

Improving Teacher Evaluation by Walking the Talk of Standards-based Grading: Communicating Educator Growth using Proficiency Scales

Chad Lang

Glenwood Community School District, USA

Matt Townsley

University of Northern Iowa, USA

ABSTRACT

Teachers and school leaders frequently express a disconnect in the purpose and importance of teacher evaluation, particularly as it relates to educator growth. At the same time, some schools are beginning to communicate student growth through a standards-based grading philosophy. One way schools might “walk the talk” of their grading reform efforts designed to communicate student growth is through the use of proficiency scales to prioritize growth in teacher evaluation. This paper describes implications of simultaneously utilizing a growth model for teacher evaluation and a student growth model via standards-based grading.

Keywords: teacher evaluation, proficiency scales, standards-based grading

Historically, teacher evaluation is a tool with distinct purposes: evaluating teachers and assessing teachers (Marzano, 2012; Stiggins, 1986). When school leaders use teacher evaluation to merely evaluate teachers, many educators meet standards in a check-box-like binary manner. Unsuccessful teachers are deemed *in need of improvement* without so much as a continuum to see their own growth. Bradford and Braaten (2018) have framed teacher evaluation as contributing towards the “demoralization of teachers” in which tensions surrounding accountability supersede pedagogical improvement. School leaders have suggested accountability-driven teacher evaluation contributes towards declining morale and a “sense of their own limited professional value in the district” (Paufler, 2018, p. 11). Conversely, teacher evaluation intended to assess teachers provides a continuum of feedback ranging from unacceptable to excellence (Danielson, 2011; Wind et al., 2018). Effective teacher evaluation is based upon a set of professional teaching standards in which educators are provided opportunities to self-assess and receive feedback along a continuum (Darling-Hammond et al., 2012; Weems & Rogers, 2010).

Teachers are posed with a similar quandary about whether their grading practices are designed to sort or develop students (Guskey, 2015). Grades are often based upon an accumulation of points and do not accurately communicate learning (Brookhart et al., 2016). When the student focus in the classroom is earning points rather than learning, students may overly focus on recognizing point-earning opportunities while being blind to the level at which they are learning. Alternatively, grading models designed to honor student growth over time reveal the learning goals for which students have demonstrated understanding as well as the ones for which additional instruction is needed. Experts define grading models, which communicate learning relative to mastery of standards, using the term: *standards-based grading* (SBG; O’Connor, 2018; Welsh, 2019). A majority of principals in one state report implementing SBG as part of their five-year vision (Townsley et al., 2019) and school districts in additional states have also reported a strong consideration for moving in this direction (Townsley, 2021). Moreover, teachers in SBG models identify benefits such as: focused instruction, improved learning climates, and a student shift toward growth mindsets (Knight & Cooper, 2019).

Perhaps the irony for schools implementing SBG is that these standards-based practices are not modeled in the feedback school leaders provide for teachers, notably in the teacher evaluation process. While students are provided just-in-time feedback disaggregated by learning goals, teacher evaluations are conducted infrequently and often using a binary *meets* and *does not meet* communication system that does not promote professional growth (Weisberg et al., 2009). The purpose of this essay is to suggest schools should consider simultaneously implementing standards-based teacher evaluation models and SBG models in order to walk the talk of facilitating improvement of teachers and students within the school system.

STANDARDS-BASED TEACHER EVALUATION

In response to the federal government's 2011 Race to the Top (RTT) initiative, which tied federal funding to student achievement and teacher evaluation, some states have focused on value-added models (VAM) of teacher evaluation (McCaffrey et al., 2004; Wright et al., 2018). While the federal government's passage of the Every Student Succeeds Act (ESSA) in 2015 lessened the required elements of standardized assessment scores as related to teacher performance, it is still unclear whether today's teacher evaluation prioritizes accountability or teacher growth (Close et al., 2020; Close et al., 2019).

Standards-based teacher evaluation models are systems in which teacher performance is evaluated in accordance to a description of competencies within each standard (Heneman et al., 2006). Two recent literature reviews document tenets of standards-based teacher evaluation models to include utilizing rating scales and rubrics, which are likely to promote professional development and improved teacher performance (Clenchy, 2017; Steinberg & Donaldson, 2016). Common models, such as Danielson's Framework for Teaching (2007) and the Marzano Teacher Evaluation Model (Marzano & Toth, 2013), have attempted to establish evaluation instruments focused on standards of effective teaching and criteria based on continuums of performance (Borman & Kimball, 2005). Several teacher evaluation models have attempted to stimulate growth in classroom practitioners (Hamilton et al., 2008). Yet, school leaders and teachers may not have fully understood the benefit of these continuum-based models because of their past affiliation with compensation and accountability. According to Close et al. (2020), post-ESSA states' teacher evaluation models are trending towards observational and research-based component driven models utilized for formative purposes. Ross and Walsh (2019) indicated 41 states were utilizing standards-based teacher evaluation with more than one rating category, and 33 states required some elements of an improvement planning for less than satisfactory ratings.

Teacher evaluation should promote teacher growth, but in reality, concurrent evaluation and assessment purposes often distract from their intent of identifying pedagogical improvement (Williams & Hebert, 2020). Recent standards-based evaluation models have not substantially improved the utility of teacher evaluations (Holloway & Brass, 2018; Kraft & Gilmour, 2016). Therefore, inconsistencies in the interpretation of teacher evaluations demand a call for stronger teacher voices and redesigning teacher assessment and evaluation systems (Goe et al., 2008). Similarly, inconsistencies in grading have beckoned a call for reformed practices by experts in recent years (Brookhart et al., 2016). Schools aspiring to implement both standards-based teacher evaluation and standards-based grading are well leveraged to capitalize from the mutual benefits these models offer to the intended learner.

Standards-Based Grading Practices to Communicate Student Learning

Standards-based grading (SBG) is intended to communicate growth in student learning. Although definitions of SBG vary, Welsh (2019) suggests this philosophy of grading differs from previous grading practices in that student achievement is communicated using levels of learning such as *beginning*, *progressing*, *proficient*, and *exemplary*. Rather than using points or percentages, teachers communicate students' current levels of learning separately from information related to non-cognitive factors such as effort, homework completion, and class participation (Guskey, 2020; O'Connor, 2018). Teachers reference scales, often referred to as proficiency scales, to determine the extent to which students are performing at one of the school-determined levels of learning. As such, teachers must unpack the standards and create consensus on each level of student learning (Hoegh et al., 2020; Marzano & Haystead, 2008). Teachers create enhanced clarity for students and parents on the learning outcomes and success criteria and then communicate accordingly in the grade book or report card. Through looking in the grade book (or report card), students and parents are able to see learner strengths and areas of growth over time. As noted in Figure 1, assignments and assessments are communicated using points in a traditional grade book whereas in a standards-based grade book the learning students demonstrate on the assessment are disaggregated and communicated by level of learning.

National headlines describing increased student failure rates resulting from COVID-19 highlight the need for grading reform (Natanson, 2020; Strauss, 2020). Thomas Guskey questions whether these failure rates are a result of decreased learning or lower rates of compliance (Sawchuck, 2020). As school leaders continue to discern the purpose of grades, SBG models may become a solution for responding to the unique challenges presented by interrupted, remote, and hybrid learning (Townasley, 2020; 2021).

Figure 1

Sample Traditional Grade Book and Standards-based Grade Book Adapted from Townsley (2019)

| Traditional Grade Book | | Standards-Based Grade Book | |
|------------------------|--------------|---------------------------------------|--------------------------|
| <i>Assignment</i> | <i>Score</i> | <i>Standard</i> | <i>Level of Learning</i> |
| Worksheet 1.1 | 5/10 | Prove theorems about triangles | Proficient |
| Worksheet 1.2 | 9/10 | Use sine and cosine | Exemplary |
| Quiz 1.1-1.2 | 20/24 | Use volume formulas to solve problems | Progressing |
| Chapter 1 Test | 86/100 | Describe rotations of quadrilaterals | Proficient |
| Kleenex Extra Credit | 3/0 | Derive equations of parabolas | Beginning |

IMPLICATIONS FOR PRACTITIONERS: BI-DIRECTIONAL BENEFITS FOR SCHOOLS DEPLOYING SBG AND STANDARDS-BASED TEACHER EVALUATIONS

At first glance, the comprehensive criteria in a standards-based model for teacher evaluation appear straightforward; however, it is the tailoring of the model to local schools that is necessary for fidelity and growth (Marzano & Kendall, 1996). Common evaluation standard categories include: “professionalism,” “content knowledge,” “instructional planning and delivery,” “curriculum and assessment,” and “classroom environment.” The primary implication is that school districts should leverage their acute knowledge of SBG practices, which communicate clarity of learning goals for students, to maximize teacher growth via standards-based teacher evaluation systems; most notably by incorporating rubrics and scales. Marshall (2005) noted schools as early as the 2000s began to implement teacher evaluation rubrics and scales because they were more informative by communicating to teachers where they stand on a scale with a detailed description of what performance looks like at each level of proficiency. In the following paragraphs, we offer considerations related to teachers’ co-creation of proficiency scales, accelerating teachers’ self-reflection skills, and improved means of identifying teacher growth. We end by describing the benefits for practitioners in adopting this bi-directional approach.

Teachers Co-Creating Proficiency Scales

To assess teacher quality consistently and accurately through teacher evaluation systems, there must first be a shared definition of teacher effectiveness among teachers, school leaders, and district personnel (Danielson, 2011; Goe et al., 2008; Stiggins, 1997). Close et al. (2020) found, in a survey of two-thirds of U.S. state education department officials, that 67% identified “increased teacher input” as a major strength of post-ESSA teacher evaluation models. Accordingly, school leaders should involve teachers in co-creating proficiency scales integrated with elements from standards-based teacher evaluation models. First, schoolwide groups should agree on the design of the district scale template and when it is to be used. Second, small group teams should utilize backward design (Wiggins & McTighe, 1998) to develop consensus to the question, “What would a Proficient teacher look like if they demonstrated knowledge or performance of (a specific teacher evaluation standard) in our school/district?” In doing so, teachers co-create elements of what it means to be Distinguished in a teaching standard and document it on the proficiency scale. Furthermore, teachers brainstorm how being Distinguished is different from demonstrating Proficiency in the teaching standard. This unpacking process mirrors the SBG protocols teachers often facilitate to create proficiency scales of curricular standards for students.

Figure 2

Sample Standards-based Teacher Evaluation Using Proficiency Scale

| Smith Fork School District | |
|---|---|
| Teaching Standard: The teacher will assess all learning outcomes by designing and deploying assessments including formative assessment as a part of the instructional process. | |
| Level | Components |
| Distinguished | No deficiencies related to the <i>Proficient</i> elements below AND incorporates additional complex elements such as: <ul style="list-style-type: none"> ● Analyze student achievement data to justify modifications to curriculum, pacing, and intervention/enrichment. |
| Proficient | No deficiencies related to the <i>Minimally Proficient</i> elements below AND: <ul style="list-style-type: none"> ● Analyze how learning components are aligned to rigor of formative and summative assessment methodology. ● Create and evaluate appropriate assessment types that match the level of rigor identified in the learning component. ● Justify scoring criteria for <i>all</i> assessments to demonstrate what proficiency. ● Explain instructional adjustments made based on formative data. |
| Minimally Proficient | No deficiencies related to the <i>Not Yet Proficient</i> elements below AND: <ul style="list-style-type: none"> ● Identify <i>all</i> learning components within formative and summative assessments. ● Explain scoring criterion for formative and summative assessments. ● Examine formative assessment data with a collaborative team. ● Conduct formative assessment on <i>all</i> learning components. |
| Not Yet Proficient | Partial demonstration/knowledge of components at the <i>Minimally Proficient</i> AND <i>Proficient</i> levels |

Note. Adapted from Danielson (2007) and Marzano et al. (2017)..

Accelerate Teachers’ Self-Reflection

Using a tool in which teachers in a SBG environment are familiar with, such as a proficiency scale, may accelerate their ability to be self-reflective of teacher evaluation standards. Proficiency scales aim to provide a platform for increased

opportunities of embedded support and collaboration with peers and school leaders regarding effective teaching practices (Taylor & Tyler, 2012). School leaders should emphasize the use of teacher proficiency scales for self-assessment, collaborative teams, and instructional support personnel to foster robust conversations during the school year to supplement the teacher evaluation process. A sample proficiency scale adapted from Danielson’s (2007) Framework for Teaching evaluation model is outlined in Figure 2. Compare Figure 2 with Figure 3, a mock traditional binary evaluation instrument excerpt, regarding level of detail and descriptors for teacher improvement and effectiveness.

An Improved Means of Identifying Teacher Growth

Weisberg et al. (2009) found in one study when school leaders evaluated teachers utilizing binary measures of proficiency, school leaders rated teachers satisfactory compared to unsatisfactory up to 99% more often in evaluation categories. When compared to evaluation rubrics, proficiency scales serve as a pragmatic assessment tool because they are clear and customized to the school evaluation model. The traditional implementation of teacher evaluation relies heavily upon observations and artifacts; both of which are retrospective, and thus reduce the opportunity for just-in-time pedagogical adjustments. Darling-Hammond (2013) argued while standards-based teacher evaluations are critical in identifying teacher strengths and weaknesses, the evaluation itself does not facilitate teacher effectiveness. Consequently, rubrics and scales should be continuously reviewed during the school year rather than solely at the traditional end of the school-year evaluation timeframes (Biggers et al., 2012; Reibel & Thede, 2020). Co-creating common teacher proficiency scales from clear standards before and during the school year maximizes the chances for goal setting, reflection, and in-year adjustments prior to formal evaluation.

Proficiency scales, on the other hand, are more aptly utilized for formative functions of assessment of the teaching process as it occurs. Proficiency scales are written to establish clarity of learning targets and to derive feedback required to build upon what a practitioner demonstrates within a progression of growth versus traditional deficit identification models (Hoegh et al., 2020; Kluger & DeNisi, 1996). This, too, parallels the classroom use of proficiency scales in SBG environments where students are encouraged to track their own progress, performance, and goal-setting.

Figure 3

Sample Excerpt of Traditional Binary Evaluation Instrument

| | |
|---|---|
| Standard 1: Demonstrates ability to enhance academic performance and support for implementation of the school district’s student achievement goals | |
| Provide a narrative and additional documentation/artifact below to support MEETING or NOT MEETING the standard and criteria. | Circle or Highlight One: |
| | <p>Meets Standard</p> <p>Does Not Meet Standard</p> |

Note. Adapted from Iowa Model Educator Evaluation System (2019)

Walking the Talk.

Aligning standards-based teacher evaluation and SBG practices offers multifaceted benefits. First, co-creating proficiency scales in standards-based teacher evaluation and SBG classrooms enables educators to deepen pedagogical knowledge and adjust their instructional decisions while mitigating one of the largest challenges related to SBG implementation: the development of proficiency scales (Proulx et al., 2012). In doing so, teachers will become more fluent

in the district's expectations of instructional effectiveness. This collaborative approach reduces teacher stress and combats the ill-effects of a top-down approach associated with system change (Vogel 2012). Second, utilizing common standards-based tools, as both a teacher and adult-learner, increases the likelihood of transformative learning (Mezirow, 1997; Taylor, 2008). Short and Hirsh (2020) posit teachers need to experience curriculum, instruction, and resources to see how deployment benefits students rather than just hearing about it; often referred to as transformational learning. Even the most committed teachers need to experiment with ongoing trial and error pedagogy to come to grips with a compelling reason to improve or change without fear of evaluative repercussions (Elmore, 2008).

Third, teachers who have the opportunity to implement consistent use of SBG tools, such as a proficiency scale with their students, will have more self and collective efficacy as a staff on how those tools are used to assess their respective teacher effectiveness. For example, a science department deliberating what it means for a student to be "proficient" through the use of a proficiency scale of "demonstrating use of the scientific method" might simultaneously benefit from a discussion of what it means to be Proficient in a teacher evaluation standard, such as "demonstrate a safe classroom environment." If school leaders desire teachers as adult learners to self-reflect and demonstrate lasting growth, schools are obligated to make sure they understand proficiency of the expected standards. As with students, teachers too desire a common understanding and visual of communicating success (Stiggins, 1997).

Finally, the parallel congruency of standards-based teacher evaluation and SBG offers practical student and teacher social, motivational, and psychological benefits. When standards are clearer for teachers, it presents transferrable opportunities to enhance clarity for students. With the provision that proficiency scales are utilized bi-directionally, the practice of self-reflection becomes a shared practice for teacher and student; thus, fostering a commonality as an emphatic learning experience. Empathic pathways to learning are often reciprocal in nature, with the teacher feeling into the experience of student and student feeling into the experience of teacher (Feshbach & Feshbach, 2009). This highlights an idyllic practice for school leaders to maximize effective practices associated with learning versus teaching. The litany of research in recent years associating social-emotional learning (SEL) competencies to positive student outcomes, namely academic achievement (Durlak et al., 2011; Taylor et al., 2017), suggests schools leverage practices to enhance integration of SEL competencies. When educators and students alike feel comfortable in their local context, they are more likely to engage; thus, emphasizing the importance of social interaction in the learning environment (Demetriou, et al., 2009).

IMPLICATIONS FOR RESEARCHERS

The authors concur with Darling-Hammond (2014) that schools need "a conception of teacher evaluation as part of a teaching and learning system supporting continuous improvement, both for individual teachers and for the professional as a whole" (p. 5). Schools currently implementing both standards-based teacher evaluation models and SBG models offer researchers a starting point to understand the effects on educator and student mindsets. Andrade and Brookhart (2019) theorized that classroom assessment is the co-regulation of learning by teachers and students within their local context. In order to apply this theory and build upon the ideas presented in this paper, researchers might consider building upon previous investigations focused on the nexus of teacher mindset and classroom assessment (DeLuca et al., 2019) by exploring the relationship between teachers' mindset towards teacher evaluation and approaches to classroom grading. Furthermore, previous reviews of research suggest fidelity of implementation in assessing teacher evaluation continues to be a challenge (Hallinger et al., 2014). Similarly, Guskey and colleagues (2020) note secondary schools may not be yet implementing grading reform with fidelity. Additional investigation is needed to understand schools' fidelity of implementing teacher evaluation systems and grading reform before any meaningful correlations between the two can be determined. Yet, increased alignment between grading reform efforts and standards-based teacher evaluation presents an opportunity for teachers and school leaders alike to "walk the talk" of identifying and assessing growth versus mere evaluation.

REFERENCES

- Andrade, H. L. & Brookhart, S. M. (2019). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, 27(3), 1-23.
- Biggers, K., Croft, A., & Goe, L. (2012). Linking teacher evaluation to professional development: Focusing on improving teaching and learning. National Comprehensive Center for Teacher Quality. <http://www.gtlcenter.org/sites/default/files/docs/LinkingTeacherEval.pdf>
- Borman, G., & Kimball, S. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *Elementary School Journal*, 106(1), 3-20.

- Bradford, C., & Braaten, M. (2018). Teacher evaluation and the demoralization of teachers. *Teaching and Teacher Education, 75*, 49-59.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research, 86*, 803–848.
- Clenchy, K. R. (2017). *Teacher evaluation models: Compliance or growth?* [Doctoral dissertation, Northeastern University]. <https://repository.library.northeastern.edu/files/neu:cj82qm547/fulltext.pdf>
- Close, K., Amrein-Beardsley, A., & Collins, C. (2020). Putting teacher evaluation systems on the map: An overview of state's teacher evaluation systems post-Every Student Succeeds Act. *Education Policy Analysis Archives, 28*(58). <https://doi.org/10.14507/epaa.28.5252>
- Close, K., Amrein-Beardsley, A., & Collins, C. (2019). Mapping America's teacher evaluation plans under ESSA. *The Phi Delta Kappan, 101*(2), 22-26.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). ASCD.
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35-39.
- Darling-Hammond, L., Amrein-Beardsley, A., Heartel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan, 93*(6), 8-15.
- Darling-Hammond, L. (2013). *Getting teacher evaluation right*. Teachers College Press.
- Darling-Hammond, L. (2014). One piece of the whole: Teacher evaluation as part of a comprehensive system for teaching and learning. *American Educator, 38*(1), 4-13.
- DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019). Assessment mindset: Exploring the relationships between teacher mindset and approaches to classroom assessment. *Studies in Educational Evaluation, 61*, 159-169.
- Demetriou, H., Wilson, E., & Winterbottom, M. (2009). The role of emotion in teaching: Are there differences between male and female newly qualified teachers' approaches to teaching? *Educational Studies, 35*(4), 449-473.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.
- Elmore, R. F. (2008). *School reform from the inside out: Policy, practice, and performance*. Harvard Education Press.
- Feshbach, N. D., & Feshbach, S. (2009). *Empathy and education*. In J. Decety & W. Ickes (Eds.), *Social neuroscience. The social neuroscience of empathy* (p. 85–97). MIT Press.
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. <http://www.tqsource.org/publications/teacherEffectiveness.php>
- Guskey, T. R. (2015). *On your mark: Challenging the conventions of grading and reporting*. Solution Tree.
- Guskey, T. R. (2020). Breaking up the grade. *Educational Leadership, 78*(1), 40-46.
- Guskey, T. R., Townsley, M., & Buckmiller, T. (2020). The impact of standards-based learning: Tracking high school students' transition to the university. *NASSP Bulletin, 104*(4), 257-269.
- Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability, 26*(5), 5-28.
- Hamilton, L. S., Stecher, B. M., & Yuan, K. (2008). *Standards-based reform in the United States: History, research, and future directions*. Center on Education Policy, RAND.
- Heneman, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay* (CPRE Policy Briefs No. RB-45). Philadelphia: Consortium for Policy Research in Education. http://www.cpre.org/images/stories/cpre_pdfs/RB45.pdf
- Hoegh, J. K., Heflebower, T., & Warrick, P. B. (2020). *A handbook for developing & using proficiency scales in the classroom*. Marzano Resources.
- Holloway, J., & Brass, J. (2018). Making accountable teachers: The terrors and pleasures of performativity. *Journal of Education Policy, 33*(3), 361–382
- Iowa Model Educator Evaluation System. (2019). Iowa Department of Education. <https://educateiowa.gov/sites/files/ed/documents/IaMEES.pdf>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.
- Kraft, M. A., & Gilmour, A. F. (2016). Can principals promote teacher development as evaluators? A case study of principals' views and experiences. *Educational Administration Quarterly, 52*(5), 711–753.

- Knight, M. & Cooper, R. (2019). Taking on a new grading system: The interconnected effects of standards-based grading on teaching, learning, assessment, and student behavior. *NASSP Bulletin*, 103(1), 65-92.
- Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86(10), 727-735.
- Marzano, R. J., & Kendall, J. S. (1996). *A comprehensive guide to designing standards-based districts, schools, and classrooms*. ASCD.
- Marzano, R. J. & Haystead, M. W. (2008). *Making standards useful in the classroom*. ASCD.
- Marzano, R. J., Norford, J. S., Finn, M., & Finn, D., III (2017). *A handbook for personalized competency-based education*. Marzano Resources.
- Marzano, R. J., & Toth, M. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. ASCD.
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14-19.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29(1), 67-101.
- Mezirow, J. (1997). Transformative learning: Theory to practice. In P. Cranton (Ed.), *Transformative learning in action: Insights from practice. New directions for adult and continuing education*, 74, (pp. 5-12). Jossey-Bass.
- Natanson, H. (2020, November 24). Failing grades spike in Virginia's largest school system as online learning gap emerges nationwide. *The Washington Post*. https://www.washingtonpost.com/local/education/fairfax-schools-more-failing-grades/2020/11/24/1ac2412e-2e34-11eb-96c2-aac3f162215d_story.html
- O'Connor, K. (2018). *How to grade for learning, K-12 (4th ed.)*. Corwin Press
- Paufler, N. A. (2018). Declining morale, diminishing autonomy, and decreasing value: Principal reflections on a high stakes teacher evaluation system. *International Journal of Education Policy & Leadership*, 13(8), 1-15.
- Proulx, C., Spencer-May, K., Westerberg, T. (2012). Moving to standards-based grading: Lessons from Omaha. *Principal Leadership*, 13(4), 30-34.
- Reibel, A. R., & Thede, M. (2020). *Small changes, big impact: Ten strategies to promote student efficacy and lifelong learning*. Solution Tree Press.
- Ross, E., & Walsh, K. (2019). *State of the States 2019: Teacher and Principal Evaluation Policy*. National Council on Teacher Quality.
- Sawchuck, S. (2020, December 11). Should schools be giving so many failing grades this year? *Education Week*. <https://www.edweek.org/leadership/should-schools-be-giving-so-many-failing-grades-this-year/2020/12>
- Short, J., & Hirsh, S. (2020). *The elements: Transforming teaching through curriculum-based professional learning*. Carnegie Corporation.
- Steinberg, M., & Donaldson, M. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy*, 11, 340-359.
- Stiggins, R. J. (1986). Teacher evaluation: Accountability and growth systems - different purposes. *NASSP Bulletin*, 70(490), 51-58.
- Stiggins, R. J. (1997). *Student-centered classroom assessment (2nd ed.)*. Prentice-Hall.
- Strauss, D. (2020, December 6). More students than ever got F's in first term of 2020-21 school year - but are A-F grades fair in a pandemic? *The Washington Post*. <https://www.washingtonpost.com/education/2020/12/06/more-students-than-ever-got-fs-first-term-2020-21-school-year-are-a-f-grades-fair-pandemic/>
- Taylor, E. W. (2008). Transformative learning theory. *New Directions for Adult and Continuing Education*, 119, 5-15.
- Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *The American Economic Review*, 102(7), 3628-3651.
- Taylor, R. D., Oberle, E., Durlak, J. A., Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, 88, 1156-1171.
- Townsley, M. (2019). Considering standards-based grading: Challenges for secondary school leaders. *Journal of School Administration Research and Development*, 4(1), 35-38.
- Townsley, M. (2020). Grading principles in pandemic-era learning: Recommendations and implications for secondary school leaders. *Journal of School Administration Research and Development*, 5(S1), 8-14.
- Townsley, M. (2021). Grading in the midst of a pandemic. *School Administrator*, 78(5), 28-31.
- Townsley, M., Buckmiller, T., & Cooper, R. (2019). Anticipating a second wave of standards-based grading implementation and understanding the potential barriers: Perceptions of high school principals. *NASSP Bulletin*, 103(4), 281-299.

- Vogel, L. R. (2012). A leader's journey: The challenges of implementing standards-based student grouping. *NASSP Bulletin*, 96, 323-349.
- Weems, D. M. & Rogers, C. B. (2010). Are US teachers making the grade? A proposed framework for teacher evaluation and professional growth. *Management in Education*, 24(1), 19-24.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. The New Teacher Project.
- Welsh, M. (2019). Standards-based grading. In T. R. Guskey and S. M. Brookhart (Eds.), *What do we know about grading: What works, what doesn't, and what's next* (pp. 113-144). ASCD.
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. ASCD.
- Williams, K., & Hebert, D. (2020). Teacher evaluation systems: A literature review on issues and impact. *Research Issues in Contemporary Education*, 5(1), 42-50.
- Wind, S. A., Tsai, C., Grajeda, S. B., & Bergin, C. (2018). Principals' use of rating scale categories in classroom observations for teacher evaluation. *School Effectiveness and School Improvement*, 29(3), 1-26.
- Wright, K. B., Shields, S. M., Black, K., Banerjee, M., & Waxman, H. C. (2018). Teacher perceptions of influence, autonomy, and satisfaction in the early Race to the Top era. *Education Policy Analysis Archives*, 26(62). <http://dx.doi.org/10.14507/epaa.26.3449>

CHAD LANG, EdD, is the Director of School Improvement & HR at Glenwood Community School District in Glenwood, Iowa. His major research interests are grading and assessment, educational leadership, teacher recruitment & retention, and extracurricular school activities. Email: chadmichaellang@gmail.com

MATT TOWNSLEY, EdD, is an Assistant Professor of Educational Leadership in the College of Education, University of Northern Iowa. His major research interests lie in the areas of educational leadership, assessment and grading. Email: matt.townsley@uni.edu

Manuscript submitted: January 8, 2021

Manuscript revised: May 7, 2021

Accepted for publication: June 25, 2021