

MIDDLE SCHOOL SCIENCE ASSESSMENTS: COMPARISON
OF THREE GRADING METHODS.

by

Monica Kim Tomayer

A professional paper submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2024

©COPYRIGHT

by

Monica Kim Tomayer

2024

All Rights Reserved

TABLE OF CONTENTS

1. INTRODUCTION & BACKGROUND	1
Context of the Study	1
Background	3
Focus Statement/Question	6
2. CONCEPTUAL FRAMEWORK	7
History of Grading	7
History of Science Standards	9
Standards-Based vs. Traditional Grading	11
Why Use Standards-Based Grading in Education	13
Implementation and Methodologies in Standards-Based Grading	16
Summary	20
3. METHODOLOGY	21
Demographics	21
Treatment	21
Instrumentation	24
Data Collection and Analysis Strategies	25
Data Collection	25
Data Analysis	30
4. DATA ANALYSIS	32
Results	32
Letter Grades or Percentage Grades	35
Student Perceptions About Grading	37
Student Confidence and Motivation	38
5. CLAIM EVIDENCE AND REASONING	42
1. Claims From the Study	42
Student Perceptions About Grading	45
Proficiency-Based Scoring and Standards-Based Scoring Methods	46
Student Confidence and Motivation	49
Overall Grading	50
Value of the Study and Consideration for Future Research	50
Impact of Action Research on the Author	52
REFERENCES CITED	54

TABLE OF CONTENTS CONTINUED

APPENDICES	57
APPENDIX A: Informed Consent	58
APPENDIX B: IRB Approval and Exemption	61
APPENDIX C: Natural selection and adaptation three scoring methods rubric	64
APPENDIX D: Student self-navigation tool natural selection and adaptation	70
APPENDIX E: Natural selection and adaptation standards-based assessment	75
APPENDIX F: Pre-, mid-, post survey student attitudes and perceptions on grading in science class	82
APPENDIX G: Eighth grade year at a glance standards and learning targets	89
APPENDIX H: Grading Interview Student Questions	101
APPENDIX I: Sample student grading information for two assessment units	103

LIST OF TABLES

Table	Page
1. Three Grading Techniques Scale.	24
2. Data Triangulation Matrix.....	28
3. Six student scores over five summative assessment units using three scoring methods, girls and boys, low, average, high ranking based on MAP testing scores	43

LIST OF FIGURES

Figure	Page
1. Five assessment units scored using three grading methods, ($N=31$).	33
2. Averages of three scoring methods for high, average, and low scoring, ($N=31$)	35
3. Survey results about percentage and letter grades, (Pre- $N=31$, Post- $N=30$)	36
4. Pre and post survey results related to learning targets, (Pre- $N=31$, Post- $N=30$)	37
5. Survey results on student confidence based on using learning targets, (Pre- $N=31$, Post- $N=30$)	39
6. Student confidence levels on assessments and grades, (Pre- $N=31$, Post- $N=30$).	40

ABSTRACT

What do grades really mean? This study was completed to compare three different grading methods for middle school eighth grade science assessments: Traditional Grading (number correct divided by number possible in a percentage), Conrad Proficiency-Based (rubric based on learning targets 5-10 score, 0 (missing)), and Standard-Based Score (rubric based on Next Generation Science Standards (NGSS) standards 0-4 score). Learning targets developed from NGSS standards were the main scoring method for proficiency-based and standards-based grades. Comparisons were made over five science units. Pre- and post- assessments, three surveys (pre, mid and post), and individual and group interviews were used to collect data on student's assessment scores, confidence levels, perceptions, motivation and student attitudes towards the various grading methods. The proficiency-based and standards-based scores were graded per learning target and each unit had between one and four learning targets. Traditional grading was completed over the full unit. The results showed a significant difference between the average scoring of the proficiency-based method and the other two methods. Proficiency-based scoring averages were higher than the other two methods, possibly because proficiency doesn't always mean that students have to get all the questions correct. To be considered proficient a student must have an overall understanding of the learning targets. It is possible that the proficiency-based scoring method could be inflating student grades when compared to the other two methods. Reported student confidence didn't show any changes when using learning targets to clarify scientific topics on the summative assessments. Letter grades were important to the students at the beginning of the study, but not as important towards the end of the study. This demonstrated that students were understanding that a score per learning target was giving them more information about what they actually understand about each science topic. Students also reported wanting the ability to re-test, which is an important curricular methodology for both proficiency-based and standards-based scoring.

CHAPTER ONE

INTRODUCTION AND BACKGROUND

Context of the Study

The basic purpose of this action research project was to understand the validity and reliability in grading scores using a variety of scoring methods. These different approaches included traditional methods of 0-100% with one score per science unit, standards-based methods of 1-4 score based off a rubric with one score per learning target, several scores per unit, and Conrad School District's specific proficiency-based method, which involves possible scores of 0 and; 5-10 per learning target with several scores per unit with students able to miss some questions and still be considered proficient. Both standards-based and proficiency-based methods use rubrics to score students, and are based on a single score per learning targets that come from science standards. The main differences between standards-based methods and proficiency-based methods for this research project are that standards-based scoring uses a 1-4 scale, where proficiency-based methods can use any scoring scale, for this study a 0, 5-10 was used. Standards-based scoring is also called competency-based, mastery-based, proficiency-based and many other names; regardless, they typically use a "four-category scale to communicate a student's level of mastery" (Otus, 2024, p. 1).

In a traditional method assessment, students would earn a score based on how many questions they answered correctly versus how many questions were on the assessment. Those numbers are then divided and multiplied by 100 to get a percentage. The traditional method is also not always connected to learning targets and standards, but potentially whatever the teacher

wants to teach. Currently, that really doesn't tell the student what content they actually learned. Holden & Retzlaff (2023) stated "Too often, students want to do well, but they don't have a clear understanding of what they need to do to be successful" (p. 46). Many teachers give the assessments back after grading, but they don't always go back over the correct answers, or let the student individually know what they need to practice more on. Also, some assessments in the traditional model had "mystery questions," questions that might have been in the textbook or other resources, but were never discussed in class. Such questions were not always fair and assessments in the traditional model didn't always have to follow national and state standards.

Examining and changing our grading practices has been long overdue. Grading in the United States has remained essentially intact since the Industrial Revolution, even though many common grading practices contradict contemporary research on motivation, adolescent development, effective communication, and feedback, and they even violate sound mathematics principles. Many schools changed century-old grading practices because the pandemic shone a bright light on how many of those common practices were not only ineffective and inaccurate, but often undermined our best teaching and harmed our students (Feldman, 2022, p. 8).

Proficiency-based and standards-based instruction and scoring methods are being implemented in the Conrad School District, as a means to avoid teacher bias and increase students' understanding of standards and remediate other issues that affect traditional grading. Conrad Junior/Senior School is a 7-12 school building with approximately 320 students. The school is located in north central Montana and is a rural community with low diversity of mostly middle-class white community of around 2,700 people. Currently, Conrad School District has 25% of families qualifying for free or reduced meals. I currently teach integrated life, physical and earth science to seventh grade and eighth grade students. I also teach two middle school science elective classes which have a mix of seventh and eighth grade students. I have taught middle school science for 28 years and love teaching science. Middle school students at this

developmental stage are learning life skills and this is a good time to try to shift their concept of learning from just a letter grade to “What science content do I understand?”

Background

Standards-based and proficiency-based scoring is a way to give more information to students about what they have learned, but more importantly what they still need to work on. Standards-based and proficiency-based instructional methods make it clear by stating to the student the learning target that needs to be accomplished on an assessment. The framework doesn't include grades for any formative or practice work, just scoring the final summative assessments. In an interview with mathematics teacher Heidi Horak (2023), she stated “Grading is a traditional focus and scoring is utilizing a rubric to classify the quality of the work” (p. 1). Standards-based and proficiency-based scoring requires the use of rubrics to make what is being evaluated clearer to both students and parents.

Both standards-based and proficiency-based scoring allow students to re-test on material that they didn't score well on the first time, while a normal traditional grading system doesn't allow for re-testing. In general, allowing students to retest is a good teaching practice for not leaving any child behind in science education. Using proficiency-based and standards-based teaching methods will assist my teaching, by laying out exact learning targets matched to Next Generation Science Standards (NGSS) and State of Montana science standards, by aligning assessments to those standards, by removing some teacher bias and by improving overall instruction and practices that the students work on to learn the concepts.

How does the actual grading influence the student? Do the scores on summative assessments really matter, or do teaching practices make more impact on student learning? This

action research project attempted to see if there is a difference between traditional grading and the various aspects of standards-based 1-4 scoring and proficiency-based scoring 0, 5-10.

Conrad School District has implemented proficiency-based scoring, and continues making additional steps towards standards-based grading methods. Therefore, this project assisted me in understanding how grading methods affected student achievement, confidence, motivation and perception about grading.

This research has been shared with teaching staff and administrators. I have actual data about how using proficiency-based methods affects the scoring of students in the classroom, what students felt about the process, and how the grades in the current traditional model compared to the standards-based and proficiency-based model. There are many questions about how proficiency-based and standards-based scoring will affect scholarships, student confidence, student motivation and more. This action research project attempts to find the answers.

Brookhart (2011) stated that “Standards-based grading is based on the principle that grades should convey how well students have achieved standards. In other words, grades are not about what students earn; they are about what students learn” (p. 12).

This action research project came from the idea of truly understanding the difference between traditional grading, standard-based and proficiency-based scoring. They are distinctly different scales. Understanding why they are so different, but also how either method can truly define what a student knows and understands.

Grading is a very sensitive topic for some students, including me. This summer I took a class on proficiency-based grading from Dr. Jeanette Chipps. The assignments were graded using the proficiency scale of 1-4. I was horrified when I was receiving 3’s on my assignments,

which in the proficiency scale means I was proficient, yet in the traditional scale it meant that I was at 75%. It was a terrific experience for me to actually feel what my students will feel when and if we truly transition to a proficiency or standards-based system.

I had to remind myself that it was not the percentage I should focus on, but what I have learned. I was proficient in obtaining knowledge, and I even did earn a 4 on an assignment which meant I was advanced. This experience really helped me to understand what I will need to do as a teacher to help my students understand these changes in grading. It really is NOT about the grade, but is about what you learn and how that learning will help you in the future. Students will need to be taught how to refocus and realign what they understand about grading. Therefore, I am glad that Conrad School District is moving towards a proficiency-based scoring system using a slow process. This will help students and parents get used to the idea in a more step by step way.

Students need to understand that an 82% really doesn't specify what they truly understand, whereas a three, when aligned to an "I can" statement or learning targets with a rubric can explain more of what a student knows. Standards-based and proficiency-based scoring and teaching methods that go with those practices to me is just good teaching. It is a good idea to allow students to re-test, to not grade practice or formative work, to differentiate and scaffold the learning for students and make the students understand that they are in control of their learning. Learning is not just about memorization of facts that can be regurgitated back on an assessment but the use of that learning to solve problems, inquire, and critically think.

While I am in favor of adopting these new grading policies, I have considered whether there really was an overall difference in the actual grade for students when averaging all the

learning targets together? On an assessment that has three or more learning targets, each target would get a proficiency score and the traditional would lump all the learning targets into one percentage grade. The standards-based or proficiency-based score would let the student know what they need to re-learn or re-test more specifically, but when averaged all together do they add up to the same as a traditional score. What truly is the difference in the overall grade? Is the overall grade really different using a traditional grading method versus a standards-based or proficiency-based scoring method?

Focus Statement/Question

My focus question was, how does using three different assessment scoring techniques, standards-based scale, proficiency scale scoring and traditional percentage grades, affect middle school science grades?

My sub-questions include the following:

1. What are students' perceptions and experiences with proficiency-based scoring?
2. How does using proficiency-based scoring improve student motivation and confidence in science learning?
3. How does using proficiency-based scoring impact me as a teacher in methods and grading?

CHAPTER TWO

CONCEPTUAL FRAMEWORK

History of Grading

Lehman et al. (2018) shared “grading is one of teachers’ greatest challenges and most important professional responsibilities” (p. 1). The current commonly accepted grading system in the United States was based on a pair of systems developed at Harvard and Mount Holyoke in 1877 and 1897 respectively (Durm, 1993; Townsley & Buckmiller, 2020). Educational grading practices that haven’t changed since 1897 do not match with today’s educational needs. Yet, these scales are almost identical to the scale used at most middle and high schools in Montana today. The Mount Holyoke scale from 1897 reads “An excellent, equivalent to percents 95-100, B Good, equivalent to percents 85-94 (inclusive), C Fair, equivalent to percents 76-84 (inclusive), D Passed (barely) equivalent to percent 75 and E Failed (below 75)” (Durm, 1993, p. 296). Today, schools have adjusted that slightly and don’t use the letter E, but the scale is still very similar. Earning a grade below 75 constitutes many more points that relate to failing a class than passing a class. When and why did we get rid of E to use F instead? (Durm, 1993).

At first, grading in the 19th century was more descriptive, where teachers used “expressive adjectives” to let parents know about how their child was doing in school (Durm, 1993, p. 295). Oral reports were given to parents about their child in their home during this time; teachers also included written narrative descriptions of student’s performances on critical concepts (Brookhart et al., 2016). This brings up many questions including: is the current grading system correctly calibrated to student achievement? It was stated in the 20th century that

a way to make grading faster was needed due to the number of students attending school, so that is partially why the grades went away from being descriptions. Whether or not numbers and letters really identify what a student has learned in the classroom has been a question for some time; why has it not changed before now? (Brookhart et al, 2016, Durm, 1993; Townsley & Buckmiller, 2020).

Brookhart et al. (2016) shared, “grading refers to the symbols assigned to individual pieces of student work or to composite measures of student performance on student report cards” (p. 804). What do grades really mean? Should just a symbol be used to grade student work? What does that symbol suggest about the work? Grades can affect a student’s future success, college admission, scholarships, dropout rate, type of pre-requisite classes and more. Grades also affect students’ self-esteem and confidence. Studies have found that what students think about themselves can affect how they learn in various subjects (Brookhart et al., 2016). It was also found that student work habits, effort, and motivation are all strong indicators of what grades students will achieve in both traditional and standards-based grading (Brookhart et al., 2016). Grading should have its purpose in communicating about student progress and academic achievement on specified learning targets. The 1983 A Nation at Risk report, recommended that schools adopt more rigorous and measurable standards; this study spurred the standards-based methodology process in schools (Lehman et al., 2018). Townsley & Buckmiller (2020) stated that “A guaranteed and viable curriculum is the foundation for any school improvement effort” (p. 1). In order to have a good grading system, it is necessary to begin with a good curriculum and assessments for students to complete. The move to aligning curriculum with standards began with A Nation at Risk in 1993 and continued with No Child Left Behind, Every Child

Succeeds Act, and the Common Core State Standards. All of these national initiatives emphasize learning targets based on standards and aligning those learning targets with the assessments given to students. Grades should be based on achievement targets rather than all the other factors that go into a student's day including attendance, behavior, and participation; thus, the growing need for standards-based grading practices (Hooper & Cowell, 2014). Each of these pushes educators to separate behavior from the actual learning of content, to not grade homework, and to allow multiple attempts at summative assessments. (Townsley & Buckmiller, 2020; Zimmerman, 2016).

History of Science Standards

With more and more students attending college in the late 1800's to early 1900's the education system needed a change in science education. Therefore, the National Education Association started requiring all students to take science classes including laboratory work in 1892 (Belcher, 2018).

After World War II, during the 1950's, the National Science Foundation (NSF) was awarded funding to increase science education by providing teacher professional development and improving curriculum (Hechinger Report, 2011). This increase in funding was due to the need to increase the science and technology workforce needed in the United States (Belcher, 2018). Then in 1985 the American Association for the Advancement of Science (AAAS) launched Project 2061 with two reports: "Science for All Americans" and "Benchmarks for Science Literacy" (American Association for the Advancement of Science, 2023; Hechinger Report, 2011). These documents started a long-term initiative to develop high quality science, mathematics and technology standards to be used in classrooms across the United States,

therefore increasing science literacy (American Association for the Advancement of Science, 2023). Project 2061 developed many different curricular tools and assessments for use in the science classroom and launched the start of a new way to look at science inquiry. The development of science learning goals based on standards and aligned assessment question banks to be used by educators was key to this initiative (American Association for the Advancement of Science, 2023).

The next steps in the development of science standards were in 1996 when the National Research Council (NRC) developed the National Science Education Standards (NSES) (Hechinger Report, 2011). These standards were to provide science literacy to all students including “what students ought to know, understand, and be able to do at various stages in their K-12 education” (American Geosciences Institute, 2023, p. 1). They also laid out a plan for study for students in K-12 grade levels to assist teachers in developing proper curricular flow and increase inquiry and critical thinking about science.

Currently, science teachers in most states in the United States use the Next Generation Science Standards (NGSS). These standards were developed in 2013 to improve upon Project 2061 and NSES which were aging documents. The NGSS were developed because “science--- therefore science education---is central to the lives of all Americans” (Next Generation Science Standards, 2023, p. 1). The NGSS developed the three-dimensions of science education including; cross cutting concepts, science and engineering practices and disciplinary core ideas to increase students’ attainment of information but also developing key science skills (Next Generation Science Standards, 2023). Those skills include communication, inquiry, collaboration, critical thinking, problem solving, inquiry and flexibility (Next Generation

Science Standards, 2023). Overall, all the various science standards built were to increase all Americans' scientific knowledge and critical thinking to help improve society into the future.

Standards-Based vs. Traditional Grading

The article “Debunking Myths of Standards-Based Grading” included good information specifically about science standards-based scoring (Wilcox & Townsley, 2022). Some of the major differences between traditional and standards-based scoring include allowing retesting, redoing assessments, not grading homework or anything that is considered practice, placing learning goals in the gradebook instead of assignment names, and grading by learning goal instead of giving a certain number of points for the assignment. The article includes examples of traditional gradebook versus a standards-based gradebook, showing that the standards-based gradebook gives the parents and students a lot more information about what is actually being graded (Wilcox & Townsley, 2022).

A traditional grading system also compares students against students. The standards-based scoring system corrects this by scoring students against themselves, and looks at students as individuals (Hooper & Cowell, 2014). In a way, traditional grading practices were meant to identify learners versus non-learners in the classroom. Students also absorb this idea from traditional grading practices. They learn quickly that they are “not a good learner” and lose motivation and self-confidence in their abilities (Brookhart, 2011).

Standards-based scoring doesn't allow teachers to include the non-achievement factors like tardies, late work, attendance, effort, cooperation, ambition, perseverance, punctuality or behavior (Brookhart et al., 2016; Townsley & Buckmiller, 2020; Wilcox & Townsley, 2022).

Standards-based scoring essentially converts an arbitrary letter grade into something that actually has meaning to the teacher, student, parent and school. This type of grading ensures that what is being assessed is clear to the student. Students work toward mastery of the learning target and summative assessments scoring that learning target can give the teacher better information about the individual student (Hooper & Cowell, 2014).

Scriffiny (2008) stated that “In the adult world, everything is a performance assessment” (p. 73). Adults have to fix mistakes and make better decisions, or face losing money or even their job. Students need to understand that there can be another chance, that learning from mistakes is a normal part of life. Teaching a growth mindset to our students is important. Scriffiny (2008) also stated that “Quality matters, and the ability to measure the quality of one’s own work is a learned skill” (p. 73). The goal is to develop the skills necessary to self-regulate and respond to external feedback in the adult world, which is made easier with standards-based scoring; students can understand what learning targets they know and which ones they need to work on to be proficient (Scriffiny, 2008). Students also need to know that they have control over their grades and take ownership of their achievements and learn from their mistakes (Scriffiny, 2008).

Traditional grading has caused parents to be caught up in the scale and not what the grade actually means. Brookhart (2011) added that “The first task in successful grading reform is to reach consensus on the purpose of grades” (p. 10). In moving towards standards-based scoring, administration can use the statements from Brookhart to start the conversation;

Brookhart stated “Which Do You Believe?

1. Grades should reflect achievement of intended learning outcomes---whether the school is using a conventional, subject-based report card or a report card that represents these intended learning outcomes as standards.

2. The primary audiences for the message conveyed in grades are students and their parents; grading policies should aim to give them useful, timely, actionable information. Teachers, administrators, and other educators are secondary audiences.
3. Grades should reflect a particular student's individual achievement. Group and cooperative skills are important, but they should be reflected elsewhere, not in an individual's academic grade.
4. Grading policies should be set up to support student motivation to learn. A student should never reach a place where there is no point in doing any more work because failure is inevitable" (p. 14).

These are strong statements, and in order to make sure that everyone is on board with standards-based grading practices, they should be discussed. Brookhart also stated that "Standards-based grading is based on the principle that grades should convey how well students have achieved standards" (p. 12). In order to shift to standards-based scoring, teachers need professional development to help them with all parts of the classroom, including formative and summative assessments, setting up learning targets, differentiation for students' learning, flow of curriculum, questioning techniques and more (Brookhart, 2011). Switching to a standards-based classroom takes a lot of work and research to truly implement the standards-based approach, but research has shown that student scores started to align better with the standardized test score the students earned (Deddeh et al., 2010).

Why Use Standards-Based Grading in Education

Across a variety of classrooms, even with all other factors being equal, students may perform differently on assessments due to various factors, including what standards the teacher is teaching, lack of specific grading criteria, and the value that teachers place on different types of content (Brookhart et al., 2016). Teacher bias and subjectivity also affect grading, including what each teacher deems is important and not important for grading. Most teachers do not want to fail

students that have given a lot of effort toward learning. Errors in grading were even found in handwritten or typed work, including rounding errors (Brookhart et al., 2016). It is suggested that in order to help teachers, professional development on writing curriculum, developing learning goals and identifying certain criteria in student work would improve the quality of grading (Brookhart et al., 2016). Teachers rarely have the time or resources to seek that kind of help or assistance to see if their assessments are valid or reliable measure of content standards, outside of district-provided professional development opportunities (Brookhart et al., 2016). Brookhart (2016) added that “Teachers often make grading decisions with little school or district guidance” (p. 826).

Standards-based scoring also gets rid of most of the teacher bias that goes into traditional grading. Inequalities in grading can come from many places; according to Hanover Research (2011), these grading practices can place inequalities in a student’s final grade: “using a point system and averages, using zeros as punishment, grading homework and other formative assignments, grading on a curve, allowing extra credit, grading for behavioral issues and incorporating teacher expectations and judgements into grades” (Lehman et al., 2018, p. 7). All these practices have been used in traditional grading and are not supposed to be part of any standards-based scoring practices. Even if teachers recognize the bias that they can put into a grade, sometimes the grade is still inflated or deflated based on the subjectivity of the test questions (Townesley & Buckmiller, 2020).

Even though standards-based grading can be more work for teachers, it is considered to be a better system by many researchers because it is more aligned with what the student actually understands and it promotes critical thinking and problem-solving (Townesley & Buckmiller,

2020). Yet, in order to have a positive change, the teacher needs to believe that standards-based scoring is the best approach. Teachers are in charge of building curriculum and assessments that will be used in their specific classrooms; if those items do not align to standards, then standards-based grading will not work (Townsend & Buckmiller, 2020). Teachers also need to consider the specific academic needs of individual students when considering standards-based scoring. It is a shift in both grading and philosophy compared to the traditional bell-shaped curve comparison of students (Townsend & Buckmiller, 2020).

In any grading process for things to be equitable communication with the student is key. Students need to understand the learning target when presented with clear student level-appropriate language. Students normally want to achieve and do well, but don't always know how to get there. Holden & Retzlaff (2023) stated that "Too often, students will give up or lose hope simply because they start to feel like they don't understand what it is the teacher is asking of them" (p. 46). It is not giving away answers to students to clearly define the learning target and make sure that students understand; it instead gives students hope that they can achieve success (Holden & Retzlaff, 2023).

Scriffiny (2008) gives seven reasons for using standards-based scoring.

1. Grades Should Have Meaning
2. We Need to Challenge the Status Quo
3. We Can Control Grading Practices
4. Standards-Based Grading Reduces Meaningless Paperwork
5. It Helps Teachers Adjust Instruction
6. It Teaches What Quality Looks Like
7. It's a Launchpad to Other Reforms (pp. 71-73)

These are well thought out reasons to choose standards-based scoring over traditional grading.

Traditional grading doesn't really demonstrate if the student has achieved mastery or even

proficiency for any learning target. Why are we assessing homework; when students are just in the learning stage, and penalizing students that don't do the homework, even if they already truly understand the learning target? Teachers also spend a great amount of time grading formative and summative papers in traditional grading without additional feedback to students (Scriffiny, 2008). That time could be better spent giving students feedback about their work.

A zero on an assessment is damaging and not proper for grading a student. O'Connor (2011 in Hopper & Cowell, 2014) noted that "the mathematical problem with zeros is that they represent very extreme scores and their effect on the grade is always exaggerated" (p. 61). Traditional grading has often used the zero to punish students, where standard-based scoring takes that zero out of consideration, allowing retesting until the score is appropriate.

Implementation and Methodologies in Standards-Based Grading

This creates a challenge for standards-based grading: how to truly calculate grades. A scale is needed that best describes what a student truly understands in relation to the learning target. Standards should be assessed multiple times for accuracy in grading (Hooper & Cowell, 2014). As a result, the teacher needs to consider whether to take an average of those two scores (which would be the more traditional method), or to take the most recent score, which is a more standards-based approach (Hooper & Cowell, 2014). In standards-based assessment, retesting is important but there should be a limited amount of time for students to complete their assessment (Deddeh et al., 2010). Students can retest as many times as they need to as long as they are making progress within that time period (Deddeh et al., 2010).

Steps to actually putting standards-based grading into practice in a science classroom include starting with an assessment and identifying what standards that assessment addresses, by

examining the quality of the questions, the level of learning for the questions, and how to create a rubric for the assessment (Wilcox & Townsley, 2022). Wilcox and Townsley (2022) clarified by adding that teachers should find out “What can these students do and do not do in relation to the standard that was addressed? Create a bulleted list for each pile. These bulleted lists can turn into a rubric for the standard” (p. 33).

Zimmerman (2016) demonstrates how a physics classroom can run using standards-based grading. The author gives steps and information for working through the whole process, including setting up learning objectives, assigning grades, keeping track of learning objectives, formatting assessments, deciding on the number of assessments for retesting, and determining final grades (Zimmerman, 2016). In this classroom, an assessment of four or five questions was given each week over one or more learning targets. The learning targets need to be the first step in starting a standards-based classroom and written from the standards broken into small chunks. Standards should be broken into “discrete skills that make up a content area standard” (Holden & Retzlaff, 2023, p. 48). Standards-based scoring takes a lot of time, time to plan the objectives, time to write the assessments, time to grade and assess grades, and time to reassess students as needed (Zimmerman, 2016). Zimmerman (2016) makes a statement that I completely agree with “While you may be concerned that reusing assessments may lead to student cheating, it is my belief that if students want to sit down and memorize several versions of an assessment, they will have learned the material fairly well despite themselves” (p. 48). The author suggests making at least three versions of the assessment to use for testing and retesting. The versions should have the same concepts but slightly different scenarios (Zimmerman, 2016).

Summative assessments placed in the gradebook should be work done by individuals and not groups, even in science. Therefore, laboratory practice should be done in groups, but the final assessment of what the students understand should be done with an individual laboratory practical or other form of summative assessment (Deddeh et al., 2010).

Homework is essential, yet not graded, and teachers should understand that some students will do the work, and some won't, but keeping track of how the students does on that work and working with the students that don't understand or don't do the work is important (Deddeh et al., 2010). Deddeh et al. (2010) suggest that students, not teachers, should keep track of their own formative work grades and completion. Their unit plan should include: the standards, the "I can" statement, vocabulary section, pages of the textbook to read, then some simple practice (Deddeh et al., 2010). This is a simple, yet straightforward way for students to truly understand what they need to accomplish. This could easily be placed in a science journal. It is important to have students keep track of what objectives they have mastered. This gives the students ownership of their grades and assists with a growth mindset (Deddeh et al., 2010; Zimmerman, 2016).

Deddeh et al. (2010) have three core beliefs: grades should communicate what the student has actually learned, homework is an important part of that learning, and students may need more than one attempt at the practice and the summative assessment.

In standards-based grading there is an emphasis on the idea that practice comes first and some students need more than one way to practice to make sure they understand the concepts. Summative assessment should not be given until the teacher knows through formative assessments "that a reasonable number of students will score proficiently, and that makes each assessment mean much more" (Scriffiny, 2008, p. 73). Summative assessments and standard-

based scores give the teacher, parent, and student more information and allow for adjustments for individual students' needs. Also, gifted students can be working on more advanced topics related to the standard instead of being bored, and students with disabilities or other issues can get modification and accommodations to help them achieve (Scriffiny, 2008).

Rubrics are a key factor in standards-based scoring; they specifically state what the teacher is grading on the assessments, making it clear for the student, parent, and teacher (Holden & Retzlaff, 2023). Rubrics help a student understand why they got the grade they got and what they need to improve to increase that grade. Teachers need to use rubrics to clarify expectations on assessments and also to assist students with areas that they need to work on, making communication of grade more successful (Holden & Retzlaff, 2023).

When moving towards standards-based grading, administration needs to consider starting the conversation with staff about the purpose of grading before they address gradebook or report card issues. Administrators need to consider the pace at which teachers, students and parents are moving towards the philosophy and implementation of standards-based scoring (Townesley & Buckmiller, 2020). It is not something that can just be changed overnight; communication is key for all parties involved in the scoring switch. Teachers will need professional development to assist with new techniques with curriculum, assessments, building rubrics, breaking down the standards into smaller pieces and writing learning targets. Parents and teachers will also need to learn and adjust to proficiency levels and what those mean instead of percentage letter grades (Townesley & Buckmiller, 2020).

Summary

In summary, grading and scoring methods need to be clearer when describing what the student knows and still needs to learn. Those descriptions should be based on learning targets that are derived from standards. Science standards have changed significantly over the years, and grading practices need to be adjusted to clarify students' understanding of those statements. Changes do need to be made, primarily by moving away from a traditional method of grading, which includes not grading formative work and not including non-academic items in grades, such as behavior or attendance. Students should only be graded against themselves and not others. Teachers need to communicate clearly with students and parents to make grades have more meaning. Ultimately, no matter which scoring or grading method is used, that method needs to clearly define student understanding of science standards.

CHAPTER THREE

METHODOLOGY

Demographics

This study included 31 middle school eighth grade students. The age range is twelve to fourteen years old. These students are in two sections, both receiving the same treatment. For this action research project there isn't a need to have an experimental group and control group since just the students' assessments are being scored differently.

Section one contains eighteen students, with eight males and ten females all in a general education setting, with two students on a 504 plan (one medical and one technology related). No additional staff or aides assists with this section. Section two contains thirteen students that will have data collected for this action research project. This section has eight males and five females in a general education setting with one female and one male having an Individualized Education Plan (IEP). I also co-teach this section with the Special Education teacher for the middle school and there is a paraprofessional assisting a student that will not have data collected.

Students and parents signed an informed consent paper (Appendix A). The research methodology for this project received an exemption by Montana State University's Institutional Review Board and compliance for work with human subjects was maintained (Appendix B).

Treatment

The treatment included starting a new grading system of proficiency-based scoring, which was a 10-5, 0 score on each learning target aligned to a NGSS standard. The students

were taught about the scoring and shown a rubric aligned to the NGSS standard(s) (Appendix C) and self-navigation tool (Appendix D) which included all the learning targets. The teacher scored each student's assessment three ways for each of the main learning targets on the unit summative assessment, a traditional overall percentage, a standards-based (1-4) and proficiency-based (0-missing), 5-10) score. For an example from the natural selection and adaptation unit see Appendix C.

During the course of the treatment which included five-unit summative assessments, (Geological Time, Natural Selection and Adaptation, Evidence of Evolution, Forces and Motion, Mechanical Energy), some with laboratory practical assessments, students were taught exactly the same way. This included not grading any formative work, allowing retesting on any assessment for any student, all assessment questions were aligned to standards and written in student friendly language in the form of learning targets, hands-on laboratories with partners were performed, videos about phenomena were watched and discussed, McGraw Hill Inspire Science workbook and activities were completed and classroom discussions were held. Students also partook in a pre-assessment at the beginning of each unit. Then the units were taught by taking one to three learning targets at a time, completing the various activities related to teaching that learning target, then moving on to the next set of learning targets related to the unit assessment. An example of a unit assessment for natural selection and adaptation can be found in Appendix E. The treatment included giving students smaller chunked summative assessments and then allowing for a second time with the learning target on a larger unit assessment. This step was only completed for the last unit in this study, mechanical energy.

Students were given three surveys (Appendix F) throughout the five-unit summative assessments that gauged their feelings, confidence level and beliefs about grading in the science classroom. These surveys were used to help triangulate the data and observe whether there was a correlation between student achievement, confidence level and understanding of grading methods. The surveys were followed up by doing individual and group student interviews (Appendix H). Additional standardized test scores using Measures of Academic Progress (MAP) testing were given at the beginning and the middle of the five-unit summative assessments for further triangulation.

The school district has been requiring teachers to input summative assessments into the grade book by the learning target, including rubrics for each learning target posted for parents and students, and additionally grade every summative assessment with the same point value. Appendix C demonstrates the three different grading methods that were used in this research. This was not a complete move to standards-based scoring four-point scale, but a step towards it.

Previous methods used by the school district were based on traditional grading. All formative and summative assessments and assignments were graded and put in the grade book. Students that earned less than 50% on an assessment were given the lower grade. Teachers didn't give retests, gave lower scores for late work and although they used standards to guide the curriculum and assessments, didn't make students aware of the learning targets that align with the standards. The formative assignments could have any value and summative assessments normally were based on 100 percent.

Other steps the district took moving toward proficiency-based scoring includes; re-testing, nothing less than a 50% (a score of 5 on the proficiency-based scale) is given if the

student attempts the summative assessment, no formative work is counted towards final grades, and advisory (intervention) time with teachers is provided for re-learning opportunities. This treatment included all items discussed since those are required by the school district. The variation of the treatment only came in the form of the different ways the assessments were scored (Table 1).

Table 1. Three grading techniques scales.

Traditional Letter Grade	Traditional Percent Grade	Standards-Based Score	Standards-Based Wording	Conrad Science Proficiency Scoring for 2023-2024 School Year
A	100	4.0	Mastery Advanced	10
A-	90	3.5	Above Proficient	9
B	80	3.2	Proficient	8
C	75	3.0	Nearing Proficiency	7
D	60	2.0	Developing	6
F	50	1.0	Basic	5

*Adapted from (Marzano, Formative Assessment & Standards-Based Grading, 2010, pp. 106-110) and (DeBruycker, 2023).

Instrumentation

The instruments used in this research project included pre-assessments from McGraw Hill Inspire Science website (McGraw Hill Education, 2020). The pre-assessments are considered reliable and valid because the source is a textbook company that pilots their materials through many school districts in the United States. The post-assessments were developed to align to learning targets in sections on each assessment and written in Google Forms for ease of grading and student accessibility. The post-assessments had questions taken from McGraw Hill Inspire Science teacher materials and website, but also from other textbook sources the teacher has gathered from twenty-eight years of teaching middle school science. The post-assessments were developed with each learning target having between five and ten summative assessment

questions of various types; multiple-choice, check boxes, matching, and short-answer questions. See Appendix E for example of a Natural Selection and Adaptation summative assessment. Depending on the number of learning targets/standards on the unit assessment, a per learning target retest was developed using McGraw-Hill Inspire science materials, Glencoe Science and other teacher materials. These assessments are addressing validity and reliability according to the review of another science teacher, previous years student data, and the use of pre-made materials from various textbook companies including McGraw Hill Inspire Science and Glencoe Science.

The three surveys used in this research project, which were a combination of Likert scale and short answer questions, were developed from viewing various standards-based grading and other science assessment grading research projects and adapting questions to fit the needs of this project. The student surveys were administered prior to treatment, at the midpoint, and post treatment (Appendix F). Some of the interview questions also came from those same sources, and from student statements on the Likert scales and short answer surveys given.

A student self-navigation tool (Appendix D) was created to assist the students by having one place where all the standards with the aligned learning targets were written out for the unit. This tool also had places for students to write notes, keep track of understanding, allow for self-reflection, set goals for the unit, and extend thinking about concepts in the unit.

Data Collection and Analysis Strategies

Data Collection

In the first step of the treatment, students were given a survey to establish their feelings about traditional grading and what they currently know about standards-based and proficiency-

based scoring (Appendix F). The surveys helped determine how the students adjusted to the standards-based/proficiency-based scoring and what they thought of the process. Students were used to being graded with percentage scores out of 100 points using traditional grading methods. During the treatment, students were shown both their Conrad proficiency-based score and their traditional score. During last two summative units included in the treatment, students were shown all three scores. Survey data assisted me with learning about student confidence and motivation towards learning the content using the new standards-based techniques.

The first step to each of the five summative unit assessments was to give students a pre-assessment on a specific NGSS and Montana State Science Standards using McGraw Hill Inspire Science website pre-assessments.

Next, the students engaged in practice, hands-on activities, classroom discussion and other methodologies to learn about one to three learning targets that are aligned to a specific standard. A science standard can have anywhere from two to eight learning targets built from that standard. See Appendix G for the sample year-at-a-glance developed during the Proficiency-Based Grading class taken this summer from MSU-Bozeman and Dr. Jeanne Chipps. This document allows the teacher to make sure the pacing is right for completing the required standards and for aligning the learning targets to the NGSS science standard. This alignment was adjusted throughout the school year as the teacher and students covered the content in each learning target. Appendix G continues to be a work in progress and was used throughout the treatment. The learning targets used for this treatment are found in Appendix G and consist of the geological time scale, natural selection and adaptation, evidence of evolution, forces and motion and mechanical energy units.

Students received a student self-navigation tool to assist them with their learning (Appendix D). This tool helped students to keep track of the learning targets and their mastery of those targets. The student self-navigation tool was a way for students to see all the learning targets related to the NGSS content standards for each unit of study. This tool was used to check for understanding of those learning targets and helped the student prepare for the assessment. It was meant to be a useful tool for students to gauge their learning progression and for the teacher to monitor the student's learning. The tool was used to introduce the learning targets addressed throughout the unit and during the review for the unit assessment. The teacher probed students' understanding and monitored whether they checked the box that indicated he or she was ready or needed more information. Information from this tool assisted the teacher in understanding student confidence levels on the science content and to see if it helped the student understand what they were graded on.

The learning targets were then taught in a systematic order using the Inspire Science by McGraw-Hill student workbook and additional activities from the teacher's wealth of knowledge (Briggs et al., 2020). This process continued until all learning targets were formatively assessed, with students filling out the student self-navigation tool sheet throughout the process (Appendix D). A unit review and connections between the learning targets were taught in class and then an overall 10 to 45 question unit assessment was given with each learning target having a specific section with questions and questions combining the learning target concepts together. Students also completed several laboratory practical unit summative assessments for specified units that were combined with the various learning targets they addressed (Appendix E). On the unit test, each section was given a standards-based score, proficiency-based score, and a traditional grade.

The unit review was completed because research states that standards should be assessed multiple times (Hooper & Cowell, 2014, p. 63). This process continued with the next unit of science standards until five complete sets of unit assessment data were obtained over a time period from October 2023 to March 2024 (Table 2).

Table 2. Data Triangulation Matrix.

Research Questions	Surveys Before, Mid and Post	Pre-assessments Scores	Unit Assessments and Learning Target Scores	Re-Testing Assessment Scores	Interviews: Individual and Small Group After Assessments	Measures of Academic Progress (MAP) Test Results
Main Question: How do middle school science traditional grades coordinate to standards-based proficiency scale grades and percentage grade using the same classroom instructional techniques?	1	2	3	3	1	4
Secondary Question 1. What are students' perceptions and experiences with proficiency-based grading?	1				1	
Secondary Question 2: How does using proficiency-based grading improve student motivation and confidence in science learning?	1				1	
Secondary Question 3: How does using proficiency-based grading impact me as a teacher in methods and grading?	1				1	

Key:

1. Qualitative and quantitative data from students' attitudes, feelings, perceptions, confidence levels and/or motivation before, during and after assessments to compare and look for changes were completed with surveys, observations and interviews.
2. Quantitative data from student learning achievement prior and during teaching methods gave information on the progress of learning for the students.
3. Quantitative data from student learning achievement at the end of the unit gave information on the final and overall grades of each student to see if there are differences in overall grade and understanding of the standards.
4. Standardized test scores from students Measured Academic Progress (MAP) Testing were used to see any growth with individual learning target scores versus complete unit scores.

Throughout this research project, teacher journaling and data collection also assisted me with understanding how I felt during the process and what curricular, assessment and other changes I made to assist students with learning. During these processes, I also considered whether or not the data matched with what I believed was the students' achievement and motivation level.

After completion of three summative units in the treatment, individual student interviews were conducted with volunteer students to determine their thoughts and feelings about the variation in grades and information that they received about the grades from the teacher (Appendix H). After completion of five units in the treatment, a group interview session was completed with volunteer students in each section and all students doing a short answer set of questions about grading. The two different sections of eighth grade science each had their own unique learning styles and thoughts, so I wanted to see the variation. The student interviews and surveys were used to help answer the sub-questions.

Data Analysis

The survey data was organized in a stacked bar graph that shows the strongly disagree, disagree, neutral, agree and strongly agree in a bar for each question. The three survey's data was analyzed for these themes; student perception of grades, student confidence and motivation towards grades, students understanding of learning targets, students view on the learning process and curricular changes. The interview questions were coded for those themes to find any trends, outliers and interesting information. The interview information that stands out and correlates best with this study was placed in a table.

Interviews were conducted with students who volunteered to be interviewed. The focus of the interviews was to see how the teacher can assist the student's learning process, what techniques could be modified, implemented or adjusted. Additionally, the interviews were used to seek information about student motivation and confidence in taking assessments, including how and why the student felt that way. The formative and summative assessments were used to look at student growth and potential, but also to see if adjustments in questions on the assessments were needed throughout the treatment.

Scores from MAP testing that was administered in September (pre-treatment) and January (mid-treatment). The scores in September helped establish a baseline of where students were at using that standardized data. In January, the scores were used to see if the students were making any progress, as well as to compare their progress to previous years. This assisted me to determine how well the new methods of learning targets, self-navigation tools and the revised grading methods assisted with overall student achievement.

Post-assessment data was analyzed for mean, mode, and range. This information was placed in a box and whisker graph. This data was viewed to consider trends in student

achievement over the various summative units and to compare the three different scoring methods. An ANOVA test was used to see if the three grading methods scoring data was significantly different and then t-test data was used to find out which scoring methods had differing results (Mertler, 2020, p. 187).

Data triangulation was completed by comparing student scores across the five summative units with their changes in confidence levels throughout the surveys and interviews. MAP testing scores were analyzed to see if students were making any changes in achievement from past scores using more traditional methods to current scores using more proficiency-based methods. MAP data was also used to place students into groups; (high, middle and low), to see if there was a scoring bias for the various levels of students. Students do not have to get every question correct to score a higher grade in the proficiency model, therefore it was thought that it would motivate students to continue to make educational strides.

CHAPTER FOUR

DATA ANALYSIS

Results

Finding from this study suggest that the Conrad proficiency-based scoring gave students a higher overall average score. As shown in Figure 1 the Conrad proficiency-based scoring had an overall higher mean of 83.5% (B) versus the means of traditional grading method 76.3% (C) and standards-based scoring 78.1% (C+) (Figure 1). The figure also shows that quartile three numbers for the traditional grading and standards-based scoring are close to the quartile two of the proficiency-scoring method. Also, of note is how the proficiency-based scoring methods quartile one is very close to the means of the other two grading methods. It should also be noted that the lowest score possible in the proficiency-based scoring is nearly double the lowest scores for the traditional grading and standards-based scoring. Students can earn much lower grades in those two methods than the proficiency-based method of grading.

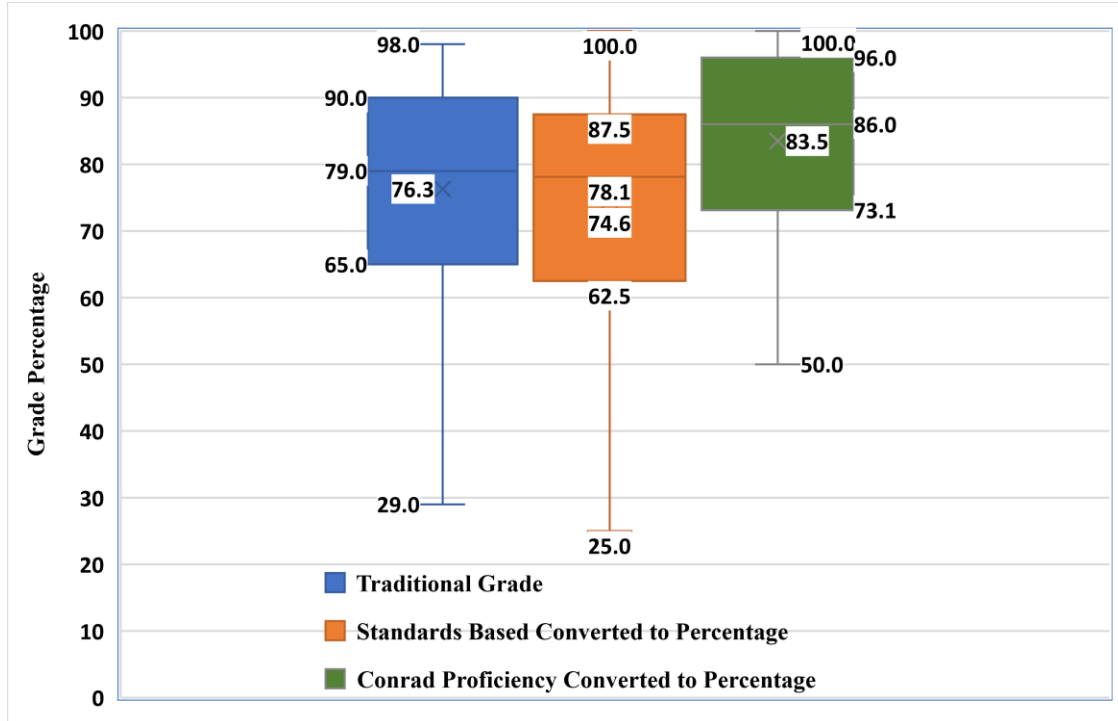


Figure 1. Five assessment units scored using three grading methods, ($N=31$).

The proficiency-based score was shown to be significantly different from the other two scoring methods (traditional grading and standards-based scoring) using ANOVA and t-tests ($p, 0.05$). A t-test comparing traditional grading versus standards-based grading showed no significant difference ($p>0.05$). The proficiency-based versus standards-based and proficiency-based versus traditional grading score showed significant differences ($p, 0.05$).

Qualitative data suggests that students in science class also preferred the proficiency-based scoring. Some comments from the post-interview included “I prefer the proficiency grading because you will have better numbers than the others.” And “Proficiency because I have better marks.” Proficiency-based also allows for students to miss a few questions and still be considered proficient. Comments from students included “I really like the ability to get 1 or 2 questions wrong on the test and still get a 10/10 because you may be close to the answer but not

get it fully right, even though you know the material.” “I feel somewhat on the fence about this. It would feel great missing a few questions yet still getting a 10/10, but I also believe that my score should be put correctly as my teacher thinks I did”. commented another student.

Additional analysis was completed to determine if there was any significant difference between high, average or low achieving students, or boys versus girls. The achievement of the student (high, average, low) was based on MAP testing scores and the comparison of NORM data from those scores. The NORM score for an eighth grader is 209-213, therefore an average student scored in that range, a low achieving student scored below that range and a high achieving student scored above that range.

T-test results showed that there wasn't a significant difference between boys and girls for any of the scoring methods. ANOVA statistical analysis of the high, average and low scoring students did show a significant difference. This was further analyzed with a t-test which showed that the proficiency-based scoring for high, average and low students was significantly different than the traditional grading and standards-based scoring methods. I believe this is because there is a good mix of boys and girls in the different levels, and that the Conrad proficiency-based scores were higher for all the levels of achievement than the standards-bases scores and the traditional grading. The statistical analysis also showed there was no evidence to reject the null hypothesis for traditional versus standards-based scoring.

These results can be seen more clearly in Figure 2, where the mean scores for each of the high, average and low scoring students are shown. The proficiency-based scoring has a higher average score than the average scores for traditional scoring and standards-based scoring.

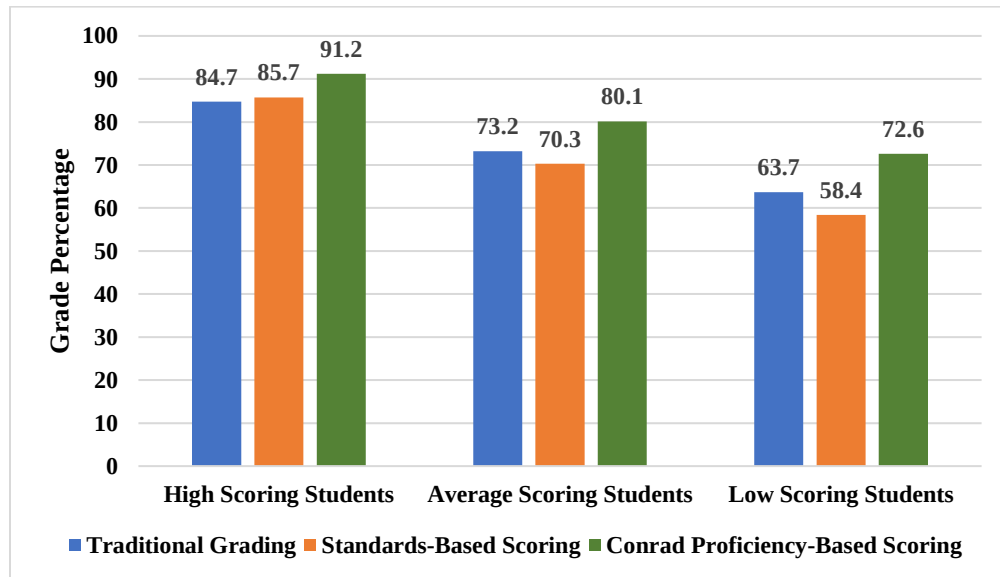


Figure 2. Averages of three scoring methods for high, average and low scoring students, ($N=31$).

Figure 2 shows there is variation in scoring when comparing the proficiency-based scores to the traditional grading method and standards-based scoring method over five units of assessments.

Letter Grades or Percentage Grades

Students also started to change their opinion of needing letter grades and percentage grades which are related to traditional grading (Figure 3). Students went from 71% (22 students) wanting percentage grades to 30% (9 students). This was proven significant using a T-test method ($p, 0.05$). The figure also shows that students started understanding the need for a score per learning target instead of a letter grade overall. This is related to the fact that proficiency-based and standards-based scoring is scored using a rubric based on NGSS and the learning targets built from the standards, where traditional grading is just one grade over the whole unit. One student stated in reference to the proficiency-based system “I like this grading system a little

more because it goes off of how familiar you are to the topic.” Another stated “I like having each learning target graded.”

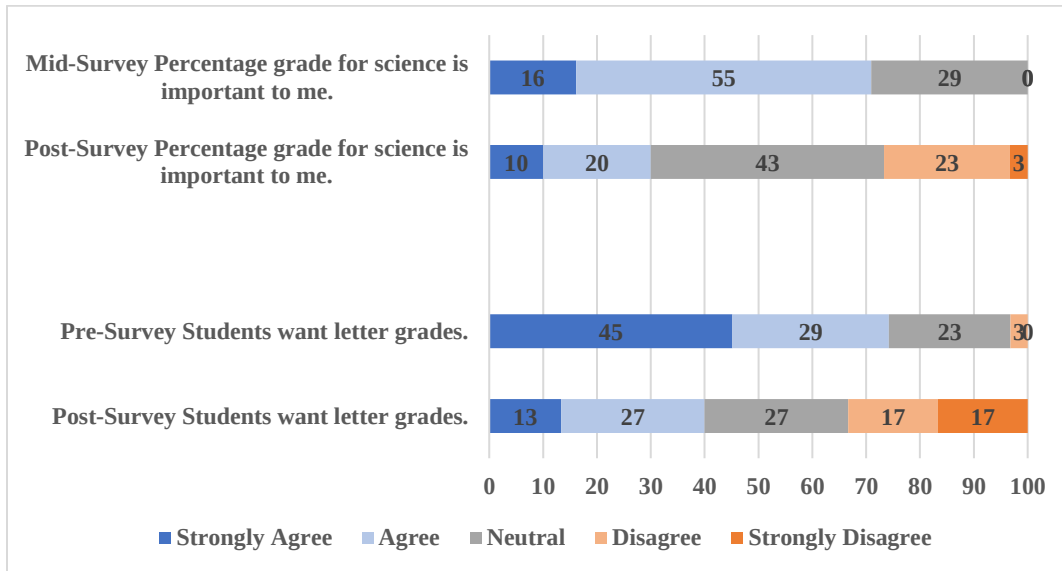


Figure 3. Survey results about percentage and letter grades, (Pre-N=31) (Mid- N=31) (Post-N=30).

The overall higher proficiency-based scores didn't specifically translate to how students scored on standardized testing when taking the science MAP assessment. At Conrad Schools, students in K-10 grade take the science MAP NGSS-aligned assessment at least twice a year, fall and spring. The MAP testing data was compared from fall to spring and the results showed that 58% (18 students) achieved a higher MAP test score from fall to spring versus 42% (13 students) who stayed the same or went down on the scoring. The majority of student improved due to learning new content throughout the school year and having more hands-on laboratory opportunities to practice with the science concept.

Student Perceptions About Grading

During this conversion for students transitioning from traditional grading to a rubric based grading method of proficiency-based or standards-based, I wanted students to see how their own perceptions and experiences changed. The basic preference was that students wanted to know more about their grade than just an overall percentage. Learning targets were used for each lesson and given to the students in the Self-Navigation Tool Appendix D. These learning targets were based on NGSS and used to guide students learning in all formative and summative assessments. Figure 4 identifies some of the survey questions related to learning targets for the students. Students believe that the teacher makes clear what the student is supposed to learn with results from pre to post survey with 90% (28 students) in the pre-survey and 84% (25 students) in the post-survey agreeing. This was guided by the teacher giving students learning targets. Yet, students had mixed feelings about learning targets when surveyed about how they help with studying and getting an individual score per learning target.

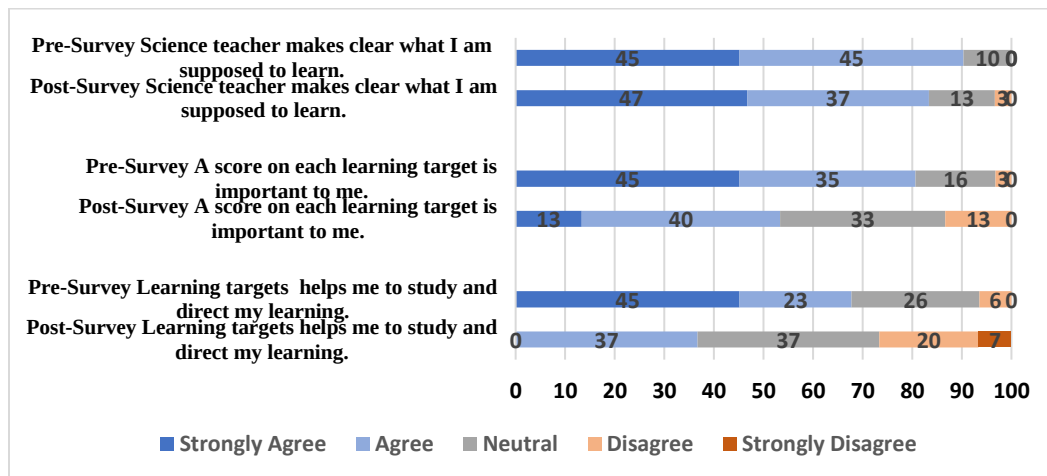


Figure 4. Pre and post survey results related to learning targets, (Pre N=31, Post N=30).

The survey showed that students originally thought that learning targets would be helpful with pre-survey results of 53% (19 students) stating that learning targets helped them study with

only 37% (11 students) post-survey agreeing. Students' perceptions about a score on each learning target also went down from pre-survey of 80% (25 students) to post-survey 43% (16 students). However, some student comments on surveys and during interviews indicated that learning targets do assist students. "I like the grade per learning target. It shows me what I am proficient at or need to study more. It also helps with note taking/ studying before retesting so I know what I need to improve." and "I like the grade per learning target so if you just fail one learning target you don't have to do the whole test all over again you just have to do that one learning target." Another student made this comment which stated they liked re-testing but find that the learning targets complicate things, "i don't like it that much because it is kind of complicated but it works well for retesting because you don't have to retake the whole test you just have to retake the section you failed." Also, students commented positively in surveys and interviews about Conrad's proficiency-based scoring about the grading system and not grading any formative work, like a traditional grading classroom. "I like this grading system a little more because it goes off of how familiar you are to the topic." "I like it. Formative work to me is like practice, where I am aloud to mess one or two things up, instead of stressing to get everything right for a grade." Many students just commented that they "liked it" or "its good".

Student Confidence and Motivation

The third question I wanted to find out about is how learning targets and the new Conrad proficiency-based grading affected students' confidence and motivation in science. Students scored 65% (20 students) on the pre-survey in agreement that learning targets boosted confidence before testing, but only 43% (13 students) were in agreement on the post-survey (Figure 5). Using learning targets to review for the test showed similar results in the mid-survey

71% (22 students) to the post-survey 74% (22 students). I believe this was due to students getting used to using learning targets to direct their learning and putting the class lessons in context of learning targets. Students were able to make connections between the learning targets and the science content discussed in class using a variety of curricular techniques.

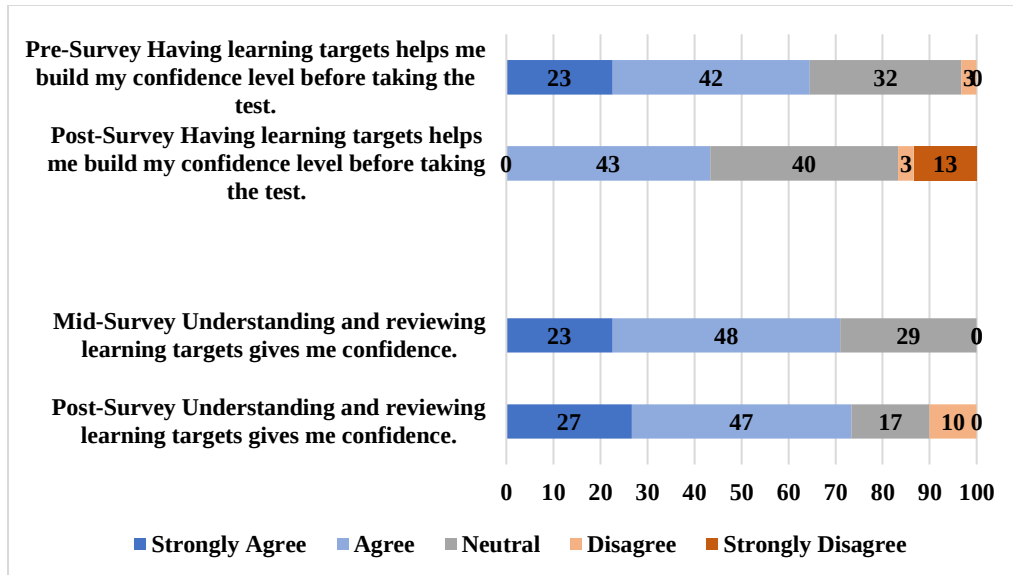


Figure 5. Survey results on student confidence based on using learning targets, (Pre-N=31, Mid-N=31, Post-N=30).

Students had this to say about learning targets and confidence on the post-survey, which related to how they answered the survey. “I think it is about the same level as with out them, but I like that Ms. Tomayer tells us what we need to learn and I can do my own practice and learning with it.” Several students stated that they didn’t use the learning targets or that the learning targets do not change their confidence level. “i don't really use learning targets” and “It doesn't really change my confidence.”

As related to student overall confidence in the pre-survey versus post-survey (Figure 6), the students had the same confidence level at the beginning of the year and end of year about

their ability to do well on science assessments with 71% (22 students) agreeing on the pre-survey and 74% (22 students) on the post-survey.

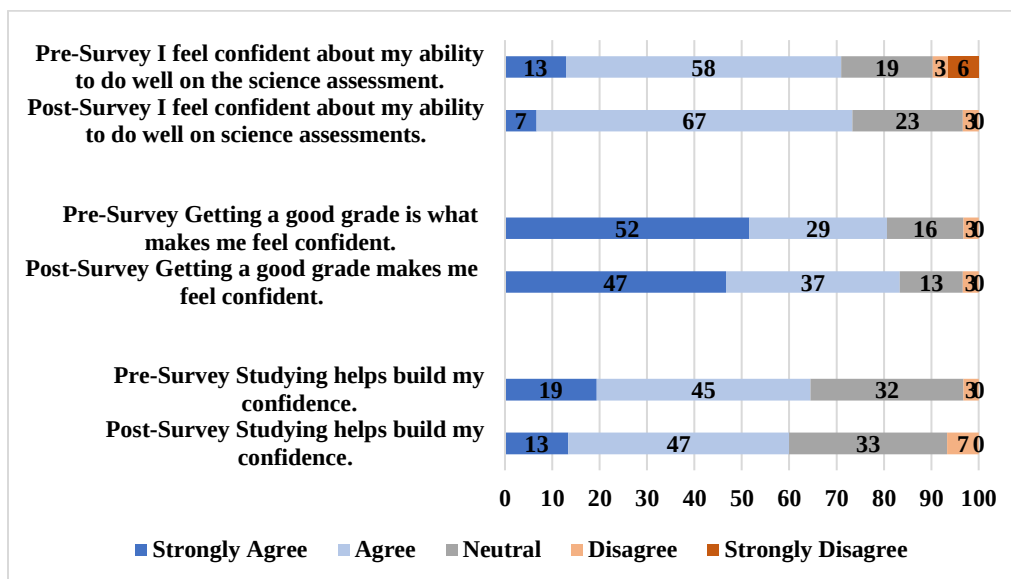


Figure 6. Student confidence levels on assessments and grades, (Pre-N=31, Post-N=30).

The same trend occurred with the question about whether getting a good grade makes students feel confident with 81% (25 students) agreeing on the pre-survey and on the post-survey with 84% agreeing (25 students). Not much changed with the question about whether studying helps build confidence, the pre-survey results showed 64% (20 students) agreeing and the post-survey had 60% (18 students) agreeing.

The comments made by students from the pre-survey about the summative assessments included; “I felt a little less confident than usual on my most recent summative, however after seeing the questions on the test I grew way more confident. I thought the test would be harder than it was.” and “I was pretty confident I didn't study that much but I studied enough to understand the chapter”. In regards to the tools that help students build confidence, comments

included many statements about taking notes and studying at home; “studying and taking notes help me understand the learning target”. Several students mentioned just paying attention in class was enough “I paid attention during class and I now know what I'm doing.” and “Studying, spending time in class learning”.

In the post-survey the several students commented about how the hands-on activities boosted their confidence, “The activities do boost my confidence. When we do them I feel more prepared and ready for the concepts”. “Some of the labs can help demonstrate the things we've learned in class. This is especially helpful to visual learners.” As far as motivation is concerned, students are motivated by several things, including parents, extracurricular activities, their future endeavors, and some even put down food and drink. Comments include; Trying to keep straight A's, to continue doing my sports.” and “I want to succeed and do my best in all that I do. Having good grades is an extremely important part of life, and I take it very seriously.”

CHAPTER FIVE

CLAIM, EVIDENCE, AND REASONING

Claims From the Study

This study can claim that proficiency-based scoring inflated students' overall grades because the proficiency-based scoring allowed students to get some questions incorrect on the learning target and still achieve a higher score. This was demonstrated in the statistical results from the three different scoring methods. The proficiency-based scoring averages were different when compared to traditional grading and standards-based scoring. This can be further demonstrated in the view of six students' scores from the various summative units as seen in Table 3. Students were randomly selected for this table except in there were three boys and three girls selected with one from each MAP assessment academic range. Students do not always show the same achievement levels when taking standardized tests versus classroom assessments. This can be demonstrated by student 3 who earned higher scores on the various classroom assessments, but only earned a low score of 197 on MAP testing, where the NORM score for eighth grade is 209-213. Yet, student 21 scores lower on classroom assessments but higher on MAP assessment with a score of 216. Student 9 is considered average because he scored a 212.

Table 3. Six student scores over five summative assessment units using three scoring methods, girls and boys, low, average, high ranking based on MAP testing scores, ($n=6$).

	Traditional Geological Time	Standards- Based Geological Time All %	Proficiency- Based Geological Time All %	Traditional Natural Selection	Standards-Based Natural Selection All %	Proficiency-Based Natural Selection All %	Traditional Evidence of Evolution	Standards-Based Evidence of Evolution All %	Proficiency-Based Evidence of Evolution All %	Traditional Forces and Motion	Standards-Based Forces and Motion All %	Proficiency-Based Forces and Motion All %	Traditional Mechanical Energy	Standards-Based Mechanical Energy All %	Proficiency-Based Mechanical Energy All %
Student 3 Boy Low	93	88	100	74	75	82	90	88	100	79	78	85	94	88	100
Student 9 Boy Average	79	75	90	61	73	82	71	75	80	69	63	70	88	83	90
Student 21 Boy High	67	50	70	66	63	72	78	75	85	62	47	60	69	71	77
Student 29 Girl Low	52	38	60	46	38	56	54	38	55	55	44	60	38	59	67
Student 2 Girl Average	79	75	80	48	38	58	76	56	75	63	56	68	69	83	87
Student 11 Girl High	98	100	100	92	98	100	83	81	100	85	88	83	94	100	100

Table 3 demonstrates what the statistical analysis found, that the proficiency scores average higher for all students than the traditional grading or standards-based scoring methods. Although, the scores are not dramatically different, for example, a student is not getting a 50% traditional grade and an 80% proficiency-based score overall. Yet, the student could earn a 50%

traditional grade and get an 80% proficiency-based score on one of the learning targets in the unit.

This creates the following challenge for proficiency-based and standards-based grading: how to truly calculate grades. A scale is needed that best describes what a student truly understands in relation to the learning target. A study conducted by Hooper and Cowell (2014) looked at the variation in grades when taking into consideration various grading methods. They compared history adjusted true score, average last two scores, average last three scores, basic average score and power law (Hooper & Cowell, 2014). Each of these methods, when four proficiency scores were calculated, gave various overall scores depending on the method (Hooper & Cowell, 2014). Therefore, some thought needs to go into exactly what type of scoring average should be used and what grading scale should be used to truly demonstrate what a student has learned.

The biggest difference for the students and teacher among the various scoring methods, is that the proficiency-based and standards-based scores are made per learning target and therefore can give the student more information about their grades. The traditional grade is just a score over the whole unit, and doesn't identify what the student specifically got correct or not (Appendix I for a sample student grading information sheet for two assessment units).

Student Perceptions About Grading

Students' perceptions about having to have a percentage grade changed over the treatment. The pre- and post-surveys demonstrated that students started to understand how a different scoring method than just a traditional percentage could help them understand what they learned about the science concept and/or still needed to study. Students viewed their two-unit grade documents (Appendix I) and had these comments; "proficiency, I like how its non pressure" and "Traditional is fine but personally not my favorite. Standards I feel like is kind of vague and would bring my grade down. Proficiency I like most." Another student commented "Traditional score is fine, it is used often so I don't mind it. Proficient scoring is my favorite because it shows how good you are at the subject. I dislike standard scoring." When asked what method they preferred after looking over the various scoring techniques, the majority of students stated that they preferred proficiency-based.

Students also liked the Conrad proficiency-based scoring, because they could get a few questions incorrect and still get a higher score. The student score is related to specific level of understanding of the learning target and not just how many questions were correct out of how many possible questions. Some comments from students included; "I think it is extremely helpful. You might have studied very hard, and had a great understanding on the topic, but you had a minor slip up on the test. Having some leeway is a good system in my opinion." "Having individual grades per learning target is helpful to understand how you're doing in certain topics, and it makes it easier to retest if needed. An overall grade is nice since it is higher numbers, this is because the lower numbers give lower scores. For example, if I get one wrong, that's immediately a 90%. However, due to the system being used currently that focuses on proficiency, that issue is resolved." "I like being able to get a few questions wrong and still get a

10/10 because it shows that one or two mistakes doesn't impact my full understanding of the topic and is fair.”

Proficiency-Based Scoring and Standards-Based Scoring Methods

I believe that part of the reason students dislike standard-based scoring is because they don't truly understand it yet. Proficiency-scoring is a way to get students to start getting used to having a rubric based score per learning target that can lead to standards-based scoring. This is the direction that Conrad School District is heading and I believe that the methods completed in the capstone are assisting students in their understanding of how much more information they can get about what they have learned when using proficiency-based or standard-based scoring.

Also, students prefer several of the curricular options that go with proficiency-based and standards-based learning, including retesting and ungraded grading formative work. In the post-survey 56.7% (17 students) strongly agreed that the ability to retest was important to them, 36.7% (11 students) agreed and 6.7% (2 students) were neutral with no students choosing disagree or strongly disagree. One student stated, “I like that I am able and have the option to retest just a section of a test as opposed to retaking the whole thing. I like that I can just see how well I did on the section as well.” This comment also relates to how the students moved to understanding that they were being graded on each learning target separately and therefore could retest on just the learning targets they didn't score well on.

Retesting is just good practice for not leaving any child behind and making sure that they understand the material in order to move on to more complicated understandings. In some ways, retesting gives students a sense of security if they don't do well the first time on the assessment, but it also can be harmful to students who don't feel like they have to study because they can

retest. This has been seen at Conrad Junior High with students not doing well on the first summative for the unit, and working with the teacher and retesting. The teachers have made adjustments to make it harder for students to retest, by completing additional formative work. This additional formative work does assist the student with understanding in order to retest, but some students would rather study for the first assessment then do the additional formative assignments.

Students quickly adjusted to the grading of only summative work and not the formative or practice work. In the pre-survey 54.8% (17 students) strongly agreed that they understood how formative work and summative assessments were used to determine their grades, with 29% (9 students) agreeing, 9.7% (3 students) neutral and only 6.5% (2 students) disagreeing. Students also agreed with the statement “I understand how formative assignments and summative assessments are used to help me learn.”; with 48.4% (15 students) strongly agreeing, 35.5% (11 students) agreeing, 12.9% (4 students) neutral and only 3.2% (1 student) disagreeing. Students were also easily able to define the meaning of summative and formative with comments; “formative is something that isn't graded in gradebook but summative is something in the grade book” and “formative don't count toward my grade but help me practice for summative assessments which shows if I learned and give me a grade”.

This study showed that in a true proficiency-based scoring system, all students should be summatively assessed twice on each learning target. The first summative assessment would be on just a single learning target, in this case kinetic energy as part of the mechanical energy unit. I also did the same for the learning target on potential energy. This was my original idea, but actually didn't occur until the last unit (mechanical energy) of my research data. The mechanical

energy unit then had three learning targets and therefore three sections on the summative assessment, one on kinetic energy, one on potential energy and one on mechanical energy. It was found that the majority of students increase their proficiency score with a normalized gain 0.55 from the first kinetic energy summative assessment to the second kinetic energy assessment that was part of the unit assessment. This was also seen with the potential energy assessments having a normalized gain of 0.25. Therefore, this is a methodology I would like to continue using; it gives students two summative assessments of the major learning targets and allows for automatic retesting opportunities for students. One student commented on the post-survey, “I like doing the unit tests instead of the big tests on everything. I feel like it helps my overall grade as well as giving me an understanding on specifically what I need to study. The Learning target tests are way better.”

Learning targets were a change this year for students. In the past, I would let the students know what subjects we were studying and why, but didn’t emphasize the ties to NGSS. During a class that I took on standards-based assessment, I developed the self-navigation tools (Appendix D) for students to use to assist them with the test. This tool was not popular with my students which showed in the survey results. The use of learning targets to help with studying for the test, to me, was a great idea, yet the students’ survey results went down from the pre-survey to post-survey (Figure 4). On the post-survey only 36.7% (11 students) agreed that learning targets helped to study and direct their learning, with no students strongly agreeing. Students generally agreed that learning targets helped them to understand what they needed to learn when discussed in class, but most didn’t care about a score per learning target (Figure 4).

Student Confidence and Motivation

Student confidence and motivation also didn't change much throughout the study. Figure 6 shows that most of the survey results stayed the same from pre-survey to post-survey when related to confidence for students. Students also had a variety of comments about what things motivate them to do well. These comments are what I would like all my eighth graders to strive for related to motivation and school work. "Knowing that the information that I'm learning will help in high school and college, maybe even my job." and "I am motivated by the fact that if I get a good grade, I feel accomplished, and that makes me want to do better." This is a goal mindset that can be taught to other students especially if we start that training in middle school. Therefore, increasing student confidence and motivation takes time and is built with trust and respect between students and teachers.

I thought that more students would gain confidence when they scored better on assessments, but the survey responses suggested that wasn't necessarily the case. More students gained confidence in the variety of learning activities given, for example, "The activities do boost my confidence. When we do them, I feel more prepared and ready for the concepts". Science is a unique subject where hands-on experiments give students a different curricular method to obtain information about a topic. This is especially true for middle school students who like to be more active in their learning. "The activities help me a lot on the test because I like to see what were doing hands on and not just from a video or a book."

Overall Grading

Brookhart stated that “Over the past 100 years, research has attempted to identify the different components of grades in order to inform educational decision making” (2016). This includes how teachers’ grade, how classroom grades relate to standardized testing, whether teachers include only academic components or include behavioral components (Brookhart et al., 2016, p. 833). Proficiency-based and standards-based scoring try to take the bias and guesswork out of grading for the teacher by aligning to specific learning targets addressed by standards. Yet, even with a rubric and specific learning target, teachers still have gray areas no matter what type of scoring system they use, for example, the teacher might give a half point here or not, which can overall make a difference. Brookhart (2016) stated “The conclusion is that grades typically represent a mixture of multiple factors that teachers value” (2016).

Value of the Study and Consideration for Future Research

The value of this study is in whether it is the scoring method used, the curricular methods used or the combination of both of those used by the teacher that makes the most impact on a student’s education. To my middle school students, the score really means nothing other than they want to get better grades, an A or B in the class. The students don’t really understand how that grade relates to their overall learning. I am still working on changing those attitudes and getting the students to understand what they are learning and why it is important for their future.

Having learning targets really can help students to identify their areas of strength and weakness when related to a science unit. Although this is something that needs further research and study, it is also something that the teacher needs to emphasize as important to students and assist students with learning how to use the learning targets to study.

Teachers' influence on students' learning is great. I believe our impact doesn't come from how we score a student, but in how we teach the curriculum and our love of the subject that makes a difference. What we say, and how we say things as teachers also makes a difference. We can emphasize the learning target or we can emphasize the grade. I believe that teachers can make a positive change for students in their understanding of what science content they truly understand.

This study showed me that I am also possibly biased to give students higher grades. I used the rubrics to grade, but still want to round up for students. Even if teachers recognize the bias that they can put into a grade, sometimes the grade is still inflated or deflated based on the subjectivity of the test questions (Townesley & Buckmiller, 2020). It might also be because I think that students with higher scores will be motivated to continue to do better on the next assessment and so forth. Yet, this also doesn't do the student any justice when they go into high school and then suddenly are not getting those high grades that they used to. My future grading techniques need to include giving the scoring rubrics to students with the self-navigation tool at the beginning of the lesson, using the rubric in more detail to score students and try not to inflate grades. Yet, I also really like the proficiency-based technique of allowing students to make mistakes on questions, but still be able to score higher on assessments. Therefore, a new balance needs to be found between these concepts and those should be explained to students and I need to continue to evaluate my methods of scoring assessments.

I believe that further study into all three scoring methods is needed to develop the best method to accurately describe student learning. Is it better to assess students on one learning target at a time, give them at least two assessments with each learning target like research says,

to let them get some questions incorrect and still be able to show proficiency? These topics still need further consideration. Also needing further consideration is how schools do report cards and transcripts. Do these documents need to change from showing a percentage score per subject to a list of learning targets and achievement levels of the student? Is there a better way to show what the student is actually accomplishing, understanding, and learning in each subject area? Grades are meant to be an indicator of a student's future success (Brookhart et al., 2016).

Impact of Action Research on the Author

This research has impacted me greatly. I have learned so far that any of the scoring methods could be used with good teaching methods. These scoring methods, although proficiency-based was found to have a higher overall average that didn't compare to standards-based and traditional, are all still within basically the same close range. Therefore, any method of grading works as long as it is explained to the student, the more important things is making the grading method reflect what the students have actually learned.

I believe that the teaching methods make the biggest amount of difference to student learning, not how they are scored on assessments. Being able to retest, not giving students a grade below 50%, doing interventions with students, giving them learning targets, hands-on inquiry activities to learn, using the claim, evidence, reasoning process, demonstrating scientific phenomena, and not scoring formative or practice work are all good techniques to assist students with their learning. I really want my students to be excited to learn about science.

Overall, this action research has already changed my teaching. I make an effort to discuss with students what they understand and am asking more questions to qualify what they really understand about the learning targets. I am also continuing to tie what we are learning to the

specified learning targets so that students are pulling out the correct information. Ties to their lives and how important understanding the content in science is something I am trying to do more of to increase the relevance for my students. Middle school students are naturally curious; therefore, I am trying to get them to express more about what they want to learn and making ties to the subject matter that interests them.

The new teaching methods also help with students wanting to redo or retest when they don't do as well on assessments. My students are starting to understand how the learning targets help them, and how science journals and self-navigation tools are a part of that learning. In their science journals and self-navigation tools, we write the learning targets and then connect those targets to the various curricular activities we do. Why did we do this hands-on laboratory? Making the students make connections with their learning. What is overall most important, the grade or the experience with the learning?

REFERENCES CITED

- American Association for the Advancement of Science. (2023, December 2). Project 2061. <https://www.aaas.org/programs/project-2061>
- American Geosciences Institute. (2023, December 2). NSES: The National Science Education Standards. <https://www.earthsciweek.org/nses>
- Belcher, D. N. (2018, February 28). A Brief History of U.S. Science Education — Leading to Modeling Instruction. <https://ntbelcher.medium.com/a-brief-history-of-u-s-science-education-leading-to-modeling-instruction-3a9bdbd801f2>
- Briggs, A. L., Keeley, P., Ortleb, E., Feathers Jr. PhD, R. M., Fisher, PhD, D., Manga, PhD, M., & Zike, MeD, D. (2020). Inspire science. McGraw-Hill Education.
- Brookhart, S. M. (2011). Starting the conversation about grading. *Educational Leadership*, 69(3), 10-14.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., . . . Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 808-848. <https://doi.org/10.3102/0034654316672069>
- DeBruycker, R. (2023, April 19). Principal. (M. Tomayer, Interviewer)
- Deddeh, H., Main, E., & Fulkerson, S. R. (2010, April). Eight steps to meaningful grading. *The Phi Delta Kappan*, 91(7), 53-58.
- Durm, M. W. (1993). An A is not an A is not an A: A history of grading. *The Educational Forum* (West Lafayette, Ind.), 57(3), 294-297. <https://doi.org/10.1080/00131729309335429>
- Feldman, J. (2022). ROLE CALL: More equitable grading for more meaningful learning. *Principal Leadership* (Middle Level Ed.), 23(3), 8-11.
- Hechinger Report. (2011, January 25). Timeline: Important dates in U.S. science education history. The Hechinger Report: <https://hechingerreport.org/timeline-important-dates-in-u-s-science-education-history/>
- Holden, K., & Retzlaff, S. (2023, February). Equity through grading. *Principal Leadership*, 45-48.
- Hooper, J., & Cowell, R. (2014). Standards-based grading: History adjusted true score. *Educational Assessment*, 19(1), 58-76. <https://doi.org/10.1080/10627197.2014.869451>
- Horak, H. (2023, March 3). High school math teacher. (M. Tomayer, Interviewer)

- Lehman, E., De Jong, D., & Baron, M. (2018). Investigating the relationship of standards-based grades vs. traditional-based grades to results of the scholastic math inventory at the middle school level. *Education Leadership Review*, 6, 1-16.
- Marzano, R. J. (2010). *Formative Assessment & Standards-Based Grading*. Marzano Research.
- Marzano, R. J., Yanoski, D. C., & Paynter, D. E. (2015). *Proficiency Scales for the New Science Standards: A Framework for Science Instruction & Assessment*. Marzano Research.
- McGraw Hill Education. (2020). *Integrated Inspire Science*. Columbus: McGraw Hill Education.
- Mertler, C. A. (2020). *Action Research: Improving Schools and Empowering Educators* (6th ed.). Los Angeles: Sage.
- National Research Council (U.S.). Committee on a Conceptual Framework for New K-12 Science Education Standards. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- Next Generation Science Standards. (2023, December 2). *Improving Science Education Through Three-Dimensional Learning*. Next Generation Science Standards For States By States: <https://www.nextgenscience.org/#:~:text=Standards%20set%20the%20expectations%20for,date%20K%E2%80%9312%20science%20standards>
- Otus. (2024, February 24). *Standards-Based Grading Scales, Calculations, and Conversions*. Otus: <https://otus.com/guides/sbg-grading-scales-calculations-and-conversions/>
- Scriffiny, P. L. (2008). Seven reasons for standards-based grading. *Educational Leadership*, 66 (2), 70--.
- Townsley, M., & Buckmiller, T. (2020). Losing As and Fs: What works for schools implementing standards-based grading? *Educational Considerations*, 46(1), 1-10. <https://doi.org/10.4148/0146-9282.2204>
- Wilcox, J., & Townsley, M. (2022). *Debunking Myths of Standards-Based Grading*. *The Science Teacher* (National Science Teachers Association), 90(1), 29-33.
- Zimmerman, T. (2016, December 20). *Grading for Understanding -Standards-Based Grading*. *The Physics Teacher*, 55, 47-50. <https://doi.org/10.119/1.4972500>

APPENDICES

APPENDIX A

INFORMED CONSENT

September 26, 2023

Dear Parents and Students of Conrad Middle School:

I am currently in my final year of my Masters of Science in Science Education after teaching for 28 years. During this program, I am tasked with completing a research project on a school related topic that interests me. I have decided to complete a project comparing proficiency-based grading against traditional grading. Data will be collected throughout this school year on a variety of formative and summative assessments, as well as through surveys and interviews. Here are my research questions for this project:

1. How do middle school science traditional grades coordinate to standards-based proficiency scale grades and percentage grade when using the same classroom instructional techniques?
2. How does completing smaller assessments over one learning target at a time assist students with understanding the overall standard?
3. Does completing one or two learning targets a week (in smaller steps) increase students' achievement?
4. What do students feel about the change to standards-based grading?
5. How does using standards-based grading improve student motivation and confidence in science learning?
6. How does using standards-based grading impact me as a teacher in methods and grading?

This letter is to inform you about what I am going to be doing in my classroom and how I will be collecting data so that I can get your permission to use data from your middle school child.

Basically, nothing will change from my normal day to day activities; students will still have formative activities, hands-on laboratories, interactive computer simulations, watch video, participate in classroom discussion, and read textbook materials. The main difference is that students will be doing their assessments in smaller chunks throughout the unit, and then at the end of a unit have a larger unit assessment. This will assist students' learning by having them assessed twice over the major content standards, as well as giving the teacher more data to understand areas of proficiency and areas that need to be covered again.

I will be collecting student test scores on pre-assessments and post-assessments and comparing those. Students will also be asked to complete a couple of different surveys on how they feel about grades, the various grading methods, teacher evaluation and curriculum instruction methods. Individual and group interviews may also take place with your child. Northwest Evaluation Association Measure of Academic Progress (MAP) testing data will also be used in this research project. The data will be used to look at student growth and how the assessment scores vary in a traditional system, proficiency-based system and the current science grading scale. This data will assist the teacher and students in understanding how grades are recorded and how the grades relate to learning of standardized content in science. Learning targets are based on the Next Generation Science Standards (NGSS) and are identified with each summative

grade in Infinite Campus. The surveys will be given to students up to four times and students could be interviewed up to two times.

Your child's name will NOT be used in the research project data or writing and your child will remain anonymous. Confidentiality for your child is of the utmost importance; data will be kept secure. If you or your child chooses not to participate, there will be no consequence in grading or otherwise.

This research will be presented in a research paper and presentation in part to obtain a Master's Degree in July 2024 on MSU-Bozeman Campus and will be published by Montana State University-Bozeman. If you have any questions concerning this study or your child's participation please contact me, Monica Tomayer, at monica.tomayer@conradschools.org or call 406-278-3285. If you have any questions on human research generally, please contact irb@montana.edu.

By signing you are giving Monica Tomayer consent to collect formative and summative scores in science, MAP testing data for science, survey your child, interview and record your child, use quotes from your child and take pictures of your child's work. Data will be collected without directly identifying your child except by the researcher (Monica Tomayer). No names will be included in final results.

My student's name is _____.

YES, I give permission (Parent Signature): _____

I have discussed this with my student, and they have also said YES to allowing any data gathered on them to be used in this research project (student signature):

NO, I do not give permission (Parent Signature): _____

APPENDIX B

IRB APPROVAL AND EXEMPTION

Printed By: Tomayer, Monica

11/14/2023 5:46:50 PM

Report Comments

Protocol Information

		Version #
Reference Number:	1015	1
2023-1015-EXEMPT		
Protocol Number:		
Protocol Title:	How do middle school science traditional grades coordinate to standards-based proficiency scale grades and	
Protocol Type:	percentage grade using the same classroom instructional techniques?	
Original		
Tomayer, Monica	Approval Date:	11/14/2023
11/8/2023	Effective Date:	11/14/2023
Tomayer, Monica	Renewal Date:	11/14/2028
Approved	Next Review Date:	11/14/2028
Principal Investigator: Submittal Date: Author: Status:		
Inactive Date:	Expiration Date: 11/14/2028	

Protocol Information 1

Definition of Human Subjects Research 1.1

Only human subjects research studies require IRB review. If you have questions about whether your work qualifies as human subjects research, contact the IRB Office at 406-994-4706 or irb@montana.edu for determination assistance.

Human Subjects Research Qualification 1.2

Does the intent and scope of your proposed study meet the definition of human subject's research?



Yes



No



Unsure - determination assistance requested

Yes	1.2.1
No	1.2.2
Unsure - determination assistance requested	1.2.3
Exempt Review	1.3

Use this form for research studies that will be reviewed at the Exempt level in accordance with 45 CFR 690.104(d).

>Some topics or methods are typically NOT eligible for Exempt Review - See Review Level Determination.

>Do not use this form for research that falls into an Expedited or Full Committee Review (FCR) category. See Types of Review.

>For students: Only certain in-class projects require IRB review. See Student Research and Class Projects.

Confidentiality Statement

APPENDIX C

NATURAL SELECTION AND ADAPTATION SCALE THREE SCORING METHOD
RUBRICS

Natural Selection and Adaptation Scoring Scale

(National Research Council (U.S.). Committee on a Conceptual Framework for New K-12 Science Education Standards, 2012)

MS-LS3-1: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial or neutral effects to the structure and function of the organism.

MS-LS4-4–Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individual's probability of surviving and reproducing in a specific environment.

MS-LS4-5–Gather and synthesize information about the technologies that have changed the way humans influence the inheritance of desired traits in organisms.

MS-LS4-6–Use mathematical representations to support explanations of how natural selection may lead to increases and decrease of specific traits in populations over time.

Competencies:

I can analyze complex natural and designed structures/systems to determine how they function.

I can investigate cause and effect relationships in order to explain the mechanisms driving change.

I can describe cause and effect relationships using probability concepts.

Summative Assessment **

Learning Targets

- I can understand and explain that genes are located on chromosomes inside the nucleus of a cell and during sexual reproduction two copies of that gene are passed to offspring.
- I can model the structure of deoxyribonucleic acid and understand the importance of its parts.
- I can analyze how DNA replicates.
- I can analyze and explain how proteins are made from a gene in DNA with the help of mRNA, tRNA, and ribosomes.
- I can identify different types of mutations and explain how mutations can alter genetic information, proteins and traits in helpful, harmful or neutral ways.
- I can understand that the two genes for every trait that sexual reproduction organisms have, can possibly code for different variations of a trait.
- I can explore and explain the variety of traits in a population of organisms.

- I can explore and explain why species interactions with the environment can affect the traits organisms have, either increase, decrease or remain the same in a population.
- I can explain how natural selection and adaptations can affect a population over time in response to environmental changes.
- I can use mathematical analysis and graphical representations to demonstrate how natural selection can lead to changes in populations over time.
- I can identify, evaluate and explain how human activity can influence the traits of an organism.

Science Scale 2023-2024 School Year--Scoring Method One--Per Learning Target or Standard

10/10	9/10	8/10	7/10	6/10	5/10	0/10
Full Proficiency in answering questions related to subjects with at least 95% accuracy.	Proficient in answering questions related to subject matter with little to no errors in general understanding.	Multiple minor errors in an understanding of the subject matter that don't contradict an overall GENERAL understanding of the subject matter.	Major errors that show the student doesn't understand the general information related to the subject.	Shows minimal understanding of the subject matter by answering very few questions correctly.	Shows minimal understanding of the subject matter by answering less than 50% of questions correctly.	Didn't complete, Missing, didn't answer questions

Proficiency Scale--Natural Selection and Adaptations Scoring Scales Method Two (Marzano et al., 2015,109-110)

MS-LS3-1: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial or neutral effects to the structure and function of the organism.

Advanced	4.0	In addition to score 3.0 performance, the student demonstrates success in content matter beyond the basics of the standard.
-----------------	------------	---

	3.5	In addition to success at the 3.0 level, partial success with some 4.0 content.
Proficient	3.0	Students will be able to use a model to explain and define the general changes in genetic material that may result in the building of different proteins (transcription and translation), which can affect the structure and function of an organism, including the organism's traits. Students can develop a model of DNA and understand its parts. Students understand that mutations can occur, they can describe the types of mutations and possible results of a mutation, harmful, beneficial or neutral effects and how that can affect the structure and function of an organism.
	2.5	No major errors or omissions regarding score 2.0 content, and partial success at score 3.0 content.
Nearing Proficient	2.0	Students will be able to recall the specific vocabulary and use it generally to answer questions, including chromosome, DNA, mutation, gene, genetic material, harmful, helpful, neutral, protein, structure and trait. Students will describe harmful, helpful and neutral effects of mutations. Students can describe the relationship between genes, chromosomes, and proteins.
	1.5	Partial success at score 2.0 content, and major errors or omissions regarding 3.0 content.
Beginning	1.0	With help, limited success at score 2.0 content.

MS-LS4-4—Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individual's probability of surviving and reproducing in a specific environment.

Advanced	4.0	In addition to score 3.0 performance, the student demonstrates success in content matter beyond the basics of the standard.
	3.5	In addition to success at the 3.0 level, partial success with some 4.0 content.
Proficient	3.0	Student will be able to describe genetic variations of traits in a population and understand that the trait will be passed on only when the individual is able to survive and reproduce to pass on traits. Students will be able to use mathematical reasoning to describe genetic variations probability. Students will be able to define and explain natural selection in relation to genetic variations of traits in populations.

	2.5	No major errors or omissions regarding score 2.0 content, and partial success at score 3.0 content.
Nearing Proficient	2.0	Student will be able to recognize or recall the specific vocabulary and use it in a general fashion to answer questions, including genetic variation, natural selection, population, probability, reproduction, traits, survival. Student can describe how genetic variations in a population can increase an individual's chance of surviving and reproducing to pass on the traits.
	1.5	Partial success at score 2.0 content, and major errors or omissions regarding 3.0 content.
Beginning	1.0	With help, limited success at score 2.0 content.

MS-LS4-5—Gather and synthesize information about the technologies that have changed the way humans influence the inheritance of desired traits in organisms.

Advanced	4.0	In addition to score 3.0 performance, the student demonstrates success in content matter beyond the basics of the standard.
	3.5	In addition to success at the 3.0 level, partial success with some 4.0 content.
Proficient	3.0	Students will be able to define, describe and explain about artificial selection and how humans use it to change the probability of traits in various species of organisms. Students will be able to describe what genetic engineering and gene therapy is and how human's use it to modify organisms. Students will be able to explain and defend the various genetic technologies that human's use on Earth and how they have led to various scientific discoveries.
	2.5	No major errors or omissions regarding score 2.0 content, and partial success at score 3.0.
Nearing Proficient	2.0	Students will be able to recognize, recall and generally use the vocabulary including, genetic modification, gene therapy, artificial selection, technology, traits, inheritance, animal husbandry. Students can describe some of the technology's humans use to influence and organism's traits. Students can describe the impacts of these technologies on society.
	1.5	Partial success at score 2.0 content, and major errors or omissions regarding 3.0 content.

Beginning	1.0	With help, limited success at score 2.0 content.
------------------	------------	--

MS-LS4-6–Use mathematical representations to support explanations of how natural selection may lead to increases and decrease of specific traits in populations over time

Advanced	4.0	In addition to score 3.0 performance, the student demonstrates success in content matter beyond the basics of the standard.
	3.5	In addition to success at the 3.0 level, partial success with some 4.0 content.
Proficient	3.0	Students will be able to explain, reason and defend their claims in case studies and mathematical models about natural selection, adaptation and artificial selection. Students will be able to use graphical data to explain trends in changes to populations over time.
	2.5	No major errors or omissions regarding score 2.0 content, and partial success at score 3.0 content.
Nearing Proficient	2.0	Students will be able to recognize, recall and use generally the vocabulary including; adaptation, case study, probability, natural selection, artificial selection, population, behavioral change. Students can describe the relationship between natural selection and population changes in traits over time.
	1.5	Partial success at score 2.0 content, and major errors or omissions regarding 3.0 content.
Beginning	1.0	With help, limited success at score 2.0 content.

Traditional Scale--Grading Method Three

Traditional grade is calculated by taking the student score over the whole summative assessment divided by total number of points, multiplied by 100.

APPENDIX D

STUDENT SELF-NAVIGATION TOOL NATURAL SELECTION AND ADAPTATION

Student Self-Navigation Tool Natural Selection and Adaptation

(National Research Council (U.S.). Committee on a Conceptual Framework for New K-12 Science Education Standards, 2012)

Standards:

MS-LS3-1: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial or neutral effects to the structure and function of the organism.

MS-LS4-4—Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individual's probability of surviving and reproducing in a specific environment.

MS-LS4-5—Gather and synthesize information about the technologies that have changed the way humans influence the inheritance of desired traits in organisms.

MS-LS4-6—Use mathematical representations to support explanations of how natural selection may lead to increases and decrease of specific traits in populations over time

Competency:

I can analyze complex natural and designed structures/systems to determine how they function.

I can investigate cause and effect relationships in order to explain the mechanisms driving change.

I can describe cause and effect relationships using probability concepts.

Student Name:

Unit: Changes Over Time--Natural Selection and Adaptation

Part One:

Learning Targets	Got It	Not Yet	Understanding Level (Beginning, Developing, Proficient, Advanced)	Student Evidence/ Resources	Teacher Notes

I can understand and explain that genes are located on chromosomes inside the nucleus of a cell and during sexual reproduction two copies of that gene are passed to offspring.					
I can model the structure of deoxyribonucleic acid and understand the importance of its parts.					
I can analyze how DNA replicates.					
I can analyze and explain how proteins are made from a gene in DNA with the help of mRNA, tRNA, and ribosomes.					
I can identify different types of mutations and explain how mutations can alter genetic information, proteins and traits in helpful, harmful or neutral ways.					
I can understand that the two genes for every trait that sexual reproduction organisms have, can possibly code for different variations of a trait.					
I can explore and explain the variety of traits in a population of organisms.					
I can explore and explain why species interactions with the environment can affect the traits organisms have, either increase, decrease or remain the same in a population.					
I can explain how natural selection and adaptations can affect a population over time in response to environmental changes.					

I can use mathematical analysis and graphical representations to demonstrate how natural selection can lead to changes in populations over time.					
I can identify, evaluate and explain how human activity can influence the traits of an organism.					

Part Two:

I am good at....
I need additional help with...
I need to practice....
I need to learn or want to learn more about....
I really liked learning about....
Other/Reflection Notes

Part Three:

Goals for Unit:

Part Four:

Extend My Learning	What	Why	How
I can....			
I can...			
I can...			

APPENDIX E

NATURAL SELECTION AND ADAPTATIONS STANDARDS-BASED ASSESSMENT

Natural Selection and Adaptation Post Assessment Tomayer 2024
(McGraw Hill Education, 2020)

Name: _____

Learning Targets and Standard Section 1: MS-LS3-1: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial or neutral effects to the structure and function of the organism.

1. A mutation cannot be inherited.
 - a. True
 - b. False
2. Small changes in traits that cause variation are frequently caused by
 - a. natural selection.
 - b. competition.
 - c. mutations.
 - d. migration.

Mary uses the process of baking a cake as a model to help her friend understand how genes function in people. The ingredients in Mary's cake recipe are: flour, milk, eggs, baking powder, sugar, butter, vanilla. The cake recipe provides directions for how to mix the ingredients together, the temperature of the oven, and the length of time the cake needs to bake. The end result is a plain, yellow cake.

3. Identify what the recipe in Mary's model represents.
4. Define "neutral mutation" and use Mary's model to provide an example.
5. A population of antelope lives on the African Savannah and is hunted by lions. Explain how a mutation that allows some antelope to run faster than others might lead to the evolution of this population.

6. What are the three main parts to a DNA molecule?
 - a. nucleotide, sugar, lipid
 - b. nucleotide, lipid, acid
 - c. lipid, acid, phosphate
 - d. nucleotide, phosphate, sugar
7. Explain the process of transcription and translation in your words.
8. How does the game Telephone (that we played in class), model the types of mutations? Explain.
9. Which of the following is NOT a type of mutation?
 - a. insertion
 - b. deletion
 - c. addition
 - d. substitution
10. All mutations are harmful.
 - a. True
 - b. False
11. Which of the following would be the correct mRNA code from this DNA sequence?
AAGCTTCTG
 - a. TTCGAAGAC
 - b. UUCGAAGAC
 - c. UUGCAACUC
 - d. UUCGUUGUC

Learning Targets and Standard Section 2: MS-LS4-4—Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individual's probability of surviving and reproducing in a specific environment

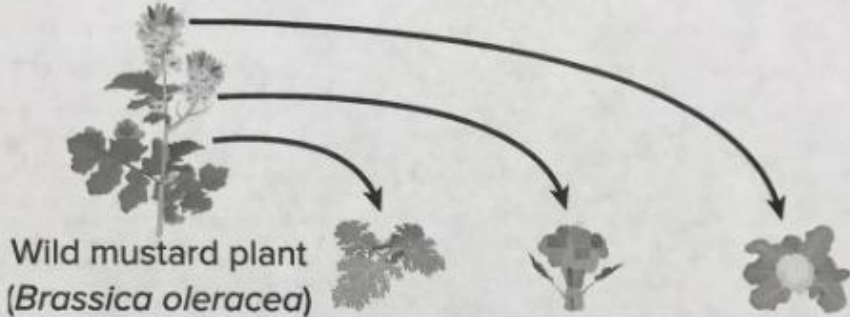
12. Which of the following is a behavioral adaptation?
- a. Blue whales migrating to mate and give birth.
 - b. Tree frogs having long, sticky tongues.
 - c. Alligators' eyes being on top of their heads to help them see above the water.
 - d. Hedgehogs having stiff spines sticking out of their bodies.
13. Which of the following is an inherited trait that increases an organism's chance of survival and reproduction?
- a. adaptation
 - b. mutation
 - c. natural selection
 - d. selective breeding
14. Animals with _____ that benefit them, allowing them to compete better for food, live longer than other animals in the same species.
- a. variations
 - b. gene therapy
 - c. competition
 - d. selection
15. Suppose there is a large population of ptarmigan in Alaska. These birds have mostly white feathers in winter and mostly brown feathers in summer. Now, because of a mutation, a few birds stay white all year. What would happen if the climate changed permanently, and the ground was snow-covered all year?
16. Which of the following would be a structural adaptation for frogs?
- a. Smooth skin has camouflaged colors.
 - b. Frog goes into hibernation during the winter and freezes solid.
 - c. Frog estivates during the summer to keep it cool in the hot sun.
 - d. Frogs hunt on the side of the lake to get more insects.

17. The ability to resemble another species, like the butterfly picture resembling an owl, is called

- a. camouflage.
- b. herding
- c. mimicry
- d. assimilating

Learning Targets and Standard Section 3: MS-LS4-5—Gather and synthesize information about the technologies that have changed the way humans influence the inheritance of desired traits in organisms.

Humans have used selective breeding to produce several types of vegetables from a single wild mustard species. Variations of kale, broccoli, and cauliflower have been produced in this way.



Wild mustard plant (*Brassica oleracea*)

Strain	Kale	Broccoli	Cauliflower
Modified Trait	Leaves	Flower buds and stem	Flower buds

Choose one of the vegetables (kale, broccoli, or cauliflower). Using the chart as guidance, explain how you think humans used selective breeding to produce the vegetable from a population of wild mustard.

Three-Dimensional
COAC
This is an example of artificial selection

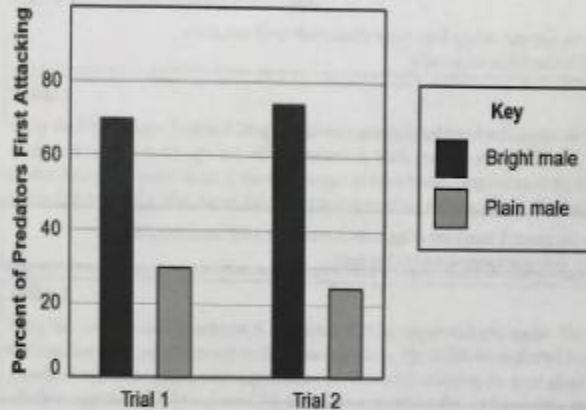
18.

19. A dog breeder prefers very pointy noses for his dogs. So, he breeds dogs with the most pointed noses he can find in hopes of getting puppies with very pointy noses. This process is called
- heredity.
 - genetic engineering.
 - selective breeding.
 - gene therapy.
20. Some insulin is made by _____ in bacteria because it is easier to produce than taking it from a human body.
- selective breeding
 - genetic engineering
 - mimicry
 - genetic manipulation
21. Design and describe a genetically engineered fruit, by combining traits from two different fruits. Critique how your new fruit could actually be developed into a real fruit and what you expect will be good and bad about the new fruit. Determine how that fruit will be made and describe the process.
22. Explain in detail how farmers and/or ranchers in Montana use artificial selection or selective breeding. Define artificial selection.

Learning Targets and Standard Section 4: MS-LS4-6–Use mathematical representations to support explanations of how natural selection may lead to increases and decrease of specific traits in populations over time.

19) The tail fin of a male guppy, a small tropical fish, may be brightly colored or plain. The color variation is an inherited trait. A researcher studied if the cichlid, a predator fish, is more likely to attack male guppies with brightly colored tails than those with plain tails. In Trial 1, the researcher introduced brightly colored guppies and plain guppies, one at a time, into a clear tube in an aquarium containing cichlids. The tube protected the guppies but allowed the cichlids to easily see them. The researcher measured how often the cichlids tried to attack the guppies. In Trial 2, he repeated the experiment but used different guppies and cichlids. The number of each fish was the same in each trial. The graph shows the researcher's results.

Predator Attacks on Male Guppies



The following observations were made of natural ponds where guppies and cichlids interact:

- In natural ponds where cichlids eat guppies, most male guppies' tail fins are not brightly colored.
- In natural ponds where no cichlid predators are present, most male guppies' tail fins are brightly colored.

Explain the probabilities of survival for both brightly colored and plain male guppies in each natural pond. Support your answer with evidence from the results of the researcher's experiments.

23.

24. In a population of moths, the trait of color has two variations---brown and white. The moths live on dark barked trees in a wooded area and are preyed upon by birds. Predict the relative frequencies (number of white and brown moths) of each trait in the population, and explain your reasoning. How might these frequencies change if the trees are replaced with white-barked trees?

25. Complete two of the case studies in the additional sheets handed out by your teacher.

APPENDIX F

PRE-, MID-, AND POST-SURVEY STUDENT ATTITUDES AND PERCEPTIONS ON
GRADING IN SCIENCE CLASS

Pre-Survey—Student Attitudes and Perceptions About Grading in Science Class

Participation in this research is voluntary and will NOT affect student's grades or class standing. Please be as honest as possible for the teacher to gain good information. This survey is part of Monica Tomayer's Capstone project for the 2023-2024 school year.

Research Questions Addressed:

1. What do students feel about the change to standards-based grading?
2. How does using standards-based grading improve student motivation and confidence in science learning?

Using the following scale, please rate the questions.

Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Comments from the teacher about how hard I tried in class are important to me.	5	4	3	2	1
A letter grade for science is important to me.	5	4	3	2	1
I understand what my grade means about my comprehension of the science content.	5	4	3	2	1
My score on assessments matches how well I've studied and learned in class.	5	4	3	2	1
Specific comments from the teacher about what I have learned and retained this year are important to me.	5	4	3	2	1
I understand how formative assignments and summative assignments are used to determine my grade in class.	5	4	3	2	1
I understand how formative assignments and summative assessments are used to help me learn.	5	4	3	2	1
I try my best on all formative assignments so that I can do better on the summative assessments.	5	4	3	2	1
Why did you answer the way you did on the three questions directly above?					
What is the difference between formative assignments and summative assessments?					

I don't always complete the formative assignments.	5	4	3	2	1
My science teacher makes clear what I'm supposed to learn in class.	5	4	3	2	1
A list of the learning targets to be accomplished during a unit are important to me.	5	4	3	2	1
A list of the learning targets helps me to study and direct my learning.	5	4	3	2	1
A score to show how well I've learned each learning target is important to me.	5	4	3	2	1
The ability to be able to re-test is important to me.	5	4	3	2	1
I always study before science tests.	5	4	3	2	1
Studying helps build my confidence before taking an assessment.					
Getting a good score is what makes me feel confident in science class.	5	4	3	2	1
I feel confident about my ability to do well on the science assessments.	5	4	3	2	1
Having learning targets help me build my confidence level before taking the assessment.	5	4	3	2	1

1. Do you understand what your letter grade in science means? Explain.
2. What does your science grade mean to you? Explain.
3. Did you feel confident prior to your most recent summative assessment? Why or why not?
4. What strategies or tools helped you feel confident and prepared for the summative assessments?

Mid-Survey—Grading in Science

Participation in this research is voluntary and will NOT affect student's grades or class standing. Please be as honest as possible for the teacher to gain good information. This survey is part of Monica Tomayer's Capstone project for the 2023-2024 school year.

Research Questions Addressed:

3. What do students feel about the change to standards-based grading?
4. How does using standards-based grading improve student motivation and confidence in science learning?

Using the following scale, please rate your response to the questions.

Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Specific comments from the teacher about what I have learned and retained this school year is important to me.	5	4	3	2	1
I would like specific comments about how well I did on each learning target from the teacher, instead of a grade.	5	4	3	2	1
A letter grade for science is important to me.	5	4	3	2	1
A percentage grade for science is important to me.	5	4	3	2	1
Understanding what I have learned in science is important to me.	5	4	3	2	1
I understand what my grade means about my comprehension of the science content.	5	4	3	2	1
My score on assessments matches how well I've studied and learned in class.	5	4	3	2	1
My science teacher makes clear what I'm supposed to learn in class.	5	4	3	2	1
A list of the learning targets to be accomplished during a unit are important to me.	5	4	3	2	1
A list of the learning targets on the self-navigation tool helps me to study and direct my learning.	5	4	3	2	1
A score to show how well I've learned each learning target is important to me.	5	4	3	2	1
The ability to be able to re-test is important to me.	5	4	3	2	1

I would like to learn more techniques to help me study for science tests.	5	4	3	2	1
Studying helps build my confidence before taking an assessment.	5	4	3	2	1
Understanding and reviewing the learning targets gives me confidence that I will do well on the test.	5	4	3	2	1
Getting a good score is what makes me feel confident in science class.	5	4	3	2	1
I feel confident about my ability to do well on the science assessments.	5	4	3	2	1
Having learning targets helps me build my confidence level before taking the assessment.	5	4	3	2	1

5. Do the learning targets on the self-navigation tool help you with your science learning?
And if so, how?
6. Explain how your science grade is being calculated this year?
7. What do you think about the way your science grade is calculated this school year?
8. How does having learning targets change your confidence level?
9. What types of activities help you learn the science content the best?
10. What motivates you to do well in any school subject?

Post-Survey—Grading in Science

Participation in this research is voluntary and will NOT affect student's grades or class standing. Please be as honest as possible for the teacher to gain good information. This survey is part of Monica Tomayer's Capstone project for the 2023-2024 school year.

Research Questions Addressed:

5. What do students feel about the change to standards-based and proficiency-based grading?
6. How does using proficiency-based and standards-based grading improve student motivation and confidence in science learning?

Using the following scale, please rate your response to the questions.

Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Specific comments from the teacher about what I have learned and retained this school year is important to me.	5	4	3	2	1
I would like specific comments about how well I did on each learning target from the teacher, instead of a grade.	5	4	3	2	1
A letter grade for science is more important to me than understanding what I learned about the science topic.	5	4	3	2	1
A percentage grade for science is more important to me than understanding what I learned about the science topic.	5	4	3	2	1
Understanding what I have learned in science is more important to me than a letter grade.	5	4	3	2	1
I understand what my grade means about my comprehension of the science content.	5	4	3	2	1
My score on assessments matches how well I've studied and learned in class.	5	4	3	2	1
My science teacher makes clear what I'm supposed to learn in class.	5	4	3	2	1
A list of the learning targets to be accomplished during a unit are important to me.	5	4	3	2	1
A list of the learning targets on the self-navigation tool helps me to study and direct my learning.	5	4	3	2	1

A score to show how well I've learned each learning target is important to me.	5	4	3	2	1
The ability to be able to re-test is important to me.	5	4	3	2	1
I would like to learn techniques in science class to help me study science topics.	5	4	3	2	1
Studying helps build my confidence before taking an assessment.	5	4	3	2	1
Understanding and reviewing the learning targets gives me confidence that I will do well on the test.	5	4	3	2	1
Getting a good score is what makes me feel confident in science class.	5	4	3	2	1
Understanding that I can get a question or two incorrect and still score well on the test is important to me.	5	4	3	2	1
I feel confident about my ability to do well on the science assessments.	5	4	3	2	1
Having learning targets helps me build my confidence level before taking the assessment.	5	4	3	2	1

11. What do you think about how your science grade is calculated this year, a grade per learning target, versus an overall percentage grade per science unit?

12. What do you think about being able to get a few questions wrong and still being able to get a 10 out of 10?

13. What do you think about not having formative work being used in the calculation of your final grade in science?

14. How does having learning targets change your confidence level?

15. How do the various activities in science class assist with your confidence level in learning the science concepts?

16. What motivates you to do well in any school subject?

APPENDIX G

EIGHTH GRADE YEAR AT A GLANCE STANDARDS AND LEARNING TARGETS

Integrated 8th Grade Science, 8th Grade and Science (Life, Physical, Earth): Year at a Glance

Information Developed by Monica Tomayer with References/Information from textbook Inspire Science by McGraw Hill Education 2020 www.mheducation.com/preK-12
Competency Statements taken from New Hampshire Nationally Aligned K-8 Science Model Competencies <https://www.education.nh.gov/sites/g/files/ehbemt326/files/inline-documents/2020-04/science-k8-competencies.pdf>

Last updated Oct. 21, 2023

Proficiency or Performance Scales, Assessment Banks, and Student Self-Navigation Tools

Priority standards/Competencies https://opi.mt.gov/LinkClick.aspx?fileticket=woYjYAdWhrs%3d&portalid=182 https://opi.mt.gov/LinkClick.aspx?fileticket=woYjYAdWhrs%3d&portalid=182 https://opi.mt.gov/LinkClick.aspx?fileticket=Yl8OWcHN4uY%3d&portalid=182	I Can Statements/Learning Targets https://docs.google.com/document/d/19zzgKhivPf9uFbTwP9dUYuF-ckqKvynz1R3qpqTp9H5I/edit?usp=sharing
Competencies I can apply the various science measuring tools to get accurate metric measurements. I can demonstrate an understanding and correctly follow all safety rules in the science laboratory. I can correctly measure and convert using the metric system and laboratory equipment. I can understand and use the steps to the scientific method. I can correctly collect data and graph that data with proper graphing knowledge. I can properly light and carefully use a Bunsen burner to complete a laboratory experiment.\	• I can measure correctly using the metric system, length, volume and mass using a variety of science measurement equipment. • I can correctly determine the density of various materials. • I can correctly identify and carefully use the science laboratory equipment. • I can correctly measure utilizing the metric system, volume, mass, length and temperature. • I can understand and identify the parts and order of steps of the scientific method.

Competencies I can develop testable questions, make logical predictions, collect and analyze data, and use specific evidence to interpret and draw conclusions, communicate findings, and develop scientific explanations.	● I can design a science experiment with proper scientific method parts. ● I can identify the variables in an experiment. ● I can use the scientific method and inquiry to do a variety of experiments in science class.
Competencies: I can use graphing techniques to analyze and interpret scientific concepts.	● I can utilize graphs, evaluate data from graphs and correctly graph data from science experiments.
Priority Standards MS-ESS1-4: Construct a scientific explanation based on evidence from rock strata for how the geologic time scale is used to organize Earth's 4.6-billion-year-old history. Competencies I can utilize scientific hypotheses, theories and laws to objectively explore and describe the natural and engineered world, investigate changes over time and revise or reinterpret knowledge based on new evidence. I can observe that the function of natural and designed systems may change with scale.	History of Planet Earth ● I can model, evaluate and explain how scientists determine the age of rock layers and fossils using relative-age dating techniques. ● I can construct a scientific explanation based on evidence from rock strata for how the geologic time scale is used to organize Earth's history. ● I can analyze rock strata and fossil records to learn about the relative ages of important events in Earth's history. ● I can provide evidence that geological processes work today like they did in the past. ● I can identify, describe and evaluate how various processes have changed Earth (mountain formation, ocean basin formation, glaciation, asteroids, volcanic eruptions).

Priority Standards MS-LS3-1: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial or neutral effects to the structure and function of the organism. MS-LS4-4—Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individual's probability of surviving and reproducing in a specific environment. MS-LS4-5—Gather and synthesize information about the technologies that have changed the way humans influence the inheritance of desired traits in organisms. MS-LS4-6—Use mathematical representations to support explanations of how natural selection may lead to increases and decrease of specific traits in populations over time Competencies: I can analyze complex natural and designed structures/systems to determine how they function. I can investigate cause and effect relationships in order to explain the mechanisms driving change. I can describe cause and effect relationships using probability concepts.	Inheritance of traits Variation of traits Natural Selection Adaptation ● I can understand and explain that genes are located on chromosomes inside the nucleus of a cell and during sexual reproduction two copies of that gene are passed to offspring. ● I can model the structure of deoxyribonucleic acid and understand the importance of its parts. ● I can analyze how DNA replicates. ● I can analyze and explain how proteins are made from a gene in DNA with the help of mRNA, tRNA, and ribosomes. ● I can identify different types of mutations and explain how mutations can alter genetic information, proteins and traits in helpful, harmful or neutral ways. ● I can understand that the two genes for every trait that sexual reproduction organisms have, can possibly code for different variations of a trait. ● I can explore and explain the variety of traits in a population of organisms. ● I can explore and explain why species interactions with the environment can affect the traits organisms have, either increase, decrease or remain the same in a population. ● I can explain how natural selection and adaptations can affect a population over time in response to environmental changes. ● I can use mathematical analysis and graphical representations to demonstrate
--	--

	how natural selection can lead to changes in populations over time. ● I can identify, evaluate and explain how human activity can influence the traits of an organism.
Priority Standards: MS-LS4-1–Analyze and interpret data for patterns in the fossil record that document the existence, diversity, extinction, and change of life forms throughout the history of life on Earth under the assumption that natural laws operate today as in the past. MS-LS4-2–Apply scientific ideas to construct an explanation for the anatomical similarities and differences among modern organisms and between modern and fossil organisms to infer evolutionary relationships. MS-LS4-3–Analyze displays of pictorial data to compare patterns of similarities in the embryological development across multiple species to identify relationships not evident in the fully formed anatomy. Competencies: I can analyze and interpret data for past patterns to predict future patterns. I can visualize and model how function depends on the shapes, composition, and relationships among its parts.	Evidence of Common Ancestry and Diversity ● I can provide support evidence that living organisms have changed over time and/or become extinct using fossil records. ● I can formulate conclusions about an organism's evolutionary history by examining fossil evidence. ● I can compare and contrast anatomical structures of organism's living today with fossil evidence of organisms from the past. ● I can compare embryological development of different species to identify relationships.
Priority Standards: MS-PS2-1–Apply Newton's Third Law to design a solution to a problem involving the motion of two colliding objects. MS-PS2-2–Plan an investigation to provide evidence that the change in an object's	Forces and motion Types of interactions ● I can define, identify and understand Newton's three laws of motion.

motion depends on the sum of the forces on the object and the mass of the object. MS-PS2-4–Construct and present arguments using evidence to support the claim that gravitational interactions are attractive and depend on the masses of interacting objects. Competencies: I can model systems and their interactions, including inputs, processes, and outputs. I can demonstrate, using evidence, that a system in dynamic equilibrium is stable due to a balance of feedback mechanisms.	● I can experiment and identify what happens when two objects collide, with both moving or one stationary and one moving in one dimension. ● I can define and demonstrate the various types of motion, constant speed, acceleration, and velocity. ● I can define and demonstrate an understanding of a force (push or pull). ● I can experiment and explain that forces act in pairs and are equal and opposite. ● I can understand and explain the difference between weight and mass and use science equipment to properly find the mass and weight of various objects. ● I can identify and describe the forces acting on an object. ● I can experiment and explain how mass, acceleration, and force interact (Newton’s 2nd Law). ● I can do experiments and explain how measurements of the forces interact and the changes in motion on an object. ● I can understand and evaluate how balanced and unbalanced forces affect an object (inertia) (Newton's first law). ● I can understand and evaluate the difference between speed and velocity and the relationship to acceleration. ● I can understand and define the difference between displacement and distance. ● I can present an argument using evidence that gravity is an attractive force that is affected by mass and distance from the two interacting objects on earth and in space.

Priority Standards: MS-PS3-1–Construct and interpret graphical displays of data to describe the relationships of kinetic energy to the mass of an object and to the speed of an object. MS-PS3-2–Develop a model to describe that when the arrangement of objects interacting at a distance changes, different amounts of potential energy are stored in the system. MS-PS3-5–Construct, use and present arguments to support the claim that when the kinetic energy of an object changes, energy is transferred to or from the object. Competencies: Everfi Unit–Hockey–addresses standards in a gamification module I can find and describe proportional relationships (e.g., speed as the ratio of distance traveled to time taken) among different types of quantities and use the relationship to predict the magnitude of properties and processes. I can demonstrate how the transfer of energy drives the motion and/or cycling of matter within a natural and a designed system.	Definitions of Energy Relationship between energy and forces Conservation of energy and energy transfer • I can differentiate between kinetic and potential energy. • I can experiment and explain about how mass and speed of objects determine the kinetic energy. • I can mathematically calculate the kinetic energy understanding that kinetic energy is proportional to the mass of a moving object and grows with the square of the object's speed. • I can evaluate how the law of conservation of energy relates to potential and kinetic energy in a system. • I can demonstrate the effects of gravity at various heights due to stored potential energy. • I can understand that the height of an object affects the amount of potential energy stored in the object. • I can evaluate the energy flow in a system, what changes the energy goes through. • I can understand and explain that when the motion of energy changes, other energy changes occur.
Priority Standards: MS-PS2-3–Ask questions about data to determine the factors that affect the strength of electric and magnetic forces. MS-PS2-5–Conduct an investigation and evaluate the experimental design to provide evidence that fields exist between objects exerting forces on each other even though the objects are not in contact.	Types of interactions Definition of Energy Relationship between energy and forces • I can experiment and explain magnetic forces and magnets, understanding the poles of a magnet and the repel, and attraction forces.

MS-PS3-2--Develop a model to describe that when the arrangement of objects interacting at a distance changes, different amounts of potential energy are stored in the system. Competencies: I can investigate cause and effect relationships in order to explain the mechanisms driving change. I can interpret and defend my interpretation of the effects of different forms of energy within and between systems (e.g. energy in fields, thermal energy, energy of motion).	● I can experiment and explain electrical forces, understanding how the positive and negative interactions work. ● I can experiment to understand that magnetic and electrical forces are affected by the magnitudes of the charges and the distance between the interacting objects. ● I can understand and explain with evidence that forces act at a distance, including magnetism, electricity, and gravity. ● I can experiment and observe the field around a magnet and around an electromagnetic field. ● I can experiment and explain the basic types of circuits. ● I can investigate how a magnet can move an object without touching the object directly. ● I can evaluate the strength of forces that act at a distance through space and on Earth. ● I can investigate various forces that act through a distance, for example, static electricity, electrically charged devices, and magnets.
Priority Standards: MS-ETS1-1--Define the criteria and constraints of a design problem with sufficient precision to ensure a successful solution, taking into account relevant scientific principles and potential impacts on people and the natural environment that may limit possible solutions. MS-ETS1-2--Evaluate competing design solutions using a systematic process to	● I can design and build an egg carrying device that is within a specific size limit and can keep an egg from breaking at a distance. ● I can determine the effectiveness of my egg drop design using a description of the forces, energy and motion involved in a diagram model. ● I can evaluate the effectiveness of my egg drop design after dropping from a height. ● I can design and build a 2-liter water bottle rocket that will effectively launch a

determine how well they meet the criteria and constraints of the problem. MS-ETS1-3--Analyze data from tests to determine similarities and differences among several design solutions to identify the best characteristics of each that can be combined into a new solution to better meet the criteria for success. MS-ETS1-4--Develop a model to generate data for iterative testing and modification of a proposed object, tool, or process such that an optimal design can be achieved. Competencies: I can apply the engineering design process to optimally improve or solve problems using evidence. I can use functional and structural evidence to develop or improve natural or human-designed structures by taking into account properties of different materials and how materials can be shaped and used.	payload into the air and stay in the air for a specified amount of time and descend using an opening parachute system without destroying the payload. ● I can determine the effectiveness of my bottle rocket design using a description of the forces, energy and motion involved in a diagram model. ● I can evaluate the effectiveness of my bottle rocket design after launching. ● I can successfully build and fly a drone through a specified course. ● I can problem solve the drone build and flight patterns.
Priority Standards: MS-PS4-1–Use mathematical representations to describe a simple model for waves that includes how the amplitude of a wave is related to the energy in a wave. MS-PS4-2–Develop and use a model to describe that waves are reflected, absorbed, or transmitted through various materials. Competencies: I can create models to predict trends and explain patterns in data that support my claims.	Wave Properties ● I can use a model to describe the basic features of a wave. ● I can complete experiments to understand the wavelength, amplitude, frequency and energy of a wave. ● I can understand that a wave that travels through spaces is straight and cannot be made of matter. ● I can understand the principles of light waves (longitudinal waves). ● I can understand the principles of sound waves (compression wave) and that sound waves cannot travel without matter.

	I can experiment with light, sound and water waves to explain about reflection, absorption and transmission through various substances.
Priority Standards: MS-PS4-2–Develop and use a model to describe that waves are reflected, absorbed, or transmitted through various materials. Competencies: I can visualize and model how function depends on the shapes, composition, and relationships among its parts.	Wave Properties Electromagnetic Radiation - I can experiment with light, sound and water waves to explain about reflection, absorption and transmission through various substances. - I can understand that light travels in a straight line except when reflected, absorbed or transmitted. - I can understand that light waves can travel through space and matter waves cannot. - I can experiment to determine that light waves can have differing brightness, color, and frequency.
Priority Standards: MS-PS4-3–Integrate qualitative scientific and technical information to support the claim that digitized signals are a more reliable way to encode and transmit information than analog signals. Competencies: I can obtain, evaluate and communicate evidence about a science concept. (mt)	Information technologies and instrumentation - I can understand the basic principles, purposes and evaluate the difference between digital and analog signals. - I can understand that waves can be used for communication purposes. (fiber optics, radio waves, Wi-Fi, binary patterns, computers).
Priority Standards: MS-ESS3-4–Construct an argument supported by evidence for how increases in human population and per-capita consumption of natural resources impact Earth’s systems.	Human Impacts on Earth Systems I can evaluate and develop solutions to how the growing population of humans affects the consumption of natural resources.

Competencies: I can describe cause and effect relationships using probability concepts.	● I can analyze data about human population ages, use of natural resources and land availability related to population growth to find evidence to support my claim. ● I can discover information about urbanization, agriculture, deforestation and describe how they affect the Earth. ● I can examine and evaluate my personal impact on Earth's systems.
Priority Standards: MS-ESS1-1–Develop and use a model of the Earth-sun-moon system to describe the cyclic patterns of lunar phases, eclipses of the sun and moon and seasons. Competencies: I can create models to predict trends and explain patterns in data that support my claims.	The Universe and Its Stars Earth and the Solar System ● I can develop and use a model to explain the cyclic patterns of the Earth-sun-moon systems. ● I can develop and use a model to explain the changing of the seasons. ● I can describe the motion of Earth and the tilt of the axis and define the difference between rotation and revolution. ● I can explain how the tilt and curve of Earth affects the temperature patterns related to the amount of the Sun's energy received. ● I can develop and use a model to describe the phases of the moon. ● I can develop and explain the various types of eclipses in relation to Earth.
Priority Standards: MS-ESS1-2–Develop and use a model to describe the role of gravity in the motions within galaxies and the solar system.	The Universe and Its Stars Earth and the Solar System

MS-ESS1-3–Analyze and interpret data to determine scale properties of objects in the solar system. Competencies: I can determine an appropriate scale to observe time, space, and energy phenomena using models to study systems that are quite large or small. I can determine the limitations of a model when it represents only certain aspects of the system under study.	● I can develop and use a model to describe the role of gravity in relation to the motion of objects in our solar system. ● I can describe the role of gravity in developing the objects in the universe. ● I can define gravity and understand its effects. ● I can use data to develop a model to demonstrate the size and significance of the various objects in the solar system. ● I can differentiate between meteors, asteroids and comets.
---	--

APPENDIX H

GRADING INTERVIEW STUDENT QUESTIONS

Grading Interview Student Questions.

1. What do you know about grades in school?
 - a. Which grading method did you prefer? Why?
 - b. Do your grades represent what you think is your understanding of the subject?
2. Are grades important to you, why or why not?
3. What do you know about standards-based or proficiency-based grading?
4. What tools or strategies help you to learn?
5. How confident do you feel before a test? Why do you feel that way?
6. What methods help you to study for a test?
7. Do you understand the difference between formative and summative work?

Grading Mid Semester Interview Student Questions

1. What is one thing that would help you to study on a test?
2. Do learning targets help you to study? If so, how? If not, why?
3. What is the self-navigation tool sheet?
4. Do you use the self-navigation tool to study with? Why or why not?
5. Does it help you study when the teacher goes over the self-navigation tool in class?
6. What do you think about the unit summative assessment being two to four learning targets, therefore four different grades in one test?
7. Do you like only having to retest a specific section that you scored lower on the summative assessment, rather than retesting the whole unit assessment?
8. Would you rather have a summative assessment per learning target or continue with unit assessments?

Grading Post Interview Student Questions

1. Based on your data, which grade type(s) do you prefer? Why?
2. In your opinion, what do you think about the grading style traditional grading style? Proficiency? Standards-based?

APPENDIX I

SAMPLE STUDENT GRADING INFORMATION FOR TWO ASSESSMENT UNITS

Student Grading Information 2024 Tomayer's Capstone

Student Name **Sample Student Number 27**

Traditional Score—The total number correct divided by the total number possible multiplied by 100. Covers the whole unit.

Proficiency-Based Score—Conrad Science Score Scale, based on the accuracy of questions answered and additional information given in the short answers. Not all questions must be correct to earn 10/10 points. Scores of 0, 5-10 points possible. Grade per learning target.

Standards-Based Score—Standardized scoring, based on overall understanding of standard and learning targets. Scores of 0-4 possible, 0-novice, 1-beginning, 2-nearing proficient, 3-proficient, 4-advanced.

Forces and Motion Unit

Unit Name/Assessment Name	Traditional Score	Proficiency Score	Standards Based Score
Pre-Test	63	6	2
Full Unit Assessment	46	5.5	1.375
Colliding forces equal and opposite Newtons third law	x	5	1
Force, Mass, Newton's Second Law Acceleration	x	5	1
Gravitational force mass and weight	x	5	1
motion, speed, velocity, inertia, newton's first law	x	7	2.5

Mechanical Energy Unit

Unit Name/Assessment Name	Traditional Score	Proficiency Score	Standards Based Score
Pre-Test	50	x	x
Full Unit Assessment	69	7.3	2.7
Kinetic Energy Mini	27	5	1
Kinetic Energy Unit	x	7	2.5
Potential Energy Mini	58	6	2
Potential Energy Unit	x	8	3
Mechanical Energy Unit	x	7	2.5