Leading Leaders in Rethinking Grading: A Case Study of Implementation of Standards-Based Grading in Educational Leadership

Erin E. Lehmann
University of South Dakota

Follow this and additional works at: https://digitalcommons.uncfsu.edu/jri

Part of the Adult and Continuing Education Administration Commons, Community College Leadership Commons, Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, Educational Leadership Commons, Higher Education Commons, Higher Education Administration Commons, and the Higher Education and Teaching Commons

Recommended Citation
Available at: https://digitalcommons.uncfsu.edu/jri/vol5/iss3/11

This Best Practice is brought to you for free and open access by the Journal of Research Initiatives at DigitalCommons@Fayetteville State University. It has been accepted for inclusion in Journal of Research Initiatives by an authorized editor of DigitalCommons@Fayetteville State University. For more information, please contact dballar5@uncfsu.edu.
Leading Leaders in Rethinking Grading: A Case Study of Implementation of Standards-Based Grading in Educational Leadership

About the Author(s)
Erin Lehmann is an assistant professor in the division of educational leadership at the University of South Dakota.

Keywords
educational leadership, assessment, standards-based grading, grading

Cover Page Footnote
I would like to thank Drs. Susan Curtin, David De Jong, and Karen Card for our bi-weekly accountability chats and thoughtful dialogue. Also, a huge thank you to Constance Krueger, you truly are a gift and the best editor I know.

This best practice is available in Journal of Research Initiatives: https://digitalcommons.uncfsu.edu/jri/vol5/iss3/11
LEADING LEADERS IN RETHINKING GRADING: A CASE STUDY OF IMPLEMENTATION OF STANDARDS-BASED GRADING IN EDUCATIONAL LEADERSHIP

Erin E. Lehmann, University of South Dakota

Abstract
The purpose of this paper is to share the process of how one university instructor worked toward a shift to standards-based grading (SBG) in a graduate Educational Leadership program. Educational leadership programs use standards to guide coursework and instruction in an accountability era, but grading practices remain as subjective as they were 50 years ago. Educators of future leaders must address this need. In addition, instructors need to effectively communicate essential learning to students to understand their learning progression clearly; standards-based grading is designed to do this. The author shares best practices in grading as well as the challenges of implementation of standards-based grading.

Introduction
Significance of Standards and Grading
The recent national emphasis on accountability in K-12 education throws a focus on accountability in assessment. Assessment literacy has a new place in post-secondary education (Popham, 2018). With the progressive works of Guskey (2019), O’Connor (2017), Brookhart (2017), and Womeli (2018), the attention on grading what students learn in a PK-12 setting is gaining momentum. Because the role of instructional leadership is ever-changing, school leaders no longer serve primarily as supervisors but are now required to be redesigners of their schools and their school systems (Levine, 2005). This means that grading reform is no longer a theory to consider but a best practice to implement in teaching and learning. School leaders are charged with guiding this focus. It is up to educational leadership programs to prepare future leaders for this change. The National Educational Leadership Preparation (NELP) standards document has been an ever fluid and responsive collection of best practices. The most recent changes provided the perfect "why" and opportunity to rethink our focus on student learning outcomes.

Background
During the 20th century, traditional grading played two key roles dividing the student body into hierarchical ranks and measuring student learning in relation to the curricula instead of standards (Guskey, 2015; Tocci, 2010). Today’s traditional grading practices encompass a range of possibilities: participation, projects, homework, all of which are potentially subjective, along with exams that may or may not be more objective. Investigating these grading practices reveals the underlying values of the humans who arbitrarily create them. For example, does the instructor believe it is okay just to read about the dissection of a frog or does the instructor value the actual process of dissecting a frog? Does the instructor who wrote the textbook for a graduate-level course value other authors’ interpretation of the facts, or does the instructor place a disproportional value on the words they authored? In theory, they may espouse different authors, but in practice, is their grading system mismatched? Is their practice aligned to sound...
educational theory, or is there a disconnect (Levine, 2005; Link, 2019)? Is this what they truly want to emphasize? Blodgett (2017) claims that “to select a grading system is to select a certain set of values about teaching and learning” (p. 1).

Educational leadership courses have problems with subjective grading practices. The values of the instructors’ surface in these courses, too. The instructor may value solely reading about supervision and evaluation in an educational leadership course. However, the instructor realizes that it is the standard itself, which asks for the leader “to evaluate, develop, and implement [emphasis added] high-quality and equitable academic and non-academic instructional practices, resources, technologies, and services that support equity, digital literacy, and the school’s academic and non-academic systems” (National Educational Leadership Preparation, 2018, p. 18). Surely these future leaders need to be implementing, not just reading, the specific duties of their profession. The entire standard needs to be addressed, not just one component. Teaching is a complicated art form that requires multiple ways of responding to students. Instructors have the academic freedom to decide where and when to emphasize the learning. Instructors have substantial latitude in determining how to teach the courses for which they are responsible.

Academic freedom is a basic tenet of higher education, and instructors need to be assured their autonomy will continue. However, grading practices, which are often part of an ingrained culture and are usually performed without question, are more effective when directly aligned to the standards. SBG provides accurate feedback with clarity. The student does not have to resort to interpreting an individual instructor’s academic values (Lynch & Hennessy, 2012). The additional challenge for educational leadership graduate students is that they are typically engaged in traditional grading models in their school systems. When we ask graduate students to shift their thinking from an externally determined grade to accepting responsibility for their learning outcomes, it challenges their deeply held beliefs and understanding of the institutional culture of higher education.

Beyond the course level, higher education institutions have become committed to a more effective assessment system or policy (Sadler, 2005). A system communicating clear feedback promotes student learning. It is then vital for institutions to create policies and procedures to establish and maintain the meaning and worth of grades (Sadler, 2010). To ensure consistency, departments often generate learning outcomes from standards during assessment (Bloxham & Boyd, 2012). Universities are then also accountable to students for the standards they employ. A broad consistency across courses is desirable and achievable, especially when entire programs align the coursework to the standards (Sadler, 2010).

Levine (2005) research described how students enrolled in an educational leadership program found a lack of standards and jumping through hoops problematic. The area of assessing student mastery and grading in higher education is surrounded by many controversies, including what criteria to grade (Adinde, 2020). A significant problem with a traditional grading system is that they serve many purposes and represent a mixture of teachers’ beliefs (Brookhart et al., 2016; Guskey, 2015; O’Connor, 2017). Problems arise when school leadership programs do not engage in systematic self-assessment (Levine, 2005), integral to standards-based grading. Although 15 years old, Levine’s research has influenced countless educational leadership departments to evaluate their programs and rethink student performance. The report shared how educational administration programs are the weakest of all the programs, and most of the programs ranged from inadequate to appalling. The study concluded that students jumping
through hoops is a prevalent issue. Without self-assessment of systemic systems, educational leadership programs fail to communicate where students are falling behind and if there are gaps and holes in the program itself. Levine's research serves as a wake-up call to educational leadership departments. Improvement science can be a helpful framework for the cultural and systemic changes needed to shift to a standards-based grading mindset.

**Theoretical Framework**

These important decisions could be developed, revised, and fine-tuned (Bryk et al., 2017). Administrators need current and up-to-date best practices when leading a school. Unfortunately, they do not always take the time to seek evidence-based research to help inform meaningful leadership decisions actively. With the current "research-to-practice disconnect," improvement science works to ensure that the research is useable (Cobb, 2019, p. 289), which means leaders can make immediately improved adjustments, including leaders in higher education.

Improvement science is not a one-size-fits-all model. It is a methodological framework that helps leaders define problems, implement changes, and determine whether those changes improve teaching and learning (Crow et al., 2019). If learning-by-doing is a way to prepare future leaders, educational leadership programs can model students to engage, explain, lead, and generalize their learning. These programs are modeling the way for their future leaders.

The purpose of this paper is to share how standards-based grading could enhance post-secondary educational leadership courses by implementing improvement science principles. For example, suppose educational leadership programs prepare future leaders to lead best practices in teaching and learning. In that case, students must understand why and how we assess student learning by utilizing grading practices aligned to essential standards relevant to their future work. A further complication is a gap in literature addressing the grading practices in higher education. While the research is beginning to emerge (Sadler, 2017), the depth is not there. Instead, the emphasis has been connecting grades to instructor performance (Hu, 2005; Remedios & Lieberman, 2008; Smith & Fleisher, 2011; Svanum & Aigner, 2011); however, it is important to note that standards-based grading has been all but neglected in higher education (Buckmiller et al., 2017). Therefore, the goals of this paper are to share best practices in embedding SBG into educational leadership courses using an improvement science framework.

**Standards-Based Grading Rationale**

The first step in improving science is to establish the why. Our program used the 5 Why's to identify the need to explore standards-based grading. Standards-based grading is an intentional way to focus on student learning based on learning outcomes or standards rather than the accumulation of points, which many traditional grading systems use. For example, if an instructor gives 23 out of 25 points on an assignment with no additional comments, the student is left wondering where the two points were lost. When the instructor identifies the essential standards or learning outcomes for the course, communicates the essential learning, connects the learning with the standards, and measures the connection, standards-based grading is in effect (Sadler, 2005). Transparency is foundational when standards-based grading is in use and the learning outcomes are objective and measurable. If students are to be treated fairly, they need to know what criteria will be used to grade the quality of their work (Sadler, 2005). As one educational leader in our department shares, "The advantage of standards-based grading is that it provides in-time feedback which can inform to what extent the learner has mastered the content and presentation of content" (S. Curtin, personal communication, July 13, 2020).
Hack (2015) expressed the need for greater clarity in the assessment process, not only for students but also for the scorers. When essential standards are identified, instructors can be more strategic when targeting the standards and knowing when and where to emphasize the standard (Toledo & Dubas, 2017). This is a checks-and-balance system for program improvement and redesign. In truth, institutions should be able to vouch for the quality of their degrees, and a consistent way to do this is to align assignments to the standards (Sadler, 2017).

Standards-based grading intends to improve student outcomes by modifying how instructors communicate essential learning and restructuring how students demonstrate mastery towards a standard (Hanover, 2015). Traditional letter grades record a single grade for a course and seldom include detailed information regarding student progress and learning. In contrast, standards-based grading communicates more detailed information about the student learning progress towards mastery of each essential standard. When implementing SBG into a technology course, Buckmiller et al. (2017) found that by using standards as a guide for reporting student achievement in the technology classroom studied, instructors could provide students with better responses to queries about their grades mean and how to improve. As a result, students jumped through fewer hoops and gained more substance. S. Curtin continued with her reflection (personal communication, July 13, 2020):

SBG could shift the learner's focus to WHAT is learned/mastered/not learned/not mastered rather than trying to fulfill an instructor's expectations to earn a grade. I would be interested to see how this shift can create a true focus on the learning process. If we did this right, we would align the entire program curriculum so that the learner had an opportunity to master skills later that were not mastered initially--much like Jerome Bruner's spiral curriculum. We would also be able to show them how we have done that, which would enrich their understanding of curriculum development and their role as instructional leaders.

Grading Practices in Higher Education

The practice of standards-based grading (SBG) is not nearly as common in higher education, nor is training for future leaders (Beatty, 2013; Buckmiller et al., 2017; Link, 2019). Nevertheless, deeply engrained traditions prevail when it comes to grading in higher education. D. De Jong, the department chair of the educational leadership division, shares (personal communication, July 16, 2020):

Traditional grading is valuable because it is known. However, I struggle with two aspects of traditional grading in higher education. First, I struggle with whether to deduct points for late work. It is an effective practice, but should a school leader earn a B if they turned in perfect work, yet it was late? I don't think so. Second, I struggle with whether to deduct points for incorrect grammar and APA. I have landed on grammar and APA, counting for about 20% of each written assignment.

Historical practices almost become ritualistic, and change is unlikely to occur when not reflecting on current practices. For example, the dilution of grading standards in higher education has been a concern for the last 30 years (Foster, 2016). As a result, educational leadership programs must align the coursework content to academic standards (Levine, 2005).

Measuring student learning can become quite complicated. If an instructor assesses memorization of facts and rote recall of details, grades can easily measure this low-level learning. Higher-order thinking, such as the knowledge, skills, and dispositions of educational leadership, requires analysis, synthesis, and judgment and is much more difficult to assess;
however, these are the very areas in which future leaders need to excel. Therefore, leadership courses should be preparing candidates for roles that require complex thinking and leading (Blodgett, 2017; Weimer, 2002). Unfortunately, research on undergraduate grading practices reveals that college students are often confronted with grading practices that reflect subjective, non-standardized formats incorporating a blend of academic and non-academic components (Buckmiller et al., 2017). Because of these inconsistencies with grading, students must aim at a moving target with the hopes of meeting each instructor's instructional agenda. One way to eliminate misaligned grading practices is to ensure the purpose of grades and grading remains clear (Link, 2019). Feedback on essential standards and student learning serves this purpose: clear communication about reaching a target that does not move and the time and space to reach the target rather than one opportunity to master knowledge or skills.

**Standards-Based Grading in Postsecondary Education**

Universal scales and standardized units for measuring academic achievement do not exist in higher education (Sadler, 2010). Often instructors create their units of measure (i.e., 90, 80, 70, 60). Some instructors use a stricter scale such as 94, 88, 82, 76, thinking this increases the rigor; however, adjusting the grading scale into smaller increments does not increase the depth of the learning. This then becomes a systemic issue. One way to ensure consistency is through standards-based grading. Although SBG is not readily evident across the country, implementation is emerging in some college courses. Dr. Kruse's educational technology course employed SBG to provide rich, formative feedback to students if they met or did not meet the standards (Buckmiller et al., 2017). Dr. Kruse provided flexibility with assignments and asked students to write a reflection on why they were able or not able to meet the standards of each project. Once students met the standard, they did not have to continue to demonstrate mastery repeatedly, but they did have to demonstrate an understanding of why they met that standard. This is an example of how standards provide a clear picture of the learner instead of the vague, open interpretation based on an average of points.

Beatty (2013) incorporated SBG into an introductory physics course. He first identified specific skills or competencies that could be articulated, understood, and assessed. Beatty (2013) found using SBG with a four-point mastery scale a more straightforward method for grading, and it was much more accurate. In addition, a sound set of standards make grading relatively fast, easy, and communicative. Beatty (2013) shared a few helpful suggestions: develop assessments with standards, identify the standards on each assessment, keep re-assessment efficient, and finally, organize standards around instruction.

Among the many perspectives on grading in higher education is one that questions the use of grading at all. Some researchers believe grading should not exist in an environment focusing on creating independent, analytical, and innovative thinkers. The critics believe grading in higher education undermines intrinsic motivation and promotes passivity, obedience, and submissiveness in students (Tannock, 2017; Tedesco, 2011). On the other hand, it might be the instructor's method for grading, which is indeed the problem. SBG takes an iterative approach when aligning assessment, standards, and instruction. Eliminating grades might be radical or unrealistic (Tannock, 2017). Suppose grading practices in higher education are driven by educational goals such as providing feedback to students, motivating students, and measuring learning. In that case, instructors need to minimize the negative impact of grades (Schinske & Tanner, 2014). In Blodgett's (2017) Theology course, she had to take a step back and examine what exactly she was grading. Instead of focusing on an accumulation of points, she decided to
focus on learning objective achievement. By doing this, her grading system greatly clarified expectations among teachers and learners.

**Standards-Based Grading in Educational Leadership**

Link (2019) adds a critical but frequently overlooked dimension of instructional leadership: grading and reporting. When preparing future leaders in this era of accountability, educational leadership programs must prepare their leaders to align theory and practice with assessment (Levine, 2005; Link, 2019). Townsley (2019) implemented SBG into his educational leadership program and, more specifically, his curriculum leadership course. He provided the implementation steps as well as key considerations for embedding SBG in an educational leadership course. With over ten years of experience in PK-12 SBG, district leadership, and higher education leadership courses, Townsley’s vast experience lends credibility to implementation. Zimmerman (2017) incorporated SBG into his introductory physics course and provided these logical steps to help with implementation:

1. Write learning objectives
2. Determine how to assign grades
3. Keep track of learning objectives
4. Determine the form of assessments
5. Determine the logistics of re-assessment
6. Take a deep breath and jump in
7. Explain things to students early and often

Townsley (2019) summarized the role of educational leadership where professors will “walk the talk of providing quality feedback to learners by embedding standards-based grading in school leader preparation courses” (p. 7). With this in mind, this author decided to jump in and walk the talk.

**Walking the Talk**

Using Zimmerman and Townsley’s implementation tips, the author's priority was to identify the essential standards for a supervision course in educational leadership. Accredited universities across the nation are required to evaluate their leadership programs based on the NELP standards. Once the priority standards were identified in the crosswalk, the intricate work to develop standards-based assignments occurred. This was a key step, and it is essential to note that this task should not be completed in reverse (aligning assignments to standards). This exercise was time-consuming; however, it kept each assignment aligned to each standard.

Understanding the depth of knowledge or essential learning of standards guided the assignment alignment. It should also be noted that having multiple steps within an assignment created preciseness to the assignment itself and the alignment to the standards. For example, at times, Step 1 in an assignment was the foundational learning of the standard, which led to Step 2, the application component.

A natural next action was to create a rubric for each of the assignments. A well-designed rubric communicates enough information, so students know the criteria for the assignment and the description of the standards. Multiple conversations ensued to determine how many levels of understanding to include. As a team, a decision was made to contain the following levels of understanding: Does Not Meet the Standard, Approaching Standard, Meets standard, and Exceeds standard. Using these levels of understanding in the rubrics, the instructors agreed to
accept Mastery of the Standard as the Meets Standard level. Anything in the Exceeds Standards was a bonus of learning and application.

The use of rubrics for grading and feedback in higher education has increased in response to requirements for consistency and transparency (Akinde, 2020; Hack, 2015). Because of the strong crosswalk work, the educational leadership division created rubrics to align with the priority standards. Also, adult learners tend to excel and produce high-quality work when utilizing a well-aligned rubric (Reddy & Andrade, 2010). Finessing rubric language without losing the integrity of the standard is tedious but necessary work. It is easy to get lost in the language or complexity of the standard, and the rubric needs to provide clarification. For example, figure 1 shows how Step 1 of an assignment could be broken into multiple components and aligned to standards. A glance at the rubric by a student who earned an "Approaching Standard" mark will quickly reveal the elements that might be missing. Perhaps, they thought they included all the requirements but discovered they omitted a detailed and specific plan. This targeted feedback is accurate and effective; the student knows what to do to meet the standard.

Figure 1
Sample assignment aligned to standards (can be broken into steps)

<table>
<thead>
<tr>
<th>Section/Standard</th>
<th>Does not meet standard</th>
<th>Approaching standard</th>
<th>Meets standard</th>
<th>**Exceeds standard</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>NELP Component</em></td>
<td>Candidate does not adhere to the instructions for the assignment or does not submit this portion of the assignment.</td>
<td>School PDP does not reflect a deep understanding of building level instructional improvement as is evidenced by a vague or underdeveloped targeted professional development plan which is not based upon analysis of informal and formal evaluative data; assessment of the school improvement plan; and exploration of school-wide assessment data for all students in the building; the plan is not detailed, specific, does not address the needs of faculty/staff as well as students. Language is not professional and has 4 or more grammatical and APA 7 errors.</td>
<td>School PDP reflects an understanding of building level instructional improvement as is evidenced by a targeted professional development plan based upon analysis of informal and formal evaluative data; assessment of the school improvement plan; and exploration of school-wide assessment data for all students in the building; the plan is detailed, specific, and addresses the needs of faculty/staff as well as students. Language is professional, almost free from grammatical errors (1-3) and APA 7 compliant.</td>
<td>School PDP reflects an understanding of building level instructional improvement as is evidenced by a targeted professional development plan based upon analysis of informal and formal evaluative data, assessment of the school improvement plan, and exploration of school-wide assessment data for all students in the building; the plan is detailed, specific, and addresses the needs of faculty/staff as well as students; language is professional, free from grammatical errors and APA 7 compliant.</td>
</tr>
</tbody>
</table>

Each assignment requires this thoughtful work and alignment to standards. This is not an easy task and requires ample instructor time, especially in the initial stages. Once assignments are
created, a standards form is constructed for students. The last step is to decide how to calculate final grades.

Even with the switch to SBG, the instructor could not submit a proficiency level as a final grade for the course but was required to submit end-of-course letter grades for each student. Fortunately, Townsley (2019) provided a clear picture of converting outcome achievement into a letter grade, so this author followed his direction, as shown in Figure 2. For example, his educational leadership course covered six essential standards, and if the student demonstrated mastery of those six standards, they would receive an A.

**Figure 2**
Grading criteria

Final grades will be calculated as follows:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Evidence of “Meeting” or “Exceeding” on all 6 NELP Standards (NELP 7.2, 7.3, 7.4, 1.2, 2.1, and 5.3)</td>
</tr>
<tr>
<td>B</td>
<td>Evidence of “Meeting” or “Exceeding” on 4-5 NELP Standards (NELP 7.2, 7.3, 7.4, 1.2, 2.1, and 5.3)</td>
</tr>
<tr>
<td>C</td>
<td>Evidence of “Meeting” or “Exceeding” on 3 NELP Standards (NELP 7.2, 7.3, 7.4, 1.2, 2.1, and 5.3)</td>
</tr>
<tr>
<td>F</td>
<td>Evidence of “Meeting” or “Exceeding” on 0-2 NELP Standards (NELP 7.2, 7.3, 7.4, 1.2, 2.1, and 5.3)</td>
</tr>
</tbody>
</table>

**A Shift from a Point System to a Feedback System**

Any accredited educational leadership program that applies for national recognition has created a crosswalk for standards and courses. However, using standards aligned to rubrics is a new way of thinking about assessment instead of only creating a crosswalk to meet the requirements for accreditation. The shift is in the mindset of moving from a traditional point system to a standards-driven one: learning matters more than the accumulation of points, one where feedback informs learning and instruction. This type of feedback, based on standards, is an essential step towards transparency and clarity (Hack, 2015). Educational leadership candidates will more than likely be new to this type of grading system, so dedicated and thorough communication will be key to the success of the implementation. Also, educational leadership programs will be better prepared to demonstrate to what extent their candidates meet the standards of the field and reveal their capacity for instructional leadership.

One of the first shifts that must occur is philosophical. Beginning with an exploration of best practices, literature, successes, and failures, the instructor can begin to see the difference between traditional and standards-based grading. With this new knowledge, the instructor understands new ways to measure the success of attaining standard proficiency. The importance of feedback becomes crucial to students’ achieving success. Standards-based grading with its emphasis on feedback begins to make sense. The traditional point system comes into question. The philosophical shift can happen. Standards-based grading becomes a viable way to communicate student success. Creating rubrics aligned to the standards is the next step. Points no longer seem to feel obsolete. Reflective practice is crucial to evaluate the type of feedback.
students require based on the standards. All that needs to be done at this point is to walk the talk and give standards-based grading a chance.

**Conclusion**

The goal of standards-based grading in educational leadership courses is to provide accurate and effective feedback to candidates about their learning and guide their development as instructional leaders responsible for every student's learning. When students have this access to their learning, they have a clearer understanding of how to master the field standards and thus are better prepared to create/sustain an aligned PreK-12 assessment system. Black and Wiliam (1998) have shared the importance and influence of effective feedback on the entire learning process. Through targeted feedback regarding the attainment of each specific standard, standards-based grading provides a more transparent assessment of strengths and weaknesses to students (Toledo & Dubas, 2017). In leading leaders to rethink grading practices, higher education leadership departments can affect a needed shift toward a standards-based grading system, one that is grounded solidly upon national standards and best practices. This shift toward effective feedback, through aligned grading practices, will enrich student learning.
References
Popham, W. J. (2018). Assessment literacy for educators in a hurry. ASCD.
Sadler, D.R. (2005) Interpretations of criteria-based assessment and grading in higher education,