowledge, and power in governance, it allows teachers to find desirable solutions for similar conflicts as they arise in education. Cherryholmes (1999) summarizes this advantageous link between democracy and pragmatism with five assertions:

- 1) There are no guarantees “individual conceptions” of effectiveness are accurate.
- 2) Since the future is unknown, it is reasonable to explore a wide variety of courses of action in case others conceptions are better.
- 3) Efficient and open communication among individuals and groups makes for better consideration of outcomes.
- 4) Adherence to a particular set of principles or tradition (especially in educational research) limits the imagination of possible outcomes.

5) Pragmatism and democracy are allies because pragmatism needs democracy's openness and democracy needs the experimentation of pragmatists (p. 40).

Hence, the pragmatic teacher in this view is liberated to use the reform agent of teacher evaluation to imagine better outcomes that go beyond the limits of a philosophical conflict over best practice.

Change in Action: Assessing and Adjusting with Reading Instruction

While the decision to let go of my resistance to the evaluation process relieved much of the initial tension, there was still the work of learning the concepts and strategies, developing the materials, and the struggling through the trial and error of implementation that brought new concerns. For ELA teachers the questions of how to formatively assess, collect data on assessments, and then use the data to adjust instruction for reading, writing, speaking and listening, and sub-skills such as grammar and vocabulary are not necessarily answered the same for each strand. The choices for the type of assessment and how feedback is given may be quite different among them. Of these strands, I have found applying this to reading has been the most challenging; therefore, I will focus here on some of the steps I took for that purpose to illustrate the ebbs and flows of changing one's teaching practice.

Early on I analyzed the "Adjustment" sub-dimension as drawing from research on formative assessment and differentiated instruction, yet I did not research the literature on these areas because I thought it would be enough to rely on the professional development already provided and the collaboration with colleagues to figure out what to do. I had recently joined our building school improvement team and was involved in discussions about how to improve reading scores on the SAT. We had identified citing textual evidence as a strategy for the school improvement plan's reading goal and had planned to take a release day as a department to flesh out how we would support this strategy. Since I knew from previous conversations with my ELA colleagues they too were concerned about the references to formative assessment, data collection, and adjusting instruction in the 5D+ rubric, I suggested to them we use this opportunity to develop a reading assignment that would both focus on citing textual

evidence and satisfy this part of the evaluation rubric as well.

As we worked at this task, a new tension arose. Would we use multiple-choice for reading assessment or open-ended response? Our need to collect data and concern over additional workload made the multiple-choice option tempting. However, experience had made me wary of the usefulness of multiple-choice reading data. My models for good reading instruction have always involved a lot of writing about the reading to give both reader and teacher a window into the metacognitive processes at work. Was this merely an aesthetic concern or was it substantive? As I learned later on from research, my concern was legitimate.

Research on the role of written responses in reading assessment criticizes an over-reliance on multiple choice assessments from both assessment and pedagogical perspectives. In the area of reading assessment, Ozura, Kurby, Briner, and McNamara (2013) studied the difference in performance between both types of reading assessment and demonstrated they measure different aspects of comprehension. The authors concluded multiple-choice questions correlate to measurement of the level of prior knowledge related to the text, whereas open-ended questions correlate to the quality of self-explanations. In considering which would be more helpful for adjusting instruction of reading skills, especially in citing textual evidence, I contend that measuring the quality of self-explanations is far more useful than the students' prior knowledge related to the text.

In terms of pedagogy, a meta-analysis on the use of writing with reading instruction has shown "writing about material read improves students' comprehension of it" (Graham & Hebert, 2011, p. 710). There is a danger, then, that as teachers try to save time with multiple-choice reading assessments, even if it is to facilitate data collection and instruction adjustment, they are reducing the overall quality of reading instruction. That concern is heightened by the influence of high-stakes testing, which relies so heavily on multiple-choice assessments.

In spite of this support, I am still mulling over the tangible sense of unease among us as we deliberated on what the assessment would look like and how it would be implemented. There was a general feeling that it may be too time-intensive to be manageable. Nevertheless, we decided to use a written response as our assessment instrument. The assignment protocol we developed was to assign the students

a text-dependent question during the reading of a complex text to respond to in writing, which would be scored using a rubric we developed together to measure the students' ability to cite textual evidence to infer a logical understanding of the text. The students' score on the rubric would be recorded to be used for school improvement data analysis; meanwhile, the teacher would be able to use the assessment data to adjust instruction before a summative assessment was given on the text.

My implementation of this protocol with students revealed both tangible success and the realization of my biggest fears. To the positive, I saw my students' thinking in their responses—I saw if they were paying attention to the relevant details and linking them to logical inferences about the reading. Looking at an individual students' work, I imagined how I might intervene to help them read and understand the text better. However, on the negative side, I did not reach a level of proficiency in the six month trial period where I could read all the responses (typically around 125 for my sophomore English classes), develop interventions for adjusted instruction, give the summative assessment, and then use the summative results to evaluate the effectiveness of my intervention strategies. As the school year ended, I found I had collected useful data that I did not have time to use to adjust instruction and intervene with students. At this point, I remain uncertain if I can make this protocol work next school year.

Returning to Research Literature

To deal with this tension related to implementation, I made another pragmatic decision. Already enrolled in an online graduate program, I decided to use an assignment to research literature on formative assessment and differentiated instruction. This choice allowed me to do the required work and to continue to examine the possibility of applying formative assessment techniques and differentiated instruction for reading as prescribed by the model. As I read research literature in those areas, I was looking for misunderstandings or missed technical details that might suggest ways to improve efficiency, but also was trying to make sure I understood the practices embedded in the model correctly. Although I read widely, I strategically included two popular authors whose work is frequently promoted by the educational research

and professional development organization Association for Supervision and Curriculum Development (ASCD), Robert Marzano on formative assessment and Carol Ann Tomlinson on differentiated instruction.

In looking at formative assessment, I wondered if I had misunderstood the emphasis on data collection implied by the 5D+ rubric. If the burden were removed to collect and use formative assessment data to adjust instruction, the process might be closer to what I have always done. However, Marzano's work confirmed the expectation of using at least some numerical data scores for adjusting instruction. In *Formative Assessment and Standards-Based Grading* (2010), he traces the understanding of formative assessment to the late nineties where it generally referred to all the activities a teacher uses to gather information for feedback and modifying teaching, even those not scored like observing and talking to students. However, he positions formative assessment that uses scores as a key part of the larger innovation of standards-based grading. To emphasize this, he transitions his terminology from formative assessment to formative scores, referring to the numeric data collected from a formative assessment (p. 27). While he notes that formative assessment without scores can be used for instructional feedback (p. 31), Marzano contends formative scores should lead to changed behavior on the part of both the student and the teacher. He explains, "On the teacher's side, behavior change involves identifying content that must be reviewed or retaught" (Marzano, 2010, p. 33). So it is not enough for a teacher to say, 'I use non-scored formative assessment and adjust lessons for the next time I teach a topic.' Marzano's work links formative assessment more tightly to scored assessments than some of the earlier definitions and to adjusted instruction that involves at least some re-teaching.

Next, I wondered about the reteaching. Is that differentiated instruction? Marzano does not use that term specifically, nor does the A6 sub-dimension on adjustment in the 5D+ model. Still, the model does mention it in the Curriculum and Pedagogy subdimension #5, "Teaching Approaches and/or Strategies: Differentiated Instruction." The rubric says a proficient teacher "frequently uses strategies that differentiate for individual learning strengths and needs" (Center for Educational Leadership, 2012, p. 17). If one adds that to the wording in the "Adjustments" sub-dimension distinguished level, "Teacher uses formative assessment data to make in-

the-moment instructional adjustments, modify future lessons and give general feedback aligned with the learning target to individual students,” the similarity is notable. Finally, if those two examples from the model are compared with descriptions from expert Carol Ann Tomlinson’s latest book on the differentiation, the conclusion that the model insists on differentiation is inescapable. Tomlinson (2014) explains the link between formative assessment and differentiation this way: “[F]ormal and informal assessment yields an emerging picture of who understands key ideas and who can perform target skills, at what proficiency, and with what degree of interest. The teacher then shapes tomorrow’s lesson—and even reshapes today’s—with the goal of helping individual students move ahead from their current position of competency” (pp. 17-18). On the use of data she elaborates, “The teacher in a differentiated classroom thoughtfully uses assessment data to guide modifications to content, process, product, or learning environment” (p. 18). From the similarities I conclude that to satisfy the model, I will need to change my practice in the direction of differentiation.

However, before jumping to the conclusion that the model is recommending differentiated instruction for every lesson, every day—an impossible standard—I observed this qualifier from Tomlinson (2014) as well: “[T]eachers need not differentiate all elements in all possible ways in every unit. Effectively differentiated classrooms include many times in which whole class, nondifferentiated fare is the order of the day....” (pp. 19-21). The appearance of ambiguity in the model, if compared to the research literature, is actually an appeal for judgment on the part of the teacher. While a teacher cannot avoid using formative assessment scores to differentiate instruction, the questions of when and how are implicitly the teacher’s call. This is problematic for the evaluation process if the evaluator does not have the opportunity to observe this aspect of the teacher’s practice.

As I consider the value of reading Marzano, Tomlinson, and other authors who speak to the expertise utilized by the 5D+ model, like so much about this journey to change, the results are mixed. I did get a clearer sense of the terminology of the evaluation tool and found many examples of using formative assessment and differentiation, which are not shared here, but provide help in further experimentation. From my reading, I have become further convinced that

more time and effort will be required to read and re-read more, make new attempts at these practices, revise my practice more, and struggle further with my lack of expertise. As I think about this, I wish it were easier.

Staying Relevant in the Discourse Community

Nevertheless, I remain positive and optimistic about the future. I see no guarantees of mastering the model, but I am more confident in discussing the model and its practices. It is easy for veteran teachers to assume their position as experienced practitioners will always allow them to participate in educational discourse. My experience with the 5D+ model has reminded me that if one loses fluency in the terminology of the meta-discourse, in grasping operative meanings of formative assessment and differentiation in this case, one will not be able to effectively criticize their value for teaching practice. The set of effective, research-based practices of the 5D+ model come from a larger educational discourse community. According to Gee (1989), membership in a discourse community comes from an ability to use the community’s language, thinking, and acting (p. 18). He argues that one cannot criticize a discourse community from the outside; instead a critic must develop fluency in the meta-discourse of the community to argue from the inside. The critic must learn and acquire a powerful literacy—that is control of the meta-discourse language of the discourse community (Gee, 1989, p. 23). When discussions over the effectiveness of the evaluation model play out in local districts, the teachers being evaluated often do not realize they are not in a position to effectively argue the merits of the evaluation process until they become powerfully literate in the language of the research-based practices of the larger educational community that has generated the model.

By choosing compliance over resistance and conducting my own research into the model, I have gained the desirable outcome of being able to effectively participate in discussion of this model that has power over my professional life. Furthermore, I am able to contribute a larger, organizational effort to educate students at the building and district level. Where my sense of authority and professional satisfaction used to come from executing a personal philosophy of how to educate students from a teacher-centric mindset, I am

learning the satisfaction that comes from participating in larger building and district efforts to educate students from an organization-centric mindset. This is a good place to be right now, and I do not think it would have been possible if I had tried to slavishly stick to a personal philosophy of teaching. Pragmatism gave me permission to consider other possibilities. I think all K-12 teachers can benefit from this expansion of educational thought to help with the tensions of change and to remain relevant contributors to their schools and profession.

References

- Center for Educational Leadership (2012). *5D+ teacher evaluation rubric*, 2. Seattle, WA: University of Washington College of Education.
- Cherryholmes, C. H. (1999). *Reading pragmatism*. New York: Teachers College Press.
- Fink, S. & Markholt, A. (2011). *Leading for instructional improvement: How successful leaders develop teaching and learning expertise*. San Francisco, California: John Wiley and Sons.
- Gee, J.P. (1989). Literacy, discourse, and linguistics: Introduction. *The Journal of Education*, 171 (1), 5-176. Retrieved from www.jstor.org
- Graham, S. & Hebert, M. (2011). Writing to read: A meta-analysis and the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81 (4), 710-744. Retrieved from <http://www.studentachievement.org>.
- Legislative Council, State of Michigan. (2015). Section 380.1249. Revised school code: Act 451 of 1976. Retrieved from <http://www.legislature.mi.gov>
- Marzano, R. J. (2010). *Formative assessment and standards-based grading*. Bloomington, Indiana: Marzano Research Laboratory.
- Ozuro, Y., Briner, S., Kurby, C., & McNamara, D. (2013). Comparing comprehension measured by multiple choice and open-ended questions. *Canadian Journal of Experimental Psychology*, 67 (3), 215-227. Retrieved from <https://www.apa.org/pubs/journals>.
- Terhart, E. (2013). Teacher resistance against school reform: reflecting an inconvenient truth. *School Leadership and Management*, 3 (5), 486-500. <http://dx.doi.org/10.1080/13632434.2013.793494>.
- Thoreau, H.D. (1849). Civil disobedience. *The Thoreau reader (Part 2)*. Retrieved from <http://thoreau.eserver.org/civil2.html>.
- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners (2nd edition)*. Alexandria, Virginia: ASCD.

Scott Sheedlo is in his 25th year of public school teaching, during which time he has taught grades 7-12 in both Language Arts and Biology. A proud Marauder, he has taught at Ovid-Elsie High School the last twenty-three years, where he currently teaches sophomore English, Technical Communication, Novels, and Creative Writing. He is also a member of his building school improvement team. Presently, he is enrolled at Bowling Green State University as a graduate student in the Professional Writing and Rhetoric online MA program. He also holds a master's degree in Curriculum and Teaching and a bachelor's degree in Secondary English Education from Michigan State.