Executive Summary

Teachers' interpretations of standards-based grading and reporting: Learning a new language and culture

Sara Hagen
University of Wisconsin-Madison
April 21, 2009

Introduction

In September 2008, the Madison Metropolitan School District (MMSD) began implementation of standards-based grading and reporting in all district middle schools. This paper is a study of teachers' responses to the new system of grading, based on interviews and surveys conducted during and after the first quarter of the 2008-09 school year. I seek to provide a clear understanding of both the intentions behind the implementation and its early effects on teachers' beliefs and actions, an understanding which may facilitate further implementation within the district and have possible implications for other districts considering a transition to standards-based grading. This document is a summary of the much longer report. For elaboration, please refer to the indicated sections of the main document.

Two main purposes of standards-based grading [page 2]

In standards-based grading, academic achievement is separated from behavior, effort, and progress. Students' achievement is evaluated with respect to a grade-level standard. Instead of a single grade for a course, teachers report on multiple standards to give a more accurate picture of what students know and are able to do. Teachers and students focus on meeting those standards. The two main purposes of standards-based grading are improved accountability for student achievement and improved learning.

1) **Improved accountability**: As teachers and students develop a shared understanding of standards and criteria, grading loses its mystery and gains a direct relationship to learning. By expanding that shared understanding throughout the school and community, the system of standards-based grading becomes transparent and fair, a force for accountability.

2) **Improved learning**: Standards-based grading improves teaching and learning through improved feedback to teachers and students and a new approach to student motivation. Teachers use knowledge of how students are doing with respect to the standards to plan further instruction. Students likewise use feedback to adjust their strategies and improve their work.

Conceptualizing the transition as a process of language learning [page 5]

During the implementation of standards-based grading, teachers explained that the new grades don't feel right—something is being lost in translation from the traditional system to the standards-based system. In this paper, I examine the ways in which learning standards-based grading is like learning a new language and culture.

Methods and organization [page 6]

The following analysis is based on observations of the district's Summer Institute on Understanding by Design, two interviews each with nine teachers in two different middle schools, and responses to a self-administered questionnaire sent to a representative sample of all district middle school teachers of Language Arts, Math, Science, Social Studies, and World Languages. [See page 127 for a thorough explanation of research procedures and statistical techniques.]

Chapter 1: A New Culture of Standards-Based Grading

Grading and assessment practices are cultural practices. These practices feel right and make sense to teachers, students, and parents; yet, these practices are often unexamined, as are the beliefs that underlie those practices. In this section I examine the relationship between standards-based perspectives on grading and standards-based practices of grading in four areas: 1) beliefs about learning, 2) beliefs about students, 3) beliefs about the role of teachers, and 4) beliefs about accountability. Increased teacher consensus around these four areas should lead to new practices of planning, instruction, assessment, and grading. It is therefore important that teachers have time not just to implement new grading practices, but to reflect on the beliefs that support a system of standards-based grading.

Standards-Based Perspectives	Standards-Based Practices	Considerations and Findings of Interest
Learning: • Students do not receive knowledge (transmission model). • Students construct knowledge (constructivism). [page 9]	• Lesson plans are based on standards (key concepts). • Grade books are arranged by standard instead of assignment type (homework, quizzes, etc.)	• MMSD teachers described a focus on standards as the largest benefit of standards-based grading. • Some standards have been revised recently. • Teachers found more specific math power standards easier to understand than general standards in other areas. • 89% of teachers have changed their grade book setup as of December 2008.
Students: • All students can learn. Bell curves do not apply. Intelligence is not a fixed construct. • Students learn in different ways. [page 13]	• Provide multiple opportunities for success. • Use formative assessment results to adjust instruction. • Value mistakes as natural and helpful; encourage student risk-taking on challenging tasks.	• Consider clarifying retake policies. • May require reallocation of resources to students who need more time to learn. • Scores of 1, 2, 3, 4 alone provide insufficient feedback to students— must be combined with additional information.

Standards-Based Perspectives	Standards-Based Practices	Considerations and Findings of Interest
Teachers: • Homework is part of a formative assessment system. • Homework grades cannot be used as rewards or punishments. • Teachers have a responsibility to help students develop agency, a sense of control over their learning. [page 20]	• Learning skills (including homework completion) are reported separately from achievement. • Consistently help students to make the connection between homework and achievement. • Do not use zeroes or group grades. • Use student-centered assessment practices. • Provide all students with the opportunity for a 4 through differentiated assignments.	• Teachers and parents have similar beliefs about homework, commonly viewing it as currency—payment for effort expended. 66% of parents surveyed by the MMSD in Spring 2008 did not support separation of homework from the grade. • Students at this age may not have made the connection between homework and achievement. • Although group grades are discouraged, group work is still beneficial for students. • Student-centered assessment includes creating criteria with students, involving students in record-keeping and parent/community communication. • Many assignments must be revised or rethought. Not all assignments have a "4" level as currently written.
Accountability: • Being accountable means providing accurate grades, opportunities, fair treatment, and high expectations for all students. [page 37]	• Communicate clearly with parents and community. • Develop school cultures that promote collaboration within a systematic framework of integrated professional development.	• Teachers' professional judgment is more reliable and valid than external test results. Classroom assessment should be part of the official accountability system. • Modeling from colleagues is highly influential. • Teachers need time for both individual reflection and extensive collaboration.

I asked teachers how much they have changed their instructional activities, the amount or type of homework assigned, and the ways they collect information about students' progress since the implementation of standards-based grading [page 41]. Changes in one area were usually significantly correlated with changes in other areas, with the largest amount of change in the area of collecting information about students' progress.

About half of teachers are changing their instructional practices, while about half or not. When asked, "Do other colleagues in your content area(s) assign a score of 1, 2, 3, and 4 on the power standards the same way you do?" [page 42], two significantly different groups emerged. Those who answered "yes" to this question were much clearer on the difference between a 1 and a 2, a 2 and a 3, and a 3 and a 4 than teachers who answered "no." They reported that their students and students' parents have a clearer understanding of the report cards. They were more likely to have made changes to data collection methods, and overall they were much more familiar with standards-based grading. I believe there is a connection between the amount of time these teachers spend collaborating with their colleagues and the changes they have made. Follow-up is needed to better understand these interactions and the likely benefits of collaboration on consistent implementation of standards-based grading.

On questions about priorities for professional development, several items showed significant correlations. It may be advisable to identify teachers with an interest in specific rubrics, differentiation, and examining student work—core practices of standards-based grading. Teachers who showed the most interest in these activities were most interested in all other professional development options.

Chapter 2: A New Vocabulary of Standards-Based Grading

Implementing standards-based grading requires developing and sharing a language of standards and criteria. Language specific to the MMSD implementation of standards-based grading includes "understanding," "power standards," the labels "1," "2," "3," and "4," and "trend scores." Knowing how teachers are defining and using these terms is a key step in facilitating or redirecting conversations about them. Teachers very much want to have a shared understanding of these terms.

Understanding [page 47]

The MMSD has chosen to use "understanding" as the key descriptor for the labels 1, 2, 3, and 4 in their standards-based grading. Teachers described understanding in ways that fell into three categories: statements about the basic nature of understanding, statements describing the process of understanding, and statements about how understanding is visible in students' work. Of these, the third category has the most potential for developing a shared, practical understanding of understanding.

Embodied metaphors [page 48]

In this section I use cognitive scientists Lakoff and Johnson's (1999) work to identify teachers' metaphors for understanding that are based on human beings' experience as embodied persons. These include understanding as . . .

- Getting and processing information
- Seeing
- Movement through physical space
- Hitting a target or reaching a goal
- Mastery or being in control
- The result of strength

These metaphors provide very general, summative statements about a student's understanding that cannot be used reliably and consistently to inform a conversation on learning or to provide a basis for formative feedback.

Process metaphors [page 50]

Teachers' metaphors about the process of understanding may provide the basis for useful conversations about teaching and learning. These include understanding as . . .

- Sudden – a "Eureka" moment
- A slow process
- Cumulative
- Something forced upon the learner
- Completeness
- Having different levels

These metaphors provoke questions about how learning occurs and what teachers can do to best facilitate learning.

Understanding as visible in students' work [page 53]

Teachers also described understanding using parts of the framework presented in *Understanding By Design* (Wiggins & McTighe, 2005). This framework includes the following "Six facets of understanding":

- Explaining
- Interpreting
- Applying
- Having Perspective
- Empathizing
- Self-Knowledge

Teachers described assignments that target one or more of these six facets. Overall, use of the six facets was much less than the use of embodied or process metaphors for understanding. If teachers are to develop a shared language of criteria based on student understanding, more extensive conversations on this topic are needed.

"Full" understanding: A problematic metaphor? [page 57]

Because the term "full" understanding makes use of an embodied metaphor and does not focus on an aspect of understanding that is visible in students' work, the district may wish to reconsider use of this term for the level 3 score.

Power standards [page 58]

Douglas Reeves (1998; 2000) coined the term "power standards" as the most important standards in a given subject or grade level. The MMSD has adopted this term for a similar concept. Teachers report a grade for each student on two, three, or four power standards in each subject area. In math, these are content-specific standards such as the ability to solve problems with factors and multiples, or the ability to read and interpret graphs. In other subjects, the power standards are more general areas of thinking skill such as Content Knowledge, Analysis and Application, Conducting Investigations, or Writing. The general power standards are used with all topics in a course, while the specific math power standards correspond to units of instruction. These different systems have various pros and cons.

The more specific power standards provide more specific feedback to parents and students. Parents then know for which content areas their children are most successful or need the most remediation. The broader power standards help teachers to focus on the balance of skills in each class. For example, the Analysis and Application power standard encourages Social Studies teachers to think about how their assignments include higher level thinking skills; in World Languages, teachers must balance oral and written communication.

Power standards quickly become the standards that matter most and the categories that teachers use to think about teaching. It is therefore important to consider carefully the power standards that are chosen for the report card. Just as it took several years for teachers to understand and use the standards documents for their areas, it will take time for teachers to understand the power standards, use them in planning and assessment, and explain them to parents. Teachers who helped develop the power standards for their

subject area had a much clearer understanding of those power standards. Several teachers requested clearer definitions of the Reading power standards.

Criteria [page 64]

No system of standards-based grading can exist without fully developed criteria for student work. I asked teachers to explain the difference between scores (1, 2, 3, and 4), focusing especially on the difference between a 3 and a 4. The district is developing grading guides that are intended to help teachers distinguish among levels of performance. However, these grading guides are not intended as rubrics for everyday classroom use. Sample rubrics for each subject (and accompanying examples of student work) will need to be developed in addition to the grading guides.

Teachers' descriptions of criteria for a 3 or 4 fell into three basic categories: vague, evaluative, and feedback-oriented.
- Vague criteria [page 66] do not help to clarify specific differences.
- Evaluative criteria [page 67] are judgments about the student's work in relation to objectives of the course and about the student's likelihood of future success. As far as I know, there is no district document that specifies what a 1, 2, 3, or 4 should represent in making decisions about course selection, remediation, tutoring, or retention. Consensus on and consistent application of these meanings would enhance the overall integrity of the standards-based system.
- Feedback-oriented criteria [page 68] include enough information about the desired performance to help teachers apply them consistently and for students to understand how to improve their work from one level to the next. These include use of evidence and sources, perspective-taking, understanding significance, quality (including detail and accuracy), quantity, and student independence. These specific criteria map well onto the six facets of understanding described above.

Criteria may be general or subject-specific. Use of general criteria across subject areas can facilitate student understanding if teachers use the criteria the same way. Subject-specific criteria are often needed to provide more detailed feedback to students. Development of criteria and technically sound rubrics is a collaborative process that begins with examples of student work.

Trend scores [page 73]

Instead of an average grade for the quarter, teachers now calculate a trend score for each student for each power standard. As a summative grade for the quarter, the trend score represents that student's level of achievement. This means taking into account the most recent and most consistent information about a student. The teachers I interviewed were very clear about the method and purpose of creating trend scores. Some teachers surveyed were less clear about their purpose. In future in-services about trend scores, district administrators may wish to repeat the rationale for their use and their relationship to formative assessment. Teachers have several practical questions about the use of trend scores, including whether or how to weight different types of assignments when calculating a trend score [page 74].

Chapter 3: L1 Interference and Interlanguage

While learning the language of standards-based grading, teachers are combining first language (traditional grading practices) with new language to create an interlanguage. This "interlanguage" is a logical transitional step. The two main interfering practices from traditional grading are use of percentages and use of letter grades. In trying to create a "direct translation" or one-to-one correspondence between percentages, letter grades, and standards-based grades, teachers face added frustration. Awareness of grading practices that are still influenced by traditional grading is the first step in addressing them, so teachers can move toward a more complete adoption of standards-based grading.

Percentages [page 77]

Teachers are accustomed to assigning grades as percentages. Many teachers have created systems to convert these percentages to a 1, 2, 3, or 4. The range of 100 possible scores offers more flexibility to teachers than the four standards-based scores or the traditional five letter grades. Because teachers' interpretations of the 1, 2, 3, and 4 vary, their systems for converting from percentages also vary. I include four sample grading scales [page 79]. A score of 2 ranges from 60 to 89 percent, grades which previously corresponded to a D, C, or B. Scores including half-point intervals [see pages 81 and 82 for a sample] could increase the accuracy of grades and decrease teachers' frustration with the 4-point system.

Letter grades [page 83]

Currently, teachers assign a 1, 2, 3, or 4 for each power standard. If desired, teachers also assign weights to each power standard. The grade book software then calculates a weighted average of each student's grades to determine a letter grade for each subject. Letter grades were retained by the district in order to provide continuity for parents during the transition. Based on teachers' comments, I recommend phasing out letter grades for the following reasons:

- The purpose of standards-based grading is to provide clear information about students' progress on individual standards. Amalgamating this information into a single score seems to defeat this purpose.
- There is no shared meaning of the letter grades.
- Some teachers wonder whether a 3, which represents mastery, should be an A. Dividing lines will always be arbitrary.
- The grades produced by the new system are sometimes very different from those produced by the traditional grading system, which makes the grades seem incorrect or confusing to teachers and parents and makes it more difficult to focus on measuring student understanding.
- Learning skills do not affect the letter grade, which suggests that they are less important or unimportant.
- Teachers sometimes adjust the letter grades or standards-based scores, resulting in inconsistencies.
- Letter grades continue to be associated with a norm-referenced system of grading, and subtly reinforce teachers' expectation of a bell-shaped distribution of grades.

Letter grades and weighting of power standards [page 88]

If letter grades are eventually phased out, it would be unnecessary to have a district consensus on weighting of power standards. Nonetheless, discussions of issues raised by the question of weighting could help to increase consistency and/or clarify curricular approaches. Weighting of power standards raises the following questions:

- Are power standards in a given subject equally important?
- Should power standards be weighted based on their importance and/or based on the emphasis given to each one in class?
- Should more important power standards have more assessments?
- If a standard has not been addressed yet or a summative assessment has not been given, how should that be reported on the report card?
- Should the same grade for Reading Comprehension count in both 6th grade Reading and Language Arts?
- Should weights remain consistent from quarter to quarter, or can the weights change during the year?
- Should higher-level thinking power standards like Analysis and Application receive more weight than Content Knowledge power standards?
- Should weighting be determined at the beginning or end of the quarter?
- What is the most important goal or power standard in each course?

If priorities in any course are redefined, teachers may require additional time and resources to make their curriculum and assessments reflect those priorities.

Chapter 4: Practical Obstacles to Learning a New Language

Learning the language and developing a culture of standards-based education requires both time and access to the proper tools. Furthermore, like any group of language learners, teachers are not a homogeneous group of learners, and they need varying levels of support, resources and time in order to make the transition to standards-based grading.

Time [page 95]

Teachers need time to understand standards and power standards, to align curricula and assessments with standards, to develop criteria for assessments, and to create and master new record-keeping systems. Teachers reported an overall increase in their workload. "Time" was the fourth most common word in teacher comments on the survey, appearing fifty-four times on twenty-eight surveys. Standards-based grading requires additional time in planning stages early in a unit and additional time at the end of the quarter when recording grades. Ideally, time for collaboration and individual work would be scheduled within the standard school day.

Teachers of multiple subjects need additional time and support for this change. Overall, teacher familiarity with standards-based grading increased from August to December [page 98]. Teachers of three or more subjects rated themselves significantly less familiar, and they also rated creation of subject-specific rubrics as a much higher professional development priority than did other teachers [page 98].

Some teachers viewed standards-based grading as a distraction from other important goals, including planning good lessons, finding resources, differentiation, and communicating with parents. This apparent conflict of interests could be deemphasized through increased attention to the ways that standards-based grading is related to and should support teachers' other goals. Without the larger vision of standards-based education, teachers resent the technical difficulties and are unable to see why it could be worth their effort.

Tools for standards-based grading [page 100]

Two important tools for the implementation of standards-based grading are rubrics and grade book software. The two main purposes of standards-based grading are to improve accountability and to improve learning. These purposes create conflicts when teachers choose which tools to use. Holistic rubrics serve accountability purposes well and are best for summative assessments; analytic rubrics are better for daily classroom instruction and formative use.

Likewise, the district's current grade book software provides information about students' progress to administrators for accountability purposes; however, it is not optimized for providing formative information to teachers or parents and students [page 102]. I summarize teachers' comments about related to implementation of IC grade book for standards-based grade reporting. In general, several teachers reported that the program is neither logical nor user-friendly. Major areas of concern include the data entry process, the visual interface, calculation of grades, and communication with

parents. It should be noted that district administrators anticipated some of these challenges, and that the software company has been asked to make changes.

Teachers as learners [page 106]

To accommodate the varied learning styles of teachers, ongoing professional development opportunities should allow for more flexibility in learning activities, combining reflection and collaboration, details and big picture.

Teachers who are most interested in the change are a valuable base of support for the implementation. Reeves (1998) recommends piloting the change before moving to a district-wide implementation, to build support for the initiative, provide concrete examples of practice, and leverage investments in professional development as the pilot group leads colleagues during implementation. Materials (standards documents, rubrics, etc.) developed by the pilot group can be improved before full-scale distribution to all teachers.

Chapter 5: The Affective Filter

The affective filter is the fear or anxiety experienced by the language learner. If the affective filter is high, it is extremely difficult for students to learn a new language effectively. An understanding of teachers' feelings about change can help to identify potential roadblocks in the transition from traditional to standards-based grading. The change to new practices may be painful, somewhat restricts teachers' autonomy, and may threaten teachers' competence—at least temporarily. Confusion caused by unclear communication and limited time to study the new system amplify these feelings of frustration.

As consummate multi-taskers, teachers sustain multiple concerns at once, each of which interacts with the others in complex ways. In this chapter, I seek to illustrate those concerns and their interactions with the implementation. Some of teachers' comments in this section are strongly-worded and indicate high levels of frustration with the implementation process. However, it is clear to me that district administrators in no way intended to cause these feelings. In every conversation with administrators, I perceived only the best intentions and a commitment to standards-based education as a means to improved student achievement.

I heard teacher frustration with the implementation of standards-based grading in the following areas:

- "Cleaning out the closet" [page 109]. Planning curriculum with a focus on standards requires teachers to reassess their lessons and activities. This can be painful as teachers must change or modify favorite activities.
- "Ceding control" [page 110]. Teachers must change their approach to motivation, which may threaten their sense of control within the classroom. These teachers may need alternative strategies for classroom management and assignment design.
- Teachers may also feel that new policies are an external attempt to control their practices. Standards-based grading may appear to threaten teachers' autonomy. Reassurance that standards-based grading is not a one-size-fits-all straitjacket approach could reduce teachers' anxiety about standards-based grading.
- "Maintaining professionalism" [page 111]. Teachers who found themselves unable to answer parents' and students' questions about the new grading system felt their professional competence was being threatened. The implementation of standards-based grading was perceived by many teachers as very rapid, resulting in feelings of isolation and overload. Teachers desired more support during the implementation.
- Understanding benefits [page 113]. Many teachers are unsure of the benefits of standards-based grading and do not understand how it fits with other educational goals.

Current clarity of understanding [page 113]

I asked teachers "How clear is your understanding of the difference between a 3 and a 4 for each power standard?" (also a 2 and a 3, a 1 and a 2). The median response was "somewhat clear" for both 3 vs. 4 and 2 vs. 3. The median response was slightly (though not significantly) higher, between "somewhat" and "very" clear for the

difference between a 1 and a 2. Teachers rated students' and parents' understanding of the new report cards as significantly lower that their own understanding.

Interestingly, teachers with ten or more years of teaching experience reported significantly higher levels of parent and student understanding than teachers with one to five years of experience. More experienced teachers likely have more and better strategies for communicating with parents. Also, 6th grade teachers reported that their students' understanding of the report cards is significantly less clear than that reported by 7th and 8th grade teachers, possibly because of different developmental stages or because of different priorities at each level early in the school year.

Lowering the affective filter [page 116]

The following practices should help to lower teachers' affective filters, thereby decreasing their anxiety about standards-based grading:
- Building a shared vision of standards-based education
- Communicating procedures and expectations clearly
- Allocating sufficient time for collaboration and individual reflection

Teachers expected that it will take several years to become comfortable with the new system of grading.

Conclusion

This chapter includes a summary of recommendations made in the previous five chapters and directions for future research on standards-based grading.

Summary of recommendations

- Develop and publish a clear philosophy of grading [page 117].
- Articulate a vision of standards-based education [page 118].
- Involve parents and community [page 119].
- Develop and publish clear guidelines for grading [page 119].
- Phase out letter grades and improve Infinite Campus [page 120].
- Increase time and support for collaboration and professional development [page 120].

Directions for future research [page 122]

I offer nine questions that address classroom practices, students' and parents' perceptions of standards-based grading, strategies for implementation and communication, approaches to homework and motivation, and the overall impact of standards-based grading.

References

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* New York: Basic Books.

Reeves, D.B. (1998). *Making standards work: How to implement standards-based assessments in the classroom, school, and district.* Denver, CO: Advanced Learning Press.

Reeves, D.B. (2000, December). Standards are not enough: Essential transformations for school success. *NASSP Bulletin*, 84(620), 5-19.

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (expanded 2nd ed.). Alexandria, VA: ASCD.

Highly Recommended Resources

Middle school culture: Anne Wheelock. (1998). *Safe to be smart: Building a culture for standards-based reform in the middle grades.* Columbus, OH: National Middle School Association.

Feedback: Susan Brookhart. (2008). *How to give effective feedback to your students.* Alexandria, VA: ASCD.

Lesson Planning and Assessment: Grant Wiggins and Jay McTighe. (2005). *Understanding by design* (expanded 2nd ed.). Alexandria, VA: ASCD.

Developing Agency: Peter Johnston. (2004). *Choice words: How our language affects children's learning.* Portland, ME: Stenhouse.

Designing Rubrics: Judith Arter and Jay McTighe. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance.* Thousand Oaks, CA: Corwin.

Student-Centered Assessment: Rick Stiggins. (1997). *Student-centered classroom assessment.* Upper Saddle River, NJ: Merrill.