Immediate vs. deferred feedback during formative assessments in a 9th grade physical science course

Jordan Sonntag

University of Northern Iowa

Let us know how access to this document benefits you

Copyright ©2021 Jordan Sonntag

Follow this and additional works at: https://scholarworks.uni.edu/grp

Part of the Science and Mathematics Education Commons, and the Secondary Education Commons

Recommended Citation

Sonntag, Jordan, "Immediate vs. deferred feedback during formative assessments in a 9th grade physical science course" (2021). Graduate Research Papers. 1766.
https://scholarworks.uni.edu/grp/1766

This Open Access Graduate Research Paper is brought to you for free and open access by the Student Work at UNI ScholarWorks. It has been accepted for inclusion in Graduate Research Papers by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.
Immediate vs. deferred feedback during formative assessments in a 9th grade physical science course

Abstract
For educators, feedback has become a very important part of the teaching practice that is taught in both teacher education programs and building-level professional development. While many studies have been done on feedback itself, there is much less research on the effectiveness of certain types of feedback used within a Learning Management System (LMS). The goal of this experiment is to determine if there is any difference on summative assessment scores based on the use of different types of feedback on formative assessments using an LMS.

An LMS is a program that allows the educator to create step by step lessons, both formative and summative assessments, automatically provide feedback, and track all the data in one location. The LMS this study will use is called Moodle. The two types of feedback this study will focus on are called immediate and deferred feedback. Immediate feedback allows the student to check their answer immediately to determine if they are correct or incorrect. Deferred feedback requires students to finish the entire formative assessment before they can receive feedback as to whether they are correct or incorrect on all questions. This study aims to find if there is a difference in summative assessment scores if these two types of feedback are used on the formative assessments leading up to the summative assessment.

This open access graduate research paper is available at UNI ScholarWorks: https://scholarworks.uni.edu/grp/1766
Immediate vs. Deferred Feedback During Formative Assessments in a 9th Grade Physical Science Course

Jordan Sonntag

University of Northern Iowa

Presented by

Jordan Sonntag

Spring 2021
This Paper by: Jordan Sonntag

Entitled: Immediate vs. Deferred Feedback During Formative Assessments in a 9th Grade Physical Science Course

has been approved as meeting the non-thesis requirement for the Degree of Master of Science Education

Date: 04/27/21

Dr. Jeff Morgan, Advisor

Date: 04/27/21

Dr. Kyle Gray, Outside Reader
Abstract

For educators, feedback has become a very important part of the teaching practice that is taught in both teacher education programs and building-level professional development. While many studies have been done on feedback itself, there is much less research on the effectiveness of certain types of feedback used within a Learning Management System (LMS). The goal of this experiment is to determine if there is any difference on summative assessment scores based on the use of different types of feedback on formative assessments using an LMS.

An LMS is a program that allows the educator to create step by step lessons, both formative and summative assessments, automatically provide feedback, and track all the data in one location. The LMS this study will use is called Moodle. The two types of feedback this study will focus on are called immediate and deferred feedback. Immediate feedback allows the student to check their answer immediately to determine if they are correct or incorrect. Deferred feedback requires students to finish the entire formative assessment before they can receive feedback as to whether they are correct or incorrect on all questions. This study aims to find if there is a difference in summative assessment scores if these two types of feedback are used on the formative assessments leading up to the summative assessment.
# Table of Contents

**Chapter 1 – Introduction**  
5

**Chapter 2 – Literature Review**  
7
   Formative Assessments and Feedback  
7
   Types of Feedback, Immediate and Deferred  
9
   Learning Management Systems  
11
   Theoretical Foundation  
12

**Chapter 3 – Methodology**  
15
   Participants  
15
   Site  
15
   Learning Objectives and Formative Assessments  
16
   Data Collection Procedure  
18
   Researchers Role  
21
   Materials  
23
   Data Analysis  
23

**Chapter 4 – Results and Discussion**  
25
   Participants  
25
   Data Removal  
25
   Data  
26
      Data for LT1  
26
      Data for LT2  
27
   Discussion  
27
   Limitations  
29
   Further Research  
32
   Conclusion  
34

**References Cited**  
36

**Appendix A: Question samples from LT1**  
41

**Appendix B: Question samples from LT2**  
42

**Appendix C: IRB Letter of Approval**  
47
Chapter 1 – Introduction

For educators, feedback has become a very important part of the teaching practice that is taught in both teacher education programs and building-level professional development. Learning is very limited without the use of feedback or assessments (Race, 2005).

The two types of feedback focused on in this study will be immediate and deferred feedback. While there are different types of feedback, there are also different reasons to use each feedback type. Immediate feedback is when the student gets to immediately see if their answer is correct. If a student learns that they do not understand a concept right away, they can immediately go through a relearning or reteaching process. Oppositely, deferred feedback is when the student receives the feedback at the completion of an assignment, assessment, lesson, etc. This can be used if the educator wants students to finish the learning activity before relearning takes place which may create different learning behaviors during the formative assessment than its immediate feedback counterpart.

There are also different ways in which the student can receive feedback, such as in person or electronically. In this study, a Learning Management System (LMS) was used to deliver the feedback to the students. There are both logistical and educational benefits to using an LMS that helps both the teacher and student.
The existing literature shows benefits to feedback, both immediate and deferred, and the use of an LMS during instruction (Attali & Power, 2010; Gaona, 2018). The combination of these should allow for both growth in student learning and an increase of time available for the educator to help students who may need more intervention. There is a gap in the research this study aims to address in that no research focuses on immediate vs. deferred feedback for high school students using an LMS. Attali (2015b) points out that there is not much research or understanding on asking a student to revise their answer after finding out they were incorrect when using an LMS for learning activities.

The purpose of this study is to determine whether using immediate or deferred feedback during formative assessments has an affect on a student’s summative assessment scores. The educator will continue to teach, answer questions, and help as usual with the independent variable being the type of feedback the students will receive during formative assessments. The goal is to learn if this minor change in the learning process has any effect on learning as measured by summative test scores. The study will take place over a single unit and each student will be given the same summative assessment regardless of the method of formative assessment feedback they were assigned. For the purposes of this study, higher summative assessment scores will be equated with increased learning during formative assessments. This leads to the research question of this study: Does providing immediate or deferred feedback, if either, during formative assessments result in higher summative assessment scores?
Chapter 2 – Literature Review

Formative Assessments and Feedback

More and more, formative assessments are becoming more common classroom practices instead of only summative assessments. Formative assessments give students feedback on their progress as they learn and develop (Hanson et. al, 2001), whereas a summative assessment assesses how much content a student has learned at the end of a learning cycle (American Educational Research Association, 2014).

Formative assessment also has the goal of improving learning for the student and teaching for the educator (Rovai, 2000). Formative assessments are good for determining where each student stands individually in the learning process (Stiggins & Dufour, 2009). For the purposes of this study, it not only tells the teacher how much material the student understands relative to the objective or learning target they are studying, but it also informs the student as to how much material they understand relative to the objective or learning target.

Feedback from formative assessments can improve a student's learning by eliminating the inefficiency of random guessing, allowing students to check their answers to learn from them (Sadler, 1989). Assuming a formative assessment aligns with the learning targets of the educator, this feedback guides the student towards the end goal better than having the student guess on their own without feedback. If educators want to
improve learning, students must be at the center of the feedback process (Merry, 2013). The center of learning refers to students receiving, interacting with, and acting on the feedback to improve their own learning, potentially without assistance from the educator.

For students to receive feedback, they need to attempt activities. These activities are completed by the student in order to, in theory, gain understanding of the concepts and procedures before they can move on to the next subject (Gaona, 2018). For students, having an understanding of whether or not they should move on to the next activity requires feedback. This feedback verifies whether the student’s answer is correct or not to help them reach the correct answer (Kulhavy & Stock, 1989; Driscoll 2007). Feedback on formative assessments not only helps the student understand where they currently are in the learning process but it can also lower a student’s anxiety level, even if the feedback is negative, compared with formative assessments with no feedback (Attali and Power, 2010). If it is assumed students with lower anxiety are more open to the learning process, then offering feedback should improve student learning overall and would be another reason to study the particular effects of specific types of feedback. (This would be something to test in the future as student anxiety is not part of the purview of this study.)

For feedback to be most useful, assessments need to occur throughout each class for the entire semester and be an ongoing process (Benson, 2003). These assessments should be formative in nature as they help the learning process for both student and teacher (McTighe & O’Conner, 2005). Formative “assessment tasks are expected to function as learning events...learners are encouraged to engage in self-assessment on their
path to assuming responsibility for the quality of their own work” (Dick et al. 2005, p. 145). The feedback students receive as to whether they are correct or incorrect will guide their actions as to how they proceed in their learning.

**Types of Feedback, Immediate and Deferred**

The use of questions with feedback may be more effective in motivating students during formative assessments than questions without feedback (Attali, 2015). Both immediate and deferred feedback types used in this study allow for students to interact with a repeatable formative assessment on a deeper level than if the student were to receive no feedback. This feedback with the formative assessments allows students to engage in the learning process more than a static assessment which the teacher would not give results to the student until well after the fact (Attali and Power, 2010). Students actions are guided by feedback as both types of feedback allow the student to know whether they were correct or incorrect (Conrad & Donaldson, 2004). These next actions prompted by the correct or incorrect feedback can be the student revisiting their notes or resources, asking a classmate for clarification, or asking a teacher for help.

In the next paragraphs, Gaona has examined differences between using immediate and deferred feedback and thus is used extensively. It should be stated that Gaona’s research focuses on the college mathematics classroom while this study focuses on a high school classroom.
Immediate feedback is defined to be the student immediately receiving feedback after they have answered whatever question they are being asked. Gaona (2018) defines immediate feedback as “timely, that is, the student does not have to wait for it, as he/she receives it automatically” (pg. 1006). One positive consequence of using immediate feedback was seen when it slightly increased the student performance of those on the lower end of the grading range on summative assessments of collegiate students when compared with students using deferred feedback formative assessments (Gaona, 2018; Attali & Power, 2010). A negative consequence of using immediate feedback emerged. As soon as students answered a question incorrectly, they would voluntarily choose to start over the entire formative assessment in an effort to achieve the desired score (Gaona, 2018). This resulted in wasted time repeating the same material when the student could have found all their deficiencies going through the formative assessment once, determining what they don't know, learning it, and taking it a second time. This resulted in a statistically significant increase in time spent working with immediate feedback formative assessments as compared to deferred feedback formative assessments (Gaona, 2018).

Deferred feedback is defined as the student receiving feedback on their answers after a certain amount of time or a specific milestone has passed. In the case of this study, the student will receive their feedback after they complete the assignment and click ‘submit.’ Just like immediate feedback, students are able to have multiple opportunities to complete the formative assessment, but they do not know which questions they answered
incorrectly until after they have finished their formative assignment. In previous research, deferred feedback produced higher summative assessment scores as compared to immediate feedback amongst collegiate students (Gaona, 2018).

Learning Management Systems

Learning management systems (LMS) are defined as “software applications used by educational systems as an online method to give students access to course content, and for educators to track, develop, and integrate relevant course information” (Cedillos and Wolfe, 2015, pg. 898). LMSs allow educators to create small chunks of content for learners to digest, assess, and review before going on to the next level (Brandl, 2005). Brandl goes on to say “the lesson module allows educators to design lessons that closely control the learning path guiding learners step-by-step, and allowing for advancement only if sufficient mastery has been achieved” (Brandl, 2005, pg. 19).

Gaona and Hardy (2014) determined there was an improvement on paper and pencil assessments after performing activities over the web on an LMS. Aside from this observed improvement, the use of online activities, through an LMS, improved the confidence and learning of the students (Bonham et al. 2003; Jang 2009).

In a traditional classroom, a teacher usually employs a call and response technique to gauge understanding. As Benson (2003) points out, this only allows the first student to give the correct answer an opportunity to receive feedback. Other students are denied the opportunity to test their understanding of the material and be afforded teacher
feedback. This problem can be avoided by using paper and pencil methods such as exit tickets. While beneficial, these paper and pencil methods can be time consuming and don’t focus on where individual students are at a given time. With an LMS, “The computer can track improvements in student performance within the various problems, and thus provide the instructor with performance assessment data” (Reeves, 2000, pg. 108). For the purposes of this study, the students will receive the same assessment data as the instructor, in the form of feedback, in relative real-time depending on the feedback type assigned to the student.

LMSs are able to deliver timely feedback to students, automatically, while requiring little of the teacher so they are free to provide in-person feedback to struggling learners (Benson, 2003). While the educator is capable of giving all feedback, both correctness of an answer and more meaningful feedback of how to proceed through a concept or process, the LMS allows the teacher to focus on students who need more meaningful feedback with the LMS automatically giving feedback to the students who are correct. This meaningful feedback can be both coded into the LMS automatically or it can be done in person. If a student learns they do not understand the material via an incorrect answer on an LMS, either the student or educator can initiate a dialogue that would rectify this misconception.
Theoretical Foundation

This study hypothesizes that students approach content differently depending on the type of feedback they receive, thus changing the way they behave as they move through the material. During immediate feedback, students check to see if their answer is correct before they finish the formative assessment. If the student answers the question incorrectly, they could restart the formative assessment or continue. If the student restarts immediately, they would be learning or receiving assistance one question at a time. Oppositely, in deferred feedback, if students completed the entire formative assessment, they have the potential to learn or receive assistance on a pattern of misunderstandings. While both have the same goal of providing the students feedback, students may react or change their habits differently depending on the timing of that feedback.

The reactions to the type of question are behaviorist in nature in that the behaviors, or reactions to feedback, are learned through interaction with the environment (Myers, 2010). After the stimuli (feedback) is provided, the response the student has may be related to the timing of this stimuli. The student’s behavior may be different depending on whether the student receives immediate and deferred feedback regarding the correctness of their answer(s). The reactions to this feedback will be unique to each student, but still follow Thorndike’s Law of Effect.

Thorndike’s Law of Effect states "responses that produce a satisfying effect in a particular situation become more likely to occur again in that situation, and responses that
produce a discomforting effect become less likely to occur again in that situation” (Gray, 2011, pg. 108–109). The satisfying effect students are looking for is selecting the correct answer at which point the educator assumes the student has understanding of the material. The feedback helps students understand what they do and do not know which might imply that both immediate and deferred feedback will have the same outcome of intrinsically motivating students through the learning. Other motivations could include time to complete the formative assessment, getting the grade, or finishing all tasks to return to a more desirable task (reading, phone use, etc.).

How students react when they receive this feedback is their behavior. Behavior is defined as “the sign or symptom of inner activities, mental or physiological” (Skinner, 1957, p. 213). As discussed in the literature review, there are benefits and drawbacks to each type of feedback. Students could become motivated or frustrated which can result in asking for help or even shutting down. Skinner presented a similar set of outcomes of experimentation when interviewing students. These outcomes were how a student felt they would do, their estimate for if they would succeed, how they felt about the set of reinforcements presented, and their impression of the difficulty of the current problem set (Skinner, 1957). While one type of feedback may have an advantage over the other, these behaviors may be the difference as to why. While the analysis of these behaviors in relation to the type of feedback is outside the scope of this study, it bears mentioning as a potential difference maker.
Chapter 3 – Methodology

Participants

This study took place at Suburban High School in a Midwestern State (SHSMS), pseudonyms chosen to help anonymize the students studied. The students who were part of the study were 9th grade and enrolled in Physical Science. Students in this course were not selected according to any further criteria. All 9th grade students are required to take Physical Science, therefore they were all eligible for this study. All 9th grade students in Physical Science were eligible for their assessment data to be used.

Only one teacher conducted this study with their population of 142 9th grade students. Students who are enrolled in the course, are in 9th grade, and have agreed to participate were a part of the study. Scores on both formative and summative assessments were collected, but no other identifying information about students were included. As the study was primarily 9th graders, the mode age is 14 years old.

In Physical Science, students focused on the physical science standards of the Next Generation Science Standards (NGSS Lead States, 2013). NGSS are the science standards adopted by both the state and district in which the study takes place.

Site

Physical Science is a full year course at SHSMS. The first semester focuses on the physics part of the Physical Science standards from the NGSS while the second
semester focuses on the Earth and Space Science standards from the NGSS. The class is designed so that if a student takes both Chemistry and Biology, this Physical Science class covers the rest of the required NGSS as directed by the state.

This SHSMS has a total population of 1309 for the 20-21 school year. The school is a large school in a city labeled by the National Center for Educational Statistics (2020) as a Large Suburb. The racial/ethnic composition of the school in 2018-2019 was 979 White, 86 Black, 110 Hispanic, 72 Asian, 2 Native/Hawaiian/Pacific Islander, 2 American Indian/Alaska Native, 58 Two or More Races, which is the last time data was available (National Center for Educational Statistics, 2020). The percentage of students eligible for free and reduced lunch at this school district in 2019-2020 was 35%, the last time data was available (educateiowa.gov, 2020). The specific percentage of free and reduced lunch is an approximate value in an attempt to keep the school anonymous.

**Learning Objectives and Formative Assessments**

Learning targets are statements of instructional objectives written from the perspective of the learner. Learning Target 1 (LT1) for the Forces unit is: I can calculate the net force, mass, and acceleration of an interaction using Newton’s second law. In achieving this learning target, students learn how to calculate an unknown using Newton’s 2nd Law, and use that knowledge in the second learning target. An example question students learn to answer might be: “If a force of 14.4 N is applied to an object with a mass of 3.4 kg, what is that object's acceleration?” All questions in this learning
target ask the student to solve for the variable asked for using Newton’s 2nd Law.

Moodle inserts random numbers, and the answer is automatically calculated based on the equation provided by the educator writing the question stems. Moodle then also checks all students’ answers.

Learning Target 2 (LT2) for the Forces unit is: I can analyze a specific situation through Newton’s laws and determine what is occurring. This learning target requires students to use both mathematical and conceptual knowledge to answer questions about Forces in relation to Newton’s Laws. An example question would be: “A 75 kg ice skater and a 110 kg ice skater push off each other on the ice. Which ice skater will accelerate faster? (A) the 75 kg ice skater, (B) the 110 kg ice skater, (C) they will accelerate at the same rate”. Based on the standard, this question assesses the students’ understanding of Newton’s 3rd Law.

The feedback for both immediate and deferred is the same in terms of content. The students were told if they are “correct” or “incorrect” with either a green "check” or a red “x” respectively. The only difference between the two groups being when they receive the feedback from Moodle.

There are three formative assessments the students will be working on before their summative assessment. These formative assessments take place between the learning of each content piece. The first formative assessment is entirely focused on LT1, only asking questions about Newton’s 2nd Law. The second formative assessment is part of LT2 in that it has students calculate net forces based on “situations” as stated in the
learning target, but also requires students to use LT1. The third formative assessment is part of LT2 as it has students use Newton’s Laws to determine what is occurring in “specific situations” as stated in the learning target.

Data Collection Procedure

A Learning Management System (LMS) called Moodle is used throughout the course, and was used to facilitate data collection in the unit under study, Forces. Students took three formative assessments within the Forces unit before they took their summative assessment. Both the formative and summative assessment data were collected for this study. The only change that was made between the two groups is one received immediate feedback and the other received deferred feedback. As the educator had ten sections of the same Physical Science course, there were five sections that received immediate feedback and five sections that received deferred feedback. A coin flip determined which sections will be immediate or deferred feedback. This particular school had A and B days during the Coronavirus pandemic, and all periods that occur on an A day received one type of feedback, while all periods that occurred on a B day received the opposite type of feedback. The data collection is expected to last approximately 3 weeks.

Summative assessments pulled questions from the same teacher-created question bank as the formative assignments. The educator created quizzes, as either formative or summative assessments, for this unit and loaded questions into the quizzes for the students to answer. (“Quizzes” is the name Moodle specifically calls its assessment
Settings can be changed such as time allowed, how many retakes are allowed, the feedback method, whether the student can review the attempt, etc. Moodle also graded each student's attempt and compared it to the educator's determined proficiency percentage to determine if the student is proficient in the content and is ready to move on. Scores for formative and summative assignments were collected automatically as most are short answer questions. Moodle defines “short answer” as anything that has to be typed, but can still be graded automatically. Examples include a word such as “Force”, or a number such as “17.43”. Students were given unlimited attempts on formative assessments and can end those attempts at any point. This was after they finished the formative assessment, or if they answered the first questions and wanted to stop. Lastly, there was a 1 hour timer on the formative assessment for the sole purpose of automatically turning in the assignment. Without this timer, assignments sometimes do not register with Moodle for days. This time limit should not be a problem from the 10 question formative assessments as well as being greater than the length of a class period in this school district.

While formative assignments target a specific skill or proficiency, summative assessments cover a range of skills and proficiencies that are defined by the learning target (standard). Formative assessments only had ten questions, while unit exams had two to three dozen questions to cover all the content from the multiple formative assessments. Summative assessments did not have immediate or deferred feedback. They instead had adaptive feedback which is the most common type of feedback currently used
in the course. Adaptive feedback allowed students two attempts at each question during
the assessment. Students clicked on the check button to see if they are correct or
incorrect. If they were incorrect, Moodle deducted ½ a point and allowed them to
attempt the question again for half credit. If the student was incorrect a second time on
the question, students could no longer receive credit for that particular question.

This was done so that the summative assessment feedback type is the same for
both groups. Immediate, deferred, and adaptive are all used in this classroom. This
means students were familiar with all types of feedback and none are new. By making
the feedback type of the summative assessment the same for both feedback groups, the
summative assessment is no longer a variable between the two groups when determining
the difference in effectiveness between immediate and deferred feedback during
formative assessments.

Students were assumed proficient when they received a 90% in this particular
classroom on formative assessments. This percentage was set by the science department
of this particular school based on learning and professional development taking place
within this district. For simplicity, the formative assessments were scored out of 10
questions, pulled randomly from a large bank of similar questions, meaning students
needed to receive a 9 out of 10 to be considered proficient. Students were aware that
anything under a 9/10 or 90% is considered not proficient and needs to be reattempted to
gain a proficient score. By achieving proficiency, students received a grade within the
gradebook that is reported home to parents.
Students were allowed an unlimited amount of retakes/retries on each formative assessment. Students could start, stop, begin, and end their formative assessments after any number of questions at any time. Only the highest score for formative assessments were recorded in the grade book. As the school where this study takes place does not employ a standards-based grading system, the scores for summative assessments were neither considered proficient nor non proficient. Scores were only reported as a percentage for grading and report card purposes per the policy of the school district.

While the formative assessments could be taken outside of the classroom as they were online, students received in class work time as well. Once proficient (90% or above), the student received a completion grade for that formative assessment. This means on all formative assessments, students could either receive a proficient or a not proficient. While students could receive a retake on their summative assessment if requested, only the first score was examined for this study. The summative assessment was offered to all students at the same time after all labs, activities, and formative assessments had been given proper time within the classroom to be completed. All students were provided an opportunity, in the classroom setting, to take the summative assessment.

**Researchers Role**

The researcher was the students' current teacher. While the teacher taught students in both research groups in the same way and attempted to remove bias wherever
possible, the teacher continued to help any student in need. While the teacher’s intervention has the potential to alter the data, the teacher would help all students in need regardless of whether or not the study took place. To limit potential bias, data was not collected by the teacher. The data was collected with the LMS Moodle automatically, not by the researcher directly. The earned scores on formative assessments and unit exams were used to determine each student’s grade, but Moodle has a feature that allows the grades to be uploaded to PowerSchool (grade management software) automatically. This means the researching teacher had minimal interaction besides clicking the “upload” button. If a student asked about their grade and then asked for help, the teacher will of course then be interacting with the data, albeit on a single student basis and not looking at the pattern as a whole.

As the data collects in the background, the teacher was able to go about their day as usual. The researching teacher’s goal remained, as ever, the achievement of the student. When the time came to gather the data from Moodle, the researcher did not do the pulling of data. The school district has an employee who serves as a technology manager for Moodle. They pulled the scores and removed all identifying information. Each student’s name and any other information was replaced with a number corresponding to when they enrolled in the course, information that was not available to the researching teacher. The Moodle manager also randomized the order of the scores, the order of the students, and then combined the students based on the group that they are a part of in an extra layer of identity protection.
Materials

The materials needed for this study were the LMS Moodle and school-provided computers. This ensured everyone had access to all assessments and no sharing of materials needs to occur. Moodle is housed in a server on the school's property and is usable by any educator that wants to use it. The high school has also provided a one-to-one level of laptops to the students. These laptops are specifically ASUS Chromebooks and are sufficient to run Moodle.

Data Analysis

The data from this study was used to answer the one research question: Does immediate or deferred feedback during formative assessments result in higher summative assessment scores? Formative assessment scores and summative assessment scores were collected. Formative assessment scores were only collected to determine how a student interacted with the stimulus. Students who had not interacted either partially or completely with the stimulus, immediate or deferred feedback, were removed from the study. Partial interaction is defined as only being proficient on one or two of the three formative assessments. This needed to be defined and a line drawn as there are too many possible levels of participation. One student may have answered all 10 questions and received a 2/10. Another student may have chosen to only answer 2 of the 10 questions then stopped working to use their phone. There is no way to tell the difference between
these students, hence, only students who showed proficiency on all three formative assessments were used. Summative assessment scores were used to answer the research question.

The unit exam used in this study contained two separate Learning Targets. the t-test will be calculated independently for each learning target. Samples of the types of questions asked in both the formative and summative assessments for both learning targets are found in Appendix A and B.

A t-test will be used to determine if the differences are significant. If differences are significant, an effect size will be calculated using Cohen’s $d$ (Cohen 1992). Values of Cohen’s $d$ will be used in accordance with values described in A Power Primer (Cohen, 1992). For this research, the effect size values for Cohen’s $d$ will be 0.20, 0.50, 0.80 representing small, medium, and large respectively (Cohen, 1992). This will allow a conclusion about which type of feedback, if any, during formative assessments is more effective and if it is statistically significant.
Chapter 4 – Results and Discussion

Participants

Data for this study was collected in fall of 2020. While data was collected from all 9th grade students, due to coronavirus, only half of students were allowed in the classroom at any given time. This means data was collected over 12 periods instead of 6 periods. It is best to view all 12 periods as individual classes rather than 6 classes divided in half. For the purposes of this study, all 12 periods were treated as individual classes. All formative and summative assessments were given to students on the days that they were in class.

Data Removal

Two students requested to have their data removed from the study. In addition to the two students who requested their data not be used, only 86 of the remaining 142 students achieved proficiency on all formative assessments and therefore had their data used. As this study is looking at the effect of feedback from formative assessments on summative assessment scores, only students who were proficient on every formative assessment were included in the final data set.
Data

Table 1

<table>
<thead>
<tr>
<th>Assessment</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>t-test</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>LT1 Immediate</td>
<td>47</td>
<td>9.48</td>
<td>1.1</td>
<td></td>
<td>.237</td>
</tr>
<tr>
<td>LT1 Deferred</td>
<td>39</td>
<td>9.68</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT2 Immediate</td>
<td>47</td>
<td>8.83</td>
<td>1.3</td>
<td></td>
<td>.349</td>
</tr>
<tr>
<td>LT2 Deferred</td>
<td>39</td>
<td>9.06</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = significant at the p = .05 level

Data for LT1

Immediate feedback students earned a mean score of 9.48 while Deferred feedback students earned a mean score of 9.68. Due to the standard deviations being noticeably different, a t-test with unequal variances was used to determine if there was a significant difference in the means. The t-test was calculated, showing that the difference in the means was not statistically different ($p = .237$). The t-test was calculated for both equal and unequal variances with a minimal amount of difference. No additional tests were needed as the t-test showed there was no statistical difference in the means thus answering the research question for LT 1.
**Data for LT2**

A total of 86 students’ data was used for LT1. Immediate feedback students earned a mean score of 8.83 while Deferred feedback students earned a mean score of 9.06. While the standard deviations were closer, they still showed a difference and a t-test with unequal variances was used. The t-test was calculated, showing that the difference in the means was not statistically different (p = .349). The t-test was calculated for both equal and unequal variances with a minimal amount of difference. No additional tests were needed as the t-test showed there was no statistical difference in the means thus answering the research question for LT 2.

**Discussion**

Both LT1 and LT2 produced null results. This means no claim can be made that either immediate or deferred feedback on formative assessments produces better performance than the other in terms of summative assessment scores. While the research reviewed for this study demonstrated pros for both immediate and deferred feedback, this study cannot claim one form is superior in producing improved summative assessment performance. Immediate feedback was seen to slightly increase the student performance of those on the lower end of the grading range on summative assessments of collegiate students when compared to deferred feedback formative assessments (Gaona, 2018; Attali and Power, 2010). Deferred feedback produced higher summative assessment
scores as compared to immediate feedback amongst collegiate students (Gaona, 2018). This study did not collect sufficient data to make a conclusion as to the students on the lower end of the grading range, as in the Gaona & Attali and Power studies. Also, this study could not replicate the results of the Gaona study in the high school setting. What is interesting is while this study’s data did not show statistically significant differences between summative assessments scores of both groups, deferred feedback mean scores for both LT1 and LT2 were higher than their immediate feedback counterparts.

Even though students were removed from the study based on lack of completion of formative assessments, the data was analyzed for all students from which data was collected for the purposes of this discussion except for the two students that opted out.

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summative Assessment Data with No Removal of Students</strong></td>
</tr>
<tr>
<td>Assessment (All Students)</td>
</tr>
<tr>
<td>LT1 Immediate</td>
</tr>
<tr>
<td>LT1 Deferred</td>
</tr>
<tr>
<td>LT2 Immediate</td>
</tr>
<tr>
<td>LT2 Deferred</td>
</tr>
</tbody>
</table>

* = significant at the p = .05 level

This data is an average of all students summative assessment scores, regardless of completion of formative assessments. As the p-values are larger in the overall average, this could imply that while this study had a null result, the p-values of the study were
significantly smaller than those of the overall average. While this doesn’t prove anything, it should be mentioned as it might help inform future research studies.

**Limitations**

There were many limitations to this study. To start with the most important and unique limitation, this study was conducted during the COVID-19 pandemic. The most notable action this called for was the halving of class sizes to create room for social distancing. Smaller classes had changes to both the usual social and behavioral aspects of daily teaching. This also put the students in a unique position of having to adapt to a new system for every class. Not only is there a new system to education due to the pandemic, but at this particular school district, the modality of learning (remote, hybrid, full in-person) changed multiple times, along with the daily bell schedule changing at least four times throughout the year. While these COVID-19 changes may have affected some students more than others, all students within the study were placed under the same conditions thus probably limiting its effect on the study. It still needs to be mentioned though, as these results could differ from a replicated experiment in a non-COVID-19 year.

In addition to the systematic changes of COVID-19, district policy presented an unexpected challenge. Student to teacher interactions decreased to almost zero. This does not mean the teacher never interacted with the students, but the teacher was 6-feet from the students at all times and therefore could not interact in a normal personal
manner. Short instructional videos were created for students to watch to try to bridge this gap. Collaborative labs and activities were banned by policy at this particular school. Students were requested to not move around the room unless absolutely necessary. All students were distanced 6-feet apart. Lastly, with masks being mandated and the already limited student-to-student interaction, students were extremely hesitant to engage in questioning or any discussion. This manifested in the form of students not engaging in educational dialogue or asking the teacher questions at all. This is starkly different than previous years within this exact same unit where the atmosphere at a minimum was students asking individual questions and at a maximum of full classroom discussion/debate. It also prevented the teacher from approaching students about missing work or being not proficient at specific learning targets outside of the normal electronic communication of email and PowerSchool. During informal conversations with dozens of teachers within this same building, this was common amongst most classrooms and not localized to this particular classroom alone.

These COVID-19 related issues resulted in some very interesting and different reactions in students. As students across the board seemed non-responsive, this resulted in one of two results. The first type of students quietly got to work and finished everything in record time. In the second group, students were confused, had misconceptions, or just didn’t want to work. This resulted in students not completing their formative assessments. Students that did not complete all of their formative assessments were not included in the data analysis. While a comparison of work
completion in a normal year versus a pandemic year would need to be made, this remains a potential limitation on the data analysis of the study.

There are also multiple non-COVID-19 related limitations. Due to how proficiency works in this particular course, a student is not considered proficient on formative assessments until they are at a 90% or above. If the only data used in the analysis comes from students who completed all of their formative assessments, this means the only students analyzed in this study were considered to be proficient on both learning targets heading into the summative assessment. Also, due to time constraints and COVID-19, this study was conducted during Physical Science’s shortest unit of the year which only included three formative assessments and one summative assessment. Lastly, the number of students used in this data analysis isn't even 1/4th of the 9th grade population of this particular high school. While this is not inherently limiting, due to higher level subjects having one section at specific times, students of similar abilities are sometimes filtered into the same section due to scheduling limitations. As sections are assigned at random to teachers, this could have had an affect on the study if this particular study featured mostly higher ability students. The combination of these factors alone could have resulted in the null result of this study.

The most obvious thing to change if this study were to be done again would be to do it in a non-COVID-19 year. While this may or may not be possible, it bears stating as most educational research takes place outside of a pandemic. An increase in the number of students would make the data more reliable. As stated, less than 1/4th of the 9th grade
class was tested as a building, while students are sometimes filtered as previously mentioned. While the higher the number the better, this particular school can at least offer at least 400 incoming 9th grade students. It could be beneficial to ask multiple questions within a single study as many students were eliminated from data analysis as they did not fit the parameters of the research question. This would potentially allow the researcher to look at the data in different ways. An example would be students were removed from data analysis for not being proficient on each formative assessment. Had there been other questions with which to look at those student’s data, their data might not have been thrown out.

**Further Research**

The scope of this study was very limited by design. In carrying out data collection and analysis, more questions presented themselves. This study only looked at summative assessment scores as its dependent variable, but never actually looked at the process of learning between the formative assessments and the students.

Do students actually learn better under one form of feedback over the other? This could potentially be measured by how many attempts or questions it takes students to achieve proficiency. It could be argued if students achieve proficiency in less attempts or questions with one form of feedback over the other, then that form of feedback is better for student learning. If this assumption were to be tested, students could also have the
number of attempts limited rather than unlimited attempts. This would be difficult based on the ethics of education and building/district policy.

Do students change their behavior based on the type of feedback they receive? Based on informal observation in this particular classroom, this can manifest in how students choose to interact with learning activities. If a student receives immediate feedback and answers the first question wrong, they either restart their attempt right away or continue moving forward through the formative assessment. Potentially, students could be divided into groups based on their behavior, such as students that restart or continue after a wrong answer, seek out assistance, try to find the answer on their own through resources, or give up entirely. Students could have their summative assessment scores compared based on behavior.

Some students gave up entirely on formative assessments. Based on this behavior, one could also ask: Does the type of feedback on formative assessments affect completion rate? This can be tracked through Moodle through both attempts and time within each attempt. Moodle collects enough data that a researcher could make some judgments about whether a student “gave up” or not. An example would be if the average time to complete a formative assessment of all students is 20 minutes, and a student only spent 45 seconds within a formative assessment before closing it. A researcher could lay out parameters within the study to be able to track those that “give up” as defined by the study. The researcher could then test whether or not the type of
feedback has an affect on the motivation of students to complete these formative assessments.

**Conclusion**

The goal of this study is simple and narrow in scope. The science department at this particular school uses Moodle heavily to run their courses. The available feedback types are all randomly used based on what each teacher thinks works best in different situations. With this study, the first step could be taken in comparing two available feedback types directly against each other. The research question is: Does immediate or deferred feedback during formative assessments result in higher summative assessment scores for 9th graders?

The results of this study are detailed in the Data and Discussion sections in Chapter 4 of this research. Overall, according to the t-tests calculated for both LT1 and LT2, no statistically significant difference could be attributed ($p = .237$, $p = .349$) to the summative assessment scores of students who received immediate feedback compared with those who received deferred feedback. This leads to a conclusion that students perform essentially the same regardless of which feedback type they were exposed to. The $p$ values of .237 and .349 respectively are much larger than the statistically significant threshold of $p = .05$ set for this study. This means based on the data collected in this study, the research question cannot be answered as written. Neither immediate nor deferred feedback resulted in a higher summative assessment score over the other,
leading to the conclusion that it likely does not matter which type of feedback is employed.
References Cited


Benson, A. D. (2003). Assessing participant learning in online environments. *New Directions for Adult and Continuing Education, 2003*(100), 69–78. [https://doi.org/10.1002/ace.120](https://doi.org/10.1002/ace.120)


*Classroom Management Technique: Ask 3 Then Me.* (n.d.). Retrieved March 4, 2020, from [https://learn.teachingchannel.com/video/ask-3-then-me](https://learn.teachingchannel.com/video/ask-3-then-me)


Hanson, J., Millington, C., & Freewood, M. (2001). *Developing a methodology for online feedback and assessment.* https://repository.lboro.ac.uk/articles/Developing_a_methodology_for_online_feedback_and_assessment/9489737


[https://doi.org/10.1177/003172170909000907](https://doi.org/10.1177/003172170909000907)


[https://doi.org/10.1016/B978-0-08-097086-8.41022-6](https://doi.org/10.1016/B978-0-08-097086-8.41022-6)
Appendix A: Question samples from LT1

A force applied to the object is 4.8 N. The force results in an object accelerating at 7.6 m/s/s. What is the mass of this object?

Answer: 

Check

If a force of 14.4 N is applied to an object with a mass of 3.4 kg, what is that objects acceleration?

Answer: 

Check

If an object of mass 5.3 kg has an acceleration of 27.4 m/s/s, what is the force that is being applied to the object?

Answer: 

Check

An object has a force of 10 N pushing it to the left and a force of 5 N pushing it to the right. What direction will the object start moving?

Select one:
- a. Left
- b. Right
- c. It will not move

Check
Appendix B: Question samples from LT2

If an object has a constant force applied to it, what is true about its acceleration?

Select one:
- a. the acceleration will decrease
- b. the acceleration will be zero
- c. the acceleration will increase
- d. the acceleration will remain constant

Check

When two objects make contact with each other, they exert forces in

Select one:
- a. same directions.
- b. opposite directions.
- c. perpendicular directions.

Check

An object is in motion. It has a force of 100 N pushing it to the right, and a frictional force of 100 N pushing it to the left. What will happen to the speed of the object with these two forces on it?

Select one:
- a. The object will stop
- b. The object will speed up
- c. The object’s speed will not change
- d. The object will slow down

Check

A parent pushes both of their children on a swing set with the SAME amount of force. Child A accelerates faster than child B. Why did child A accelerate faster?

Select one:
- a. Child A has more mass than child B
- b. Child A applied a force on the parent pushing them
- c. Child A has less mass than child B

Check
Which vehicle has more inertia, a 4000 kg semi with a speed of 0 m/s OR a 4000 kg SUV with a speed of 100 m/s.

Select one:
- a. the SUV
- b. the semi
- c. they have the same inertia
- d. not enough information to determine inertia

Check

A 75 kg ice skater and a 110 kg ice skater push off of each other on the ice. Which ice skater will accelerate faster?

Select one:
- a. 110 kg ice skater
- b. they will accelerate at the same rate
- c. 75 kg ice skater

Check

Which object would be the most difficult to accelerate? (Ignore friction)

Select one:
- a. a 70 kg person
- b. a 0.01 kg piece of paper
- c. a 1500 kg rock
- d. a 1000 kg car

Check

An object speeds up to 10 m/s. Once it reaches 10 m/s, its acceleration changes to 0 m/s/s. What will happen to the speed of the object next? (Ignore friction)

Select one:
- a. The object will stop
- b. The object's speed will not change
- c. The object will speed up
- d. The object will slow down

Check

A 75 kg male ice skater and a 75 kg female ice skater push off of each other on the ice. Which ice skater will accelerate faster?

Select one:
- a. 75 kg female ice skater
- b. 75 kg male ice skater
- c. they will accelerate at the same rate

Check
If a bowling ball and a basketball are dropped off the roof at the same time, which will hit the ground first? (Ignore air resistance)

Select one:
- a. They will hit at the same time
- b. Bowling Ball
- c. Basketball

Check

When a gun is fired, the gun pushes backward on the person shooting it. This is called recoil. This recoil is the result of two different forces. The gunpowder explosion of the rifle pushes forward on the bullet. Newton's third law of motion tells us that the bullet pushes backward upon the rifle. The acceleration of the rifle is

Select one:
- a. smaller than the bullet
- b. greater than the bullet
- c. the same as the bullet

Check

Which vehicle has more inertia, a 2000 kg semi with an acceleration of 10 m/s or a 3000 kg SUV with an acceleration of 0 m/s.

Select one:
- a. not enough information to determine inertia
- b. the semi
- c. they have the same inertia
- d. the SUV

Check

Which of the statements below is true if impulse remains constant?

Select one:
- a. If the force decreases, the time decreases.
- b. If the time increases, the force decreases.
- c. If the force increases, the impulse decreases.
- d. If the time increases, the force increases.

Check

A 75 kg ice skater and a 110 kg ice skater push off of each other on the ice. Which ice skater will apply more force?

Select one:
- a. They will both apply the same amount of force
- b. 75 kg ice skater
- c. 110 kg ice skater

Check
The two forces involved in Newton's Third Law

Select one:
- a. act in the same direction.
- b. act at different times.
- c. act in opposite directions.

Check

What will the acceleration be if an object has 0 N of Force applied to it?

Select one:
- a. Negative Acceleration
- b. Positive acceleration
- c. O acceleration

Check

If an object has no acceleration, what can be true about it?

Select one:
- a. All are correct
- b. There are no forces being applied to it
- c. The net force on the object is 0 N
- d. It has a constant speed

Check

A fish is swimming forward in the ocean. Which of the following is accurate of the forces between the fish and the water?

Select one:
- a. The force of the fish's tail points back while the force of the water points forward
- b. The force of the fish's tail on the water is greater than the force of the water on the fish's tail
- c. The forces of both the fish's tail and the water point in the same direction
- d. The force of the fish's tail on the water is less than the force of the water on the fish's tail

Check
Which vehicle has more inertia, a 5000 kg semi with an acceleration of 10 m/s or a 2000 kg car with an acceleration of 0 m/s.

Select one:
- a. they have the same inertia
- b. the semi
- c. not enough information to determine inertia
- d. the car

Check

A 110 kg female ice skater and a 110 kg male ice skater push off of each other on the ice. Which ice skater will apply more force?

Select one:
- a. 110 kg female ice skater
- b. They will both apply the same amount of force
- c. 110 kg male ice skater

Check
Appendix C: IRB Letter of Approval

Research Request Form
Urbandale Community School District
Information / Data Request Form

Exh. A

Code No. 647

Your Name: Jordan Snitker

Title: Peer Grade Reviewer - 12 Grade

Physical Science

Building/Department/District: UHS/Science/UCSD

College/University Attending: University of Nebraska

Purpose of Request:
(Why do you need this information? What question are you trying to answer? What use will the information be used for?)

Masters Program Research. Which is your cohort? Immediate

Who will have access to the information provided?
Myself, Casey Clark, Jeff Morgan (UWO), and Kyle
Gray (UNL)

Detailed Description of Information Needed, Analysis Required or Report Requirements:

(please be as specific as possible)

All that is needed is the formative and summative
assessment score with NO personal identifiers attached!
Analysis will be company form the data to summative
data.

Please attach copies of supporting materials/documents concerning this request.
(e.g. guidelines, instructions, required submission forms, etc. and IRB Form and IRB Approval if applicable)

I verify that the information I have shared will be used in the manner described. The information is
for educational purposes and individual names or identifiers will not be shared with others.

Dr. Teaching & Learning Signature 10/13/20
Superintendent Signature 10/14/20

Note: The researcher making the request shall follow UCSD Policy 647.

For UCSD Use Only:

Approved/Denied: Approved

By: MF

Date: 10/15/2020