

THE EFFECT OF THREE-DIMENSIONAL INSTRUCTION ALIGNED TO THE
NEXT GENERATION SCIENCE STANDARDS ON STUDENT
LEARNING IN HIGH SCHOOL SCIENCE

by

Todd Lawrence Samet

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 2016

©COPYRIGHT

by

Todd Lawrence Samet

2016

All Rights Reserved

TABLE OF CONTENTS

1. INTRODUCTION AND BACKGROUND1

2. CONCEPTUAL FRAMEWORK.....2

3. METHODOLOGY9

4. DATA AND ANALYSIS22

5. INTERPRETATION AND CONCLUSION39

6. VALUE.....53

REFERENCES CITED.....57

APPENDICES61

 APPENDIX A Instructional Plan and Alignment – Treatment Unit 162

 APPENDIX B Instructional Plan and Alignment – Non-Treatment Unit 264

 APPENDIX C Instructional Plan and Alignment – Non-Treatment Unit 366

 APPENDIX D Instructional Plan and Alignment – Treatment Unit 468

 APPENDIX E Photosynthesis/Cell Respiration Pre-Post Test.....70

 APPENDIX F Assessment Alignment Instrument – Treatment Unit 1.....73

 APPENDIX G Laboratory Report and Scoring Rubric – Non-Treatment Unit 275

 APPENDIX H Student Questionnaire – Assessment/Instruction Alignment.....81

 APPENDIX I IRB Exemption84

LIST OF TABLES

1. Classroom Treatment Plan.....	10
2. Data Collection Matrix	17
3. Student Questionnaire–Most/Least Effective Instructional Activities	35

LIST OF FIGURES

1. Normalized Gains – Treatment v. Non-Treatment Pre- and Post-Test.....23

2. Normalized Gain Treatment v. Non-Treatment Units– by Learning Level
and Student Performance Level26

3. Laboratory Report DCI Prompts – Treatment v. Non-Treatment Units28

4. Laboratory Report DCI Prompts – by Student Performance Level29

5. Laboratory Report Scores – Total and SEP Groupings31

6. Laboratory Report Scores – by Cohort32

7. Student Perception of Instructional Alignment – Treatment v. Non-Treatment33

8. Student Perception of Overall Alignment – by Unit.....34

9. Student Perception of Assessment Difficulty38

10. Normalized Gains – Pre/Post Assessments.....39

INTRODUCTION AND BACKGROUND

The Next Generation Science Standards (NGSS) were released in 2013 and have currently been adopted by 16 states and the District of Columbia. These standards differ from the old California State Science Standards – and other standards sets – by moving from primarily content-based standards to performance expectations. Additionally, the NGSS suggest that the best route to meeting these performance expectations lies in a three-dimensional approach to learning, combining by design disciplinary core ideas (DCI), science and engineering practices (SEP), and crosscutting concepts (CCC). This differs from past efforts in the purposeful integration of content and practices. The NGSS propose that content is best learned by engaging in the science and engineering practices, while the practices are best learned in the context of specific content.

This action research focused on student learning when instruction and curriculum have been aligned with the three-dimensional learning model of the NGSS. In the author's school district, we are committed to a science curriculum that emphasizes laboratory-based instruction. However, the science and engineering practices of the NGSS suggest some different approaches to this instructional methodology. If those revisions prove effective, the author's students will have improved learning opportunities for *both* the core disciplinary content and the science and engineering practices. They may also learn these concepts at a deeper level. Consequently, this action research project measured learning at both the acquisition and application/transfer levels to evaluate the depth of student learning.

With the recent adoption of NGSS in California there is a lot of professional activity as teachers prepare for implementing the new standards. In addition to the potential benefit for the author's classroom, this study may also be applicable to on-going work in his site and district that he, as departmental teacher leader, will be facilitating.

Research Questions

The researcher's primary question states:

What is the impact on student learning of revising curriculum and instructional activities to align with the NGSS model of three-dimensional learning?

The researcher's secondary questions are as follows:

- 1) What is the impact of NGSS-aligned revisions on student learning, as measured by acquisition and application of the disciplinary core ideas?
- 2) What is the impact of NGSS-aligned revisions on student learning, as measured by application of the science and engineering practices?

CONCEPTUAL FRAMEWORK

The literature review for this topic focused in three areas. First, papers were reviewed specifically about the NGSS. These included explanations on the general structure of the new standards, as well as the theory and history of their development. Next, were papers on the nature of learning, specifically learning in science. Third, were papers that made links between learning, instruction and assessment in the science classroom. In this last area, there was a focus on references that were recent enough to use the NGSS as a lens through which to view these connections.

Direction for the Work

The nature of this action research study is based on two distinct components, each linked to the NGSS. The first component involves the alignment of curriculum and instruction with the NGSS, especially to the science and engineering practices within the new standards. The second component involves the measurement/assessment of learning as a result of the newly aligned instruction. Krajcik, Coderer, Dahsah, Bayer and Mun (2014) wrote specifically about the alignment issue, including the necessity of integrating the disciplinary core ideas with the science and engineering practices. They note:

The science and engineering practices build on what we know about inquiry to focus on students asking questions or refining problems, investigating and analyzing data, constructing models, and arguing based on evidence to build and refine explanations to understand the world. All three dimensions – DCIs, science and engineering practices and crosscutting concepts – serve as tools to build understanding. When the dimensions are blended and work together, like strands of a rope, learning is stronger (p. 158).

Krajcik, et al., (2014) also emphasize the importance of the performance expectations as better descriptors of student proficiency when compared to content based standards. They indicate that many old content standards are defined by descriptors such as “students will know” and “students will understand” without ever clearly defining what the terms “know” and “understand” mean. The authors emphasize the three-dimensional nature of the NGSS performance expectations, where effective demonstration of expectations requires aspects of the science and engineering practices, the disciplinary core ideas and the crosscutting concepts.

To put these ideas into practice, the authors offer an iterative ten-step process for developing instruction designed to get students to a given performance expectation, or a

related “bundle” of performance expectations. The process reads like an NGSS-specific version of backward planning, similar to Wiggins and McTighe’s (2005) Understanding by Design model. They illustrate their model using middle school performance expectations on the structure of matter (Krajcik, et al., 2014). In the end, the interplay between content and science practices is the crucial concept. The authors conclude:

If we want students to learn the content, they have to engage in the practice. But if we want students to learn the science and engineering practice, then they have to engage in content. Leave one out, and students will not develop proficiency in the other (Krajcik, et al., 2014, p. 159).

In an earlier paper, Krajcik, McNeill and Reiser (2008), describe a similar process of alignment in the context of project-based learning. In this model, which reads like a precursor to the NGSS performance expectations, the authors emphasize the importance of clearly articulated learning goals that incorporate both content and process. They discuss “... learning performances as a useful framework, which go beyond general notions of ‘understanding’ to specify the type of cognitive performance that is desired of learners” (Krajcik, McNeill & Reiser, 2008, p. 24). This idea of a learning performance sounds very similar to the performance expectations found in the NGSS. Further evidence of this linkage is suggested by the following:

Learning performances combine the knowing and doing of science. ... Furthermore, the learning performances allow us to look at the same content across different inquiry practices and the same inquiry practices across different content to create a more complete picture of a student’s understanding (Krajcik, McNeill & Reiser, 2008, p. 24).

In another nod to backwards planning, the authors also stress the importance of careful alignment between goals, learning tasks and assessments. Both papers by the Krajcik group confirm the strength of the three-dimensional NGSS model, suggesting its

validity in this study. Additionally, the linkage to the principles of Understanding by Design validates the authors' use of that approach to instructional design.

Cobern, et al. (2010) conducted a research study comparing the effects of direct and inquiry-based instruction on student learning. Their study attempted to measure the relative effect on student learning of direct instruction when compared with inquiry-based instruction. The authors do a very good job of identifying a number of problems with this type of comparison. They discuss how loose the definitions of inquiry and direct instruction are and, particularly, how many comparisons start with the assumption of one instructional mode as good and the other bad. They state, for example, "A less commendable argument for inquiry, however, is to contrast it with straw man caricatures of alternative modes of instruction, in particular, direct instruction cast as exposition, memorization and cookbook laboratory work" (Cobern, et al., 2010, p. 83). They also note the critical importance of specificity in describing modes of instruction. They replace "direct instruction" with "direct active" instruction, which includes components of practice believed to be critical for any good science instruction. They want to test the difference between direct instruction, done well, and inquiry instruction, also done well. The critical difference, they feel, "...lies in 'how students come to a concept'. This is, do students develop the concepts and principles from exploration, or are they told? This represents the 'active agent' distinguishing inquiry from direct" (Cobern, et al., 2010, p. 84). They also note the difficulty in maintaining the absolute fidelity of each instructional approach as a distinct treatment. In the end, however, they found no statistically significant difference between the two approaches in units on Dynamics and Light, Climate and Seasons. The difficulties they found in separating the effects of these

instructional approaches influenced the author's choice of research questions, ultimately choosing not to include this dichotomy as part of his research design.

Theoretical Framework

The theoretical framework behind the NGSS is built on work that identifies features of learning in general and learning science in particular. At the heart of the NGSS are the performance expectations, which reflect a connection between disciplinary concepts and scientific practice. The performance expectations are meant to elicit deeper learning, while current approaches to deep learning tend to revolve around inquiry-based instruction. Much of this approach can be traced to John Dewey and his work in the early 20th century. Wong, Pugh and the Dewey Ideas Group (2001) present an interesting interpretation of Dewey's influence – a Deweyan perspective in their words – on educational philosophy, particularly in relation to learning and instruction in science. Experience was a big part of Dewey's overall philosophy. According to the authors:

Teachers and researchers in education have often used Dewey's 'experience' as the pedagogical antidote to rote learning; ... Dewey's name has been used to justify hands-on activities, out-of-school learning activities, project-based learning, apprenticeships, and so on because they all purportedly involve learning through experience (Wong, et al., 2001, p. 319).

The authors, all Deweyan scholars, contend that this is a misrepresentation of Dewey's thinking. Their premise, an important semantic interpretation, suggests that Dewey was concerned not with general experience (experience in any form) but with *an* experience and *an* experience possesses specific and important attributes. Most important is that an experience requires anticipation, which helps to organize information and potentially use it in future experiences. In their view:

The central goal of a Deweyan view of education is to help students lead lives rich in worthwhile experiences. The task of the school is to provide students with transformative experiences: experiences that are valuable in themselves and valuable in their potential to lead to other worthwhile experiences (Wong, et al., 2001, p. 322).

Their view is that this cannot happen if the experience is disconnected from subject matter. Here, then, is an educational philosophy that supports the NGSS and its goal of three-dimensional learning that always connects disciplinary core ideas with the science and engineering practices.

Methodologies

The current research questions for this project are based on comparisons of student learning under different instructional treatments. Cobern, et al. (2010) conducted a research project to compare the effects inquiry-based and direct instruction on student learning. Their treatment relied on teachers implementing carefully constructed instructional activities. In order to insure the fidelity of the teaching approach, outside observers watched several of the lessons and rated the teachers' adherence to the lesson plan and overall instructional strategy. This approach is appropriate for the author's action research and was used for the purpose of checking instructional alignment.

Cobern, et al. (2010) also provided useful insights on instruments to measure student learning. Their primary assessment instrument was a pre- and post-test consisting of 24 conceptual multiple choice questions, each with four choices and a three-level confidence indicator. Test items were field-tested to determine validity. The actual questions "... involve application of principles to new situation rather than recall of factual knowledge" (Cobern, et al., 2010, p. 85). The data analysis techniques used standard descriptive statistics to compare raw scores on the assessment instruments. In

addition, performance gains were calculated as normalized gains – that is, “. . .the ratio of gain to maximum possible gain given the pre-score” (Cobern, et al., 2010, p. 87). Using a paired t-test the results were compared for statistical significance, looking at comparisons between program years, different teachers, different topics and different instructional modes. This project used a model of pre- and post-test and a similar approach to the data, using normalized gains and inferential statistics, derived from those instruments.

One of the secondary questions concerns the effect of NGSS alignment on student learning of the science and engineering practices. Lederman, et al. (2014) present a data collection instrument designed for just this question. They have developed and subsequently validated a questionnaire called the VASI (Views About Scientific Inquiry) that provides the ability to collect both pre- and post- instruction data on student thinking in this area. This questionnaire consists of seven open-ended questions, several of which have multiple sub-questions. The questions are designed to uncover student thinking around the eight components identified in the NGSS science and engineering practices. The authors provide detailed rationale for the development of this tool:

Inquiry is typically taught in science classrooms by having students conduct investigations or, in general, by ‘doing’ inquiry, Students can participate in inquiry ‘experiences’ but unless instruction explicitly addresses common characteristics of SI (scientific inquiry), students are more likely to continue to hold naïve conceptions (Lederman, et al., 2014, p. 66).

Their questionnaire has an explicit focus on the student’s understandings about inquiry, in addition to their actions, while doing inquiry work. The questions ask specifically for what a student would do in a given inquiry situation and probes further for explanations about why they would undertake specific actions. It was determined

that this specific instrument was not sufficiently aligned with this research to use in its entirety. The questions did, however, provide a useful starting point and overall structure for the instruments, primarily laboratory reports, used in the author's action research.

A critical aspect of this research is the development of instructional units that are carefully aligned to the NGSS three-dimensional model. Development of effective treatments for this project requires a tool to review current instructional activities for their alignment to the NGSS. The Achieve group, authors of the NGSS, have built a tool for the purpose of evaluating and revising lessons/units. This tool is called the "Educators Evaluating the Quality of Instructional Products" (EQuIP) rubric. Ewing (2015) presents an overview of this tool. "The purposes of the rubric and review process are to: (1) review existing lessons and units to determine what revisions are needed; ..." (Ewing, 2015, p. 13). The author notes that when considering activities/lesson developed before the NGSS, some activities are more aligned than others. She summarizes this as follows: "I have heard this (the rubric) described as a bridge to the NGSS, and that the bridge is longer for some materials than for others" (Ewing, 2015, p. 13). She also indicates that this review process, when performed collaboratively, becomes a value-added proposition. The EQuIP rubric became the basis for building a more targeted tool for evaluating alignment of specific instructional activities, and overall units, with the NGSS three-dimensional approach.

METHODOLOGY

Treatments

The primary focus of this action research is the impact of NGSS-aligned instruction on student learning. The treatment plan alternates instructional units that

have been aligned to NGSS with instructional units that have not been explicitly aligned to NGSS. The aligned units are referred to as treatments and the non-aligned units as non-treatments. All the instructional units were designed to move students through a learning cycle of acquisition, meaning making and transfer. All units used both student-centered and teacher-centered instruction, including some type of laboratory experience in each. Table 1, below, provides a summary of the topics and treatment sequence.

Table 1

Classroom Treatment Plan– Integrated Science Honors (N=54)

Topic	Treatment	Time
Photosynthesis Cell Respiration	NGSS Revised Instruction (treatment)	Oct 2015 3 weeks
Agricultural Practices Food Resources	2014-15 Instruction (non-treatment)	Oct-Nov 2015 3 weeks
Energy Basics – Forms and Transformations	2014-15 Instruction (non-treatment)	Nov-Dec 2015 3weeks
Energy Resources	NGSS Revised Instruction (treatment)	Jan 2016 3 weeks

Treatment Units

The process for developing the treatment units will be described first. This involved revising a unit – from instruction used in previous years – to explicitly align with all three dimensions of the NGSS. The process to align and revise each unit was recursive for both treatment topics, while the actual instructional revisions were specific to each topic.

Developing the treatment units started with the identification of NGSS performance expectations related to the content topic of the unit. The performance expectations (PE) are linked to specific goals from each of the three NGSS dimensions – disciplinary core ideas (DCI), science and engineering practices (SEP) and cross cutting

concepts (CCC). Each of the instructional activities was then evaluated for its alignment to these goals and, subsequently, the entire instructional sequence was evaluated for alignment and gaps. The next step was most critical to the research plan, making necessary revisions to remove existing gaps and improve the overall alignment of the unit to NGSS's three-dimensional structure.

While the revisions were specific to each unit, some commonalities exist. Alignment to the disciplinary core ideas is typically a curricular issue. In the treatment units, it was necessary to remove activities and/or content not related to the performance expectation in question. This included content that went beyond the standards, as articulated by the "clarification statements" and "assessment boundaries" in the NGSS. Content was also added where alignment indicated gaps with the NGSS core ideas. Alignment to the science and engineering practices was more often about revision of instructional practice and strategies. This typically included revising laboratory activities to be more open-ended, in general, and to include more aspects of the science and engineering practices, in particular. Direct instruction activities were similarly revised to be more inquiry-based.

The first treatment unit was on photosynthesis and cell respiration. The instructional plan/alignment instrument is presented in Appendix A. The performance expectations say that students will "Use a model to illustrate how photosynthesis transforms light energy into stored chemical energy" (Achieve, 2012, p. 69). There is a similar performance expectation related to cellular respiration. This PE has additional detail from its clarification statement that reads, "Emphasis is on illustrating inputs and outputs of matter and the transfer and transformation of energy..." while its assessment

boundary states, “Assessment does not include specific biochemical steps” (Achieve, 2012, p. 69). As a result, student readings and class lectures were revised to exclude detailed aspects of the biochemistry that had been included in the past, while material was added emphasizing inputs and outputs.

In the photosynthesis and cell respiration unit, lectures were revised to smaller chunks with the addition of concept mapping between the lectures. This provided students with the opportunity to make meaning of the lecture material in the context of model building, one of the science and engineering practices. The laboratory activity, comparing primary productivity in grass plots exposed and not exposed to sunlight, was also revised to provide students with more opportunity to engage in the practices of mathematical and computational thinking, as well as argument from evidence.

As an additional part of the planning process, assessments were also developed to align with the identified NGSS goals. Both the researcher and a colleague teaching the same course evaluated the alignment of individual assessment items. These assessments, including traditional tests and laboratory-based assessments, were administered as part of the instructional sequence and provided the primary data on student learning.

The second treatment unit focused on energy resources. The instructional plan/alignment instrument is presented in Appendix B. This unit focused on a comparative evaluation of fossil fuels and solar energy as future sources to meet human energy demands. There are two related performance expectations, both from the earth science domain, driving this sequence of instruction. The first indicates that students can, “Evaluate design solutions for developing, managing and utilizing energy resources based on cost-benefit ratios.” (Achieve, 2012, p.81). The second states that students can,

“Evaluate or refine a technological solution that reduces impacts of human activities on natural systems.” (Achieve, 2012, p. 81). As a result of the gaps identified in the instructional alignment analysis, this unit was significantly revised from the unit that was taught in 2014-15. The fossil fuels portion of the unit was changed to an inquiry based jigsaw project in which students evaluated costs and benefits of fossil fuels from the context of the resource use cycle. At the end of this investigatory process, they were asked to create evidence-based arguments for *both* the increased and decreased use of fossil fuels. The background information regarding solar energy was similarly organized through inquiry-based instruction. In this case, the students worked individually through both print and video resources to construct evidence-based arguments both for and against the increased use of solar energy. As a final instructional revision, the collaborating teachers revised the laboratory activity to be a student-designed experiment, with the analysis/lab write-up explicitly linked to the large picture evaluation of the costs and benefits associated with human energy use.

Non-Treatment Units

For the non-treatment units the instruction was implemented as it was taught in the 2014-15 school year. The process for evaluating the non-treatment units was the same as the treatment units, with the exception that the gaps in alignment were identified but *not* addressed. In considering the first non-treatment unit on agricultural practices and food resources, several misalignments were identified. The instructional alignment instrument for the first non-treatment unit is presented in Appendix C. There were two related performance expectations associated with this unit. The first states that students will “Design, evaluate and refine a solution for reducing the impacts of human activities on

the environment and biodiversity” (Achieve, 2012, p. 71). The second says that students will “Evaluate or refine a technological solution that reduces impacts of human activities on natural systems” (Achieve, 2012, p. 81). Instruction on this topic had, in prior years, viewed agriculture and agricultural practices in the context of ecosystems. In particular, prior instruction had focused on agricultural inputs – fertilizers, irrigation, and pesticides – for their role in managing these ecosystems. The NGSS does not include any disciplinary standards that contain this specific content. This alignment gap was not adjusted. Additional misalignments exist in the primary laboratory activity for this unit, an investigation about the effects of light and fertilizer on primary productivity, While this laboratory investigation does include several aspects of the science and engineering practices, it does not align well with the disciplinary content in the identified performance expectation.

The second non-treatment unit was on energy concepts. The instructional alignment instrument for this unit is presented in Appendix D. Two performance expectations were identified for this instructional sequence, both from the physical science domain. The first indicates that students will, “Create a computational model to calculate the change in the energy of one component in a system when the change in energy of the other component(s) and energy flows in and out of the system are known” (Achieve, 2012, p. 64). The second states that students will, “Plan and conduct an investigation to provide evidence that the transfer of thermal energy when two components of different temperature are combined within a closed system results in a more uniform energy distribution among the components in the system” (Achieve, 2012, p. 64). This instructional sequence is generally well aligned with the NGSS. The

application focus of the sequence, with students conducting an experimental lab based on measuring energy conversions and efficiencies, is closely matched with the performance expectations. These energy calculations also align well with the SEP of mathematical and computational thinking. The students engaged with this work both in the introduction to concepts and in the laboratory experience. The experimental work was aligned to both the DCI and the SEP, while also including some engineering design thinking. The major misalignments occurred with the introductory activity, which was not aligned with the DCI, and in the direct instruction components, which were lacking in the SEP. As an additional feature, this unit was strong in addressing the following cross cutting concepts: energy and matter; systems and system models; cause and effect; scale, proportion and quantity.

Research Design

Sample Population

The project's research design used two Integrated Science Honors classes at Redwood High School in Larkspur, California, an affluent suburb of San Francisco. The school has total enrollment of 1661 students. Ethnically, the population is composed of the following: 1372 white students (82.8%); 176 Hispanic students (10.6%); 92 Asian students (5.5%); 75 multiracial students (4.6%); 29 African-American students (1.7%); 13 Filipino students (0.8%); 4 Native American students (0.2%); and 4 Pacific Islander students (0.2%) (California Department of Education, 2015). Additionally, there are 16 students identified as English learners (1.0%) and 84 identified as low-socioeconomic status (5.1%) (Education Data Partnership, n.d.). Integrated Science Honors is a 10th grade course and is the 2nd year of the Tamalpais Union High School District's

graduation requirement for science. The two-year program integrates the biological and earth sciences. The Honors program is open to all students who have performed at a B or better level in the first year of the program. There is no qualifying entrance exam or any other prerequisites. It is designated as a laboratory science course for admission to both the University of California and California State University systems. During the 2015-16 school year, there were 3 sections of Integrated Science Honors, along with 16 sections of non-honors Integrated Science. The Integrated Science course, as a required class, was more closely aligned with the broad scope of NGSS content and, as a result, was the best option from the researcher's teaching load for implementing this action research project.

The two classes had a total of 54 students, 28 males and 26 females. Reflecting the overall population of Redwood High School, there was very little diversity within the classes. There was one student of Asian descent and no Latinos or African-Americans. There were five students for whom English is a second language and all were designated proficient in English. There was one student with an IEP plan, one with section 504 accommodations and two with general education accommodations. All of the students are in their second year of a two-year curriculum that is strongly laboratory based. In particular, the students have significant experience in experimental design and data analysis. As such, while they have not experienced instruction explicitly aligned to NGSS, aspects of their science preparation have been similar to the new standards.

Instrumentation

The overall plan for data collection is summarized in Table 2, below. Each data collection instrument is described and an example of each is included in the Appendices.

Table 2
Data Collection Matrix

Research Questions	Data Collection Instruments				
	Pre- and Post Test	Student Laboratory Reports	Student Questionnaire – Assessment Level and Alignment	Student Interviews – Instruction	Teacher Journal – Notes and Reflection
What is the impact on student learning of revising curriculum and instructional activities to align with the NGSS model of three-dimensional learning?					
What is the impact of NGSS-aligned revisions on student learning, as measured by acquisition and application of disciplinary core ideas?	√	√	√	√	√
What is the impact of NGSS –aligned revisions on student learning, as measured by application of the science and engineering practices?		√	√	√	√

The pre/post test and the laboratory reports are specific to each unit, while the other instruments are common to all units. As with the instructional activities, these assessments were checked for alignment to the NGSS using a common alignment instrument, applied by both the author and a colleague teaching the same course.

The pre- and post-instruction test was designed to address two of the research questions. The first states: What is the impact of NGSS-aligned revisions on student learning, as measured by acquisition of disciplinary core ideas? The second asks: What

is the impact of NGSS-aligned revisions on student learning, as measured by application of disciplinary core ideas? These content-based assessments were administered at the start of instruction and then again at the completion. The author commonly used pre-assessments as part of instruction so students viewed this approach as normal procedure.

The pre- and post-tests were designed with a consistent structure. Each test includes 15-20 multiple-choice questions to measure acquisition of the DCI. Additionally, each test includes three or four short answer questions. These are scored on a three-point rubric and are designed to measure application of the DCI. It is important to note that the analysis of student lab reports will permit further analysis of application level learning. The pre- and post-test from the photosynthesis and cell respiration unit is found in Appendix E.

In an effort to ensure validity, a colleague –who also teaches this course – and the researcher subjected each of the pre/post assessments to an independent alignment analysis. In this process each individual assessment item was evaluated for alignment to specific DCI that were part of the performance expectations. Additionally, each item was evaluated for its conceptual level – acquisition, meaning-making, application – and its overall difficulty level – on a scale of one through four. An example of this instrument, from the first unit, is found in Appendix F. The collaborating teachers compared independent analysis, paying careful attention to the identification of aligned standards. The two teachers had already worked on developing the assessment, so the alignment analysis demonstrated close agreement. Some items were revised where the interpretation of aligned standards differed. The difficulty levels assigned to each item and to the overall assessment were also compared for general agreement between the

evaluators. After the fact, the difficulty level scores from the teachers were compared to those assigned by the students in their post-assessment questionnaire. Difficulty levels across the different assessments were also compared for overall consistency. For example, in evaluating the assessment from the first treatment unit, the researcher and his colleague scored average difficulty levels— based on individual analysis of each item— of 2.74 and 2.92 respectively. The students’ average for overall difficulty, from the student questionnaire was 2.77. This close agreement among the three measures indicated the assessment’s overall validity. The other units, both treatment and non-treatment, produced comparable results.

To evaluate reliability, the scoring distributions from the assessments were evaluated for mean, standard deviation and normal distribution. Again, general consistency in these measures served as an indicator of assessment reliability. For example, in comparing Unit 3 – non-treatment— and Unit 4 –treatment— post instruction assessments, the percentage results were 0.813 ± 0.143 SD and 0.789 ± 0.107 SD respectively. This similarity in standard deviation and the generally normal distribution of the results was assumed to indicate assessment reliability.

After administering the post-instruction assessment, normalized gains were calculated for each student – at both acquisition and application levels – and those values were analyzed by standard descriptive statistics. Inferential statistics, in the form of the t-Test, were also used to compare treatment to non-treatment results.

The student laboratory report addresses two of the project’s research questions. The first states: What is the impact of NGSS-aligned revisions on student learning, as measured by acquisition and application of disciplinary core ideas? The second is: What

is the impact of NGSS-aligned revisions on student learning, as measured by application of science and engineering practices? These written reports are based on a laboratory experience that requires the collection and analysis of data. These laboratory activities often include aspects of experimental design as well. The laboratory reports require students to manipulate, present and analyze their data, while connecting experimental observations and results to the core disciplinary content. The prompts are aligned with specific aspects of the SEP and the DCI, permitting evaluation of application level learning on both of those NGSS dimensions. The SEP that are evaluated include: ask questions; design and carry out experiments; analyze and interpret data; mathematical and computational thinking; constructing explanations; argument from evidence. The lab reports are scored with a four-point rubric for each of the prompts. The lab report prompts and scoring rubric for the Agricultural Practice and Food Resources unit lab are included in Appendix G.

The Student Questionnaire – Assessment Level and Alignment is presented in Appendix H. It was administered to gain student perspective on the alignment between instruction and assessment, as well as to gauge student perception on the difficulty of the assessment questions. These were used to further inform the research questions listed for the pre/post test and the laboratory report.

The questionnaire was administered to all students immediately following each post-assessment, allowing the students to reflect on relationships between the assessment and prior instruction. The individual items are Likert-style questions, with the opportunity to provide open-ended explanations behind their perceptions.

The first question measured the difficulty level of the content-based assessments. In this study, randomly selected treatment groups were not possible so the treatment plan alternated between treatment and non-treatment using the same group of students. The efficacy of such a study rests largely with the consistency of the assessments across the four different topics. This question helped establish that reliability between assessments.

Questions two through five provided evidence, from the student perspective, about the alignment and value of the different instructional strategies used in the unit. The generic form of question asked: “Rate how well the [in-class/lab/homework-reading activities] prepared you for this assessment.” The categories available for response were as follows: fully prepared; most prepared; mostly not prepared and not prepared. The last two open-ended questions provided qualitative data, adding detail to the analysis of the value questions from above. These questions were evaluated using both mean and frequency distribution for the strategies identified as most effective and least effective, while using the explanations to add depth to the statistical conclusions.

The student interviews were conducted following each pair of aligned/non-aligned units. Students were selected based on their overall performance on the post-test and the laboratory report – one top performer, one middle performer and one low performer from each class. The questions were designed to follow-up on student responses to the questionnaire. The interviews were recorded and analyzed for information to help with the quantitative data analysis. The teacher reflection log was used in a similar fashion.

The research methodology for this project received an exemption by Montana State University’s Institutional Review Board – Appendix I – and compliance for working with human subjects was maintained.

DATA AND ANALYSIS

The primary research question states: What is the effect of aligning both curriculum and instruction to the NGSS three-dimensional model? The first sub-question states: What is the impact of NGSS-aligned revisions on student learning, as measured by acquisition and application of disciplinary core ideas? A pre- and post-instruction test was administered for each unit to measure learning. These tests also had separate components to analyze learning at the acquisition and application levels. The classroom research project was conducted over four instructional units – two treatment units and two non-treatment units. Normal gain scores were calculated for each student for each assessment. As a result, each student had two normal gain scores for treatment and two for non-treatment, with an overall $N= 104$. The treatment units covered the topics of photosynthesis/cell respiration and energy resources, while the non-treatment units were on agriculture/food resources and energy concepts. Figure 1 compares the pre- to post-instruction normalized gains for acquisition, application and overall performance in treatment and non-treatment units. The error bars represent 95% confidence intervals.

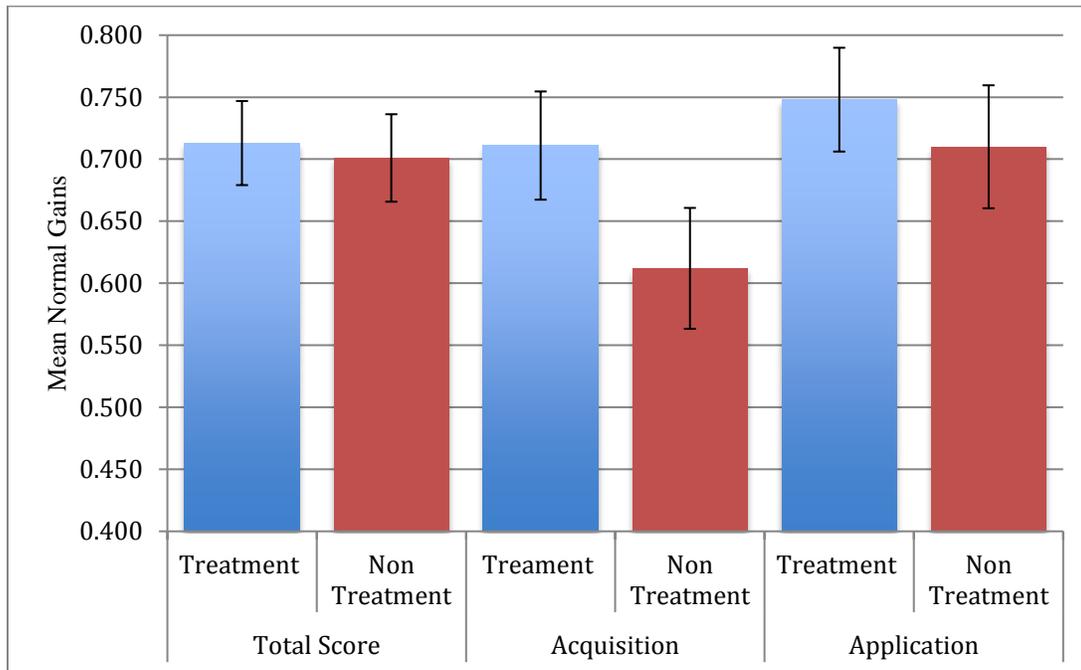


Figure 1. Normalized gains – treatment v. non-treatment pre- and post-test, ($N=104$).

The primary trend that is evident from this data is that learning occurred in both the treatment and non-treatment units, with slightly more learning demonstrated in the treatment units. This is evidenced by overall mean normal gains of 0.713 ± 0.034 CI for the treatment units and 0.701 ± 0.035 CI for the non-treatment units. When these gains are compared using a one-tail, equal variance t-Test the results are *not* statistically significant with a $p=0.3163$. A more fine-grained analysis, comparing the results of acquisition and application level learning, reveals a dichotomy between these learning levels. The treatment group saw mean normalized gains of 0.711 ± 0.044 CI compared with a non-treatment mean of 0.612 ± 0.049 CI at the acquisition level. This result *was* statistically significant by a one-tailed, equal variance t-Test, with $p=0.0014$. At the application level, the treatment group showed the highest normalized gain score in this project, 0.748 ± 0.042 CI, compared to the non-treatment group, 0.717 ± 0.050 CI. This result was *not* statistically significant based on a one-tail, equal variance t-Test, with

$p=0.1706$. It is possible that this three dimensional instruction model is most effective for basic acquisition. This is not the premise of NGSS, where learning that purposefully integrates the practices and concepts should also improve deeper/application-level learning. It may also be the case that the instructional practices for integrating the SEP are not yet well developed and thus have less impact on application level learning. It is also important to note that, while not statistically significant, the t-Test results suggest that more than 80% of the difference between the treatment and non-treatment groups is *not* random. The treatment had an effect, just not as big as the researcher would like.

The raw scores from the post-instruction assessments and the normal gains calculations were also examined for outliers, with some interesting results. In considering the raw scores from each unit quiz, the treatment units had 26 students fall greater than one standard deviation above the mean score and 7 students one standard deviation below the mean. The non-treatment units had 17 students one standard deviation above and 6 one standard deviation below the mean. This further supports the efficacy of the treatment unit instruction, showing that those units had more students towards the upper end of the distribution range. The low score on each of the unit assessments ranged from 13.5 to 15 (total possible of 27) and was a different student in each instance. Importantly, the students with the lowest scores typically showed mid-range normalized gains, showing they were learning more than some even if their overall results were low. This conclusion is also supported in the analysis, later in the paper, of students by academic performance levels. The high score on each quiz was a perfect score of 27, with 3 from the treatment units and 10 from the non-treatment units. The students reported, in the perception survey, that the non-treatment assessments were

easier than the treatment and this analysis of high scores seems to support that. This is more interesting given that student learning was greater in the treatment units.

The outliers in the normalized gains data were different. The treatment units had 14 students one standard deviation above the mean and 13 below, while the non-treatment units had 14 students one standard deviation above and 14 below. These values are based on an $N=104$ for both categories. These normalized gains were used for the bulk of this study's statistical analysis. This consistent distribution of high and low outliers across both treatment and non-treatment units indicates a normal distribution to that data and, along with the large sample size, supports the use of parametric tests for statistical significance.

Students were placed in three cohorts based on their overall course grade at the time of the post-instruction test. Cohort 1 represents the top third of students by performance ($N=36$), Cohort 2 the middle third ($N=36$) and Cohort 3 the bottom third ($N=32$). This permitted a comparison of normalized gains by cohort level, to see if the treatment effect was similar for different types of students. The results of this analysis are summarized in Figure 2.

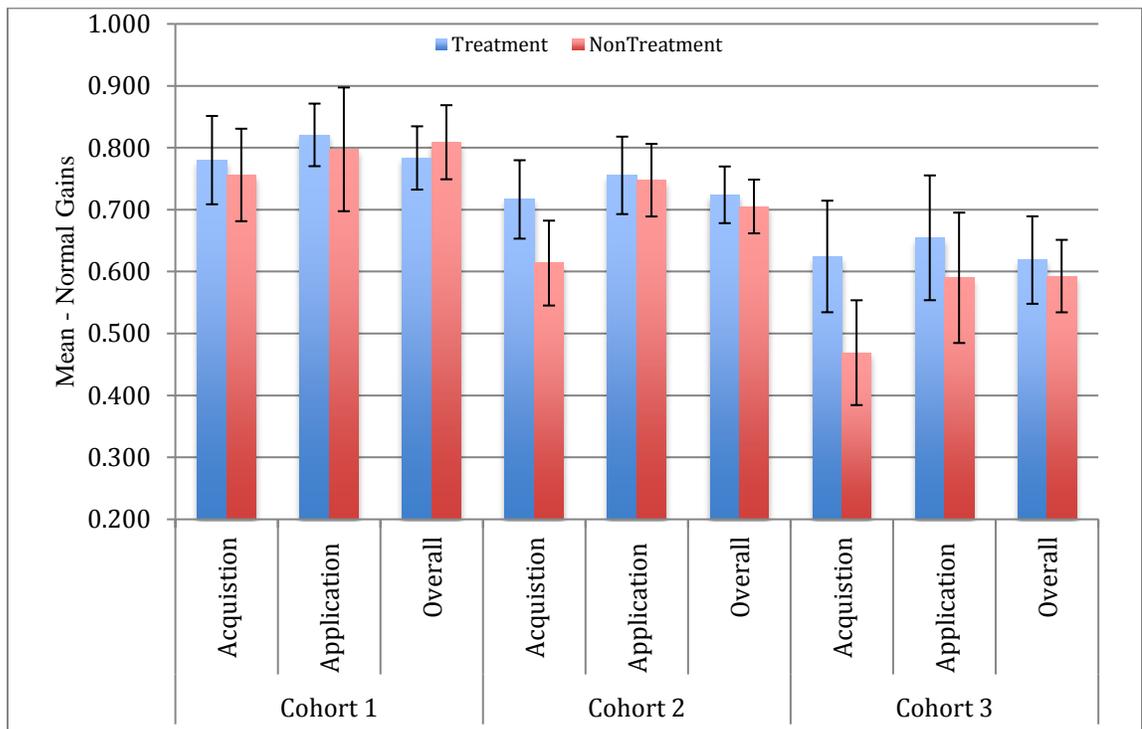


Figure 2. Normalized gain treatment v. non-treatment – by learning level and student performance level.

Note: Cohort 1 = top ($N=36$); Cohort 2 = middle ($N=36$); Cohort 3 = bottom ($N= 32$)

An interesting trend from this analysis is that there appeared to be different effects based on student performance level, specifically that learning gains were greatest in cohort 3. In cohort 1, the top performers, the difference between treatment and non-treatment units was slight in all categories. Normalized gains were higher for acquisition and application categories, but the non-treatment units showed slightly higher overall gain than the treatment unit. In cohort 2, the treatment units showed larger gains overall and at both acquisition and application levels. The difference at the acquisition level was noticeably large, with the treatment mean 0.717 ± 0.063 CI and the non-treatment mean 0.614 ± 0.069 CI. This was statistically significant with a one-tail, equal variance t-Test and a $p= 0.0142$. At the cohort 3 level, the bottom performers, the pattern was similar to cohort 2 but with bigger learning gains at all three measurements. The difference in the

normalized gains for acquisition was statistically significant with a $p= 0.0065$. The differences in normalized gains for application were not statistically significant, but the p -values declined from cohort 1 to cohort 3.

The instructional revisions to align with NGSS might help to explain these results. For example, concept mapping – as a form of modeling – was incorporated into the treatment units as an NGSS-aligned instructional strategy. In the student interviews and the questionnaires, this was mentioned consistently. A cohort 3 student noted in the interview, “Having to make concept maps made it possible to create a system that worked for me.” A second student interviewed from cohort 3, also commented on this strategy, “It wasn’t the most fun, but it helped me create a way for myself to think that wasn’t assigned as the book written for everyone and that can be confusing to some minds.” The data, in conjunction with the student comments, suggests some real advantage to three-dimensional instruction, particularly with students who are not top performers. This also matches with the conceptual predictions of the NGSS, namely that students who typically struggle are more likely to be impacted by the opportunity to engage in content through the practices. Top performers may learn effectively with one-dimensional instruction or, at the least, with less need to engage in the practices to support their learning. An interesting question for future consideration may be whether three-dimensional instruction hurts, helps or is neutral for all learners or with only certain learners.

The laboratory reports associated with each unit also provided an opportunity to analyze application level learning of the disciplinary content. An example of the laboratory report is found in Appendix G. All laboratory reports were structured with a

series of eight prompts – each scored using a four-point rubric – designed to evaluate specific DCI and specific SEP. Two of the prompts specific to the DCI were associated with each lab report. Each student wrote two treatment unit and two non-treatment unit lab reports for an $N=104$. An analysis combining the two specific DCI prompts provided a measure of application level learning for those content standards. With two prompts from each student laboratory report, these analyses had an $N=208$. The laboratory report results are summarized in Figure 3.

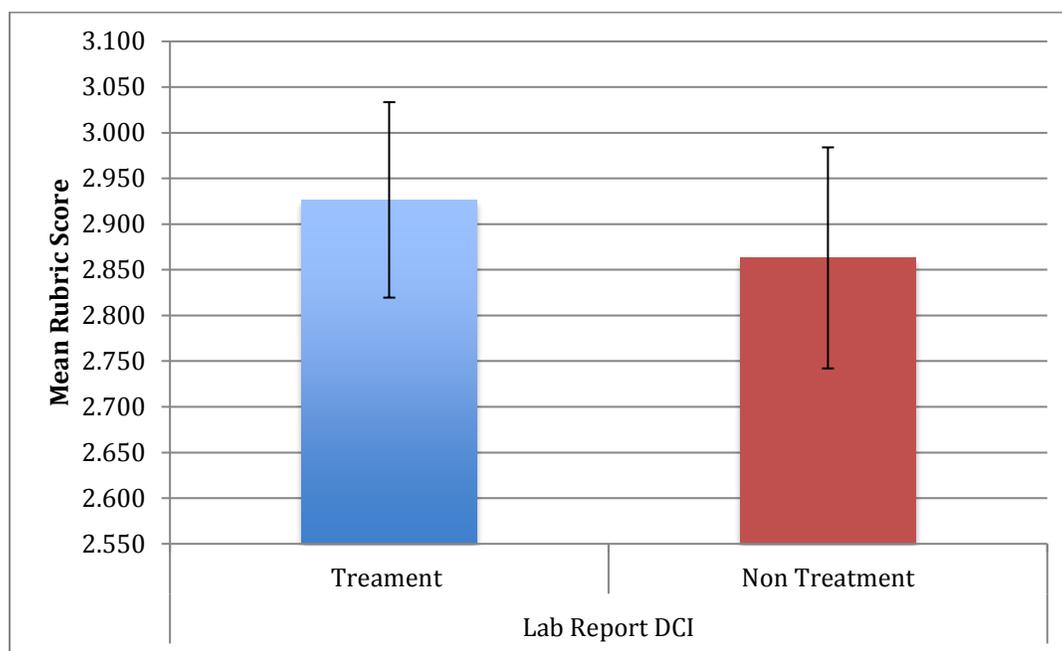


Figure 3. Lab report DCI prompts – treatment units v. non-treatment units, ($N=208$).

Student performance level cohorts were also used to analyze the DCI prompts from the laboratory reports. Cohorts 1 and 2 each contained 36 students contributing two laboratory reports with two DCI prompts in each report for a total $N=144$. Cohort 3 had only 32 students for an $N=128$. These results are presented in Figure 4.

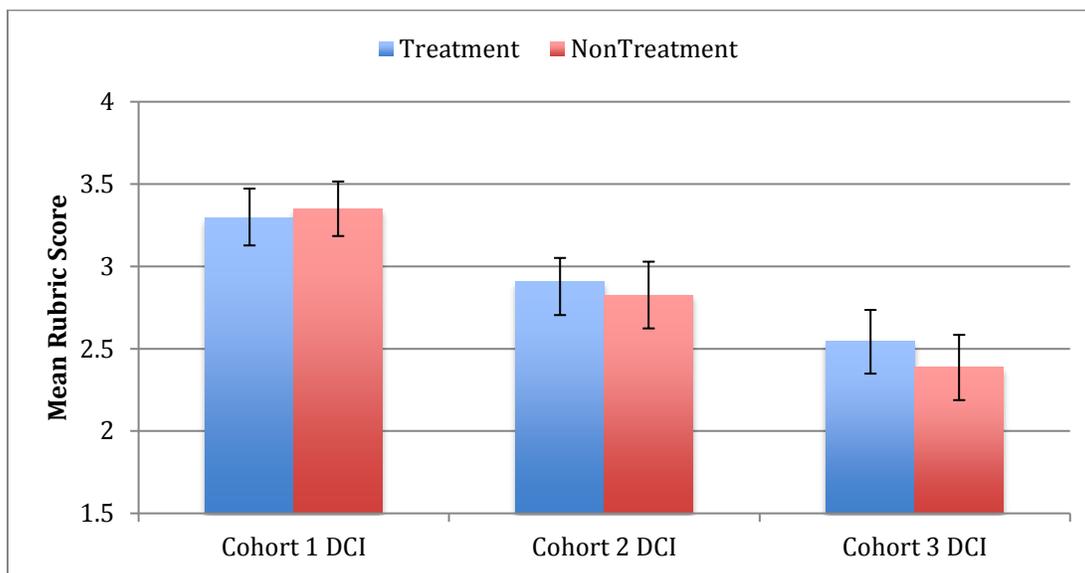


Figure 4. Lab report DCI prompts – by student performanceLevel.

Note: Cohort 1 Top Level ($N=144$); Cohort 2 ($N=144$); Cohort 3 ($N=128$)

These results are consistent with the DCI learning that was measured by normalized gains from the pre- and post-assessment tests. Overall, the treatment units show a slight benefit in application-level learning of the DCI when compared to the non-treatment units, with a mean rubric score of 2.926 ± 0.107 CI compared to a mean score of 2.863 ± 0.121 CI. This was not statistically significant, with a $p=0.2156$ from a one-tailed, equal variance t-Test.

In the cohort analysis, the results were also similar to the pattern of the normalized gains data. The treatment units had the greatest effect on cohort 3, the bottom performance level, and the least effect on cohort 1. In fact, for cohort 1 the results indicate a slightly better performance in the non-treatment units, 3.350 ± 0.166 , compared to the treatment units, 3.293 ± 0.179 . Cohort 2 showed a small difference in favor of the treatment units. The difference was, as with the test data, most pronounced in cohort 3, with a DCI treatment value of 2.547 ± 0.189 and a non-treatment value of 2.386 ± 0.198 .

None of the differences were significant, but the p-values were smallest in cohort 3. In discussing the laboratory on photosynthesis and cell respiration, from the first treatment unit, a cohort 2 student stated, “The data and the lab itself was more of a reinforcement on the concepts we were already learning without the lab.” These results seem to support the conclusions from the pre/post test, that the instructional revisions to align these topics with the three dimensions of NGSS had a greater effect on the learning of the lower performing students. This suggests the need for continued exploration on this idea, as well as whether the top learners are affected in a positive, negative or neutral manner.

The second research sub-question states: What is the impact of NGSS-aligned revisions on student learning, as measured by application of the science and engineering practices? Analysis of the student laboratory reports provides an opportunity to evaluate this sub-question. The laboratory report –Appendix G – is structured with a series of eight prompts designed to evaluate specific aspects of the SEP. Each prompt has a four-point scoring rubric associated with it. The average rubric score for the entire laboratory report provided a measure of the overall application level learning of the SEP. Each student wrote two treatment unit and two non-treatment unit lab reports for an $N=104$. It was also possible to combine the first four prompts as a measure of the following practices: plan and conduct experiments (PE); analyze and interpret data (AD); mathematical and computational thinking (MT). Similarly, the second four prompts serve as an indicator of the following practices: asking questions (AQ); constructing explanations (CE); argument from evidence (AE). Each of these groupings have an $N=416$. The laboratory report results are summarized in Figure 5.

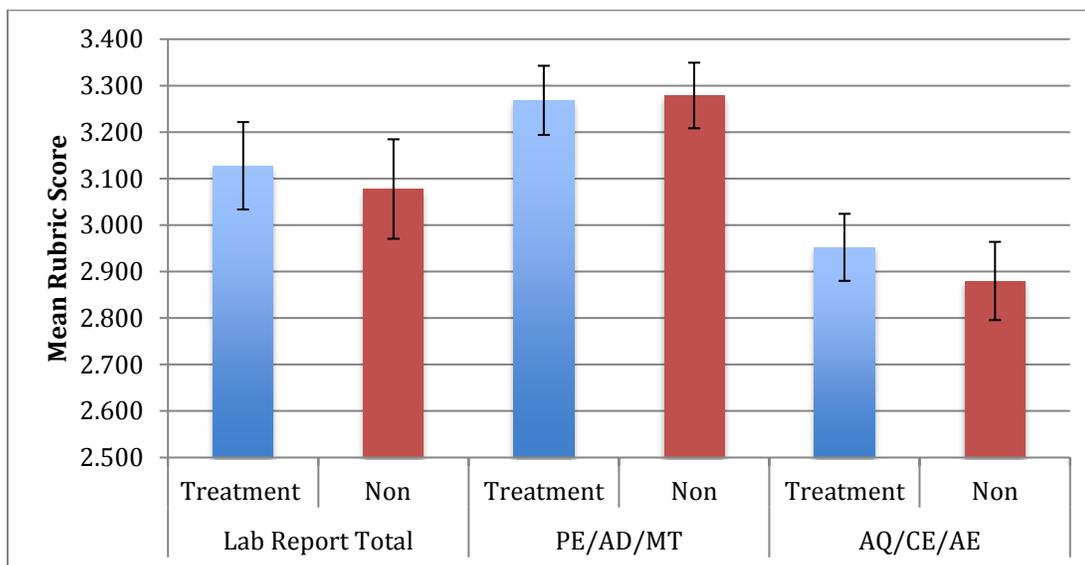


Figure 5. Laboratory report scores – total and SEP groupings.

Note: Lab Report Total N=104; PE/AD/MT N=416; AQ/CE/AE N=416

Analysis of the laboratory report scores suggest that the treatment units were, overall, more effective for student learning. This is true as an overall measure of SEP application, with the treatment group mean 3.128 ± 0.094 CI compared to the non-treatment mean of 3.078 ± 0.107 CI. The treatment had the greatest effect in prompts specifically linked to the constructing explanations and argument from evidence practices, where the treatment was 2.952 ± 0.072 CI compared to the non-treatment of 2.880 ± 0.084 CI. The planning experiments and analyzing data prompts showed a very slight difference in favor of the non-treatment group. None of these differences was judged statistically significant by a one-tail, equal variance t-Test. The p-values for each component were: $p = 0.2418$ for the overall lab report; $p = 0.4191$ for the PE/AD/MT practices and $p = 0.0962$ for the AQ/CE/AE practices. While the differences are not statistically significant, they do suggest that the treatment has a positive impact on learning the SEP. This seems a promising result, especially given that significant parts

of the treatment unit – including its newly aligned laboratory activities – were being taught for the first time. Experience suggests that the effectiveness of this instruction will improve with subsequent use.

The laboratory report results were also analyzed, as with the normalized gains, by student performance levels. This approach used the same three cohorts, based on student course grades at the due date of the laboratory report, in an attempt to determine if these instructional revisions for more effective for different groups of learners. These results are presented in Figure 6.

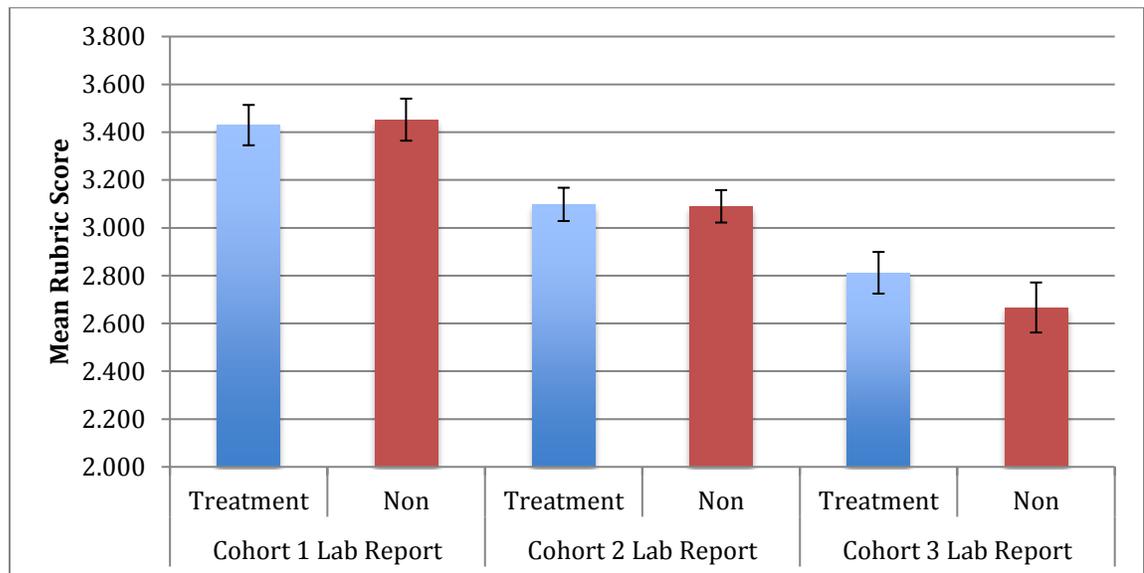


Figure 6. Laboratory report score – by cohort.

Note: Cohort 1 = top ($N=36$); Cohort 2 = middle ($N=36$); Cohort 3 = bottom ($N= 32$)

The pattern with these results is similar to the pattern noted with the normalized gains data. The treatment units had the greatest effect on cohort 3, the bottom performance level, and the least effect on cohort 1. In fact, for cohort 1 the results indicate a slightly better performance in the non-treatment units. In overall average rubric score, the cohort 3 treatment group had a mean value of 2.812 ± 0.157 CI

compared to a non-treatment value of 2.667 ± 0.188 CI. None of the cohort differences was statistically significant by a one-tail, equal variance t-Test. However, the p-values for cohort 3, $p=0.1237$, was much smaller than the tests conducted on cohort 1 and cohort 2. These results support the conclusions from the pre/post test, that instructional revisions to align the topic's instruction with NGSS had a greater effect on the learning of low performing students. This suggests a need for continued exploration on this idea. How are learners of various levels and various skill sets impacted by this instruction? Are these various groups of student affected in a positive, negative or neutral manner?

Students were given a questionnaire after each of the post-instruction tests. This was designed to probe student perception on alignment between specific aspects of instruction – readings/homework, lectures/direct instruction, activities/labs – and their learning. Each student completed two treatment and two non-treatment questionnaires resulting in an $N=104$ for this comparison.

Figure 7 represents the mean values for student perception of instructional alignment in the treatment versus non-treatment units. Figure 8 shows student perception of overall alignment separated out by the individual treatment and non-treatment units.

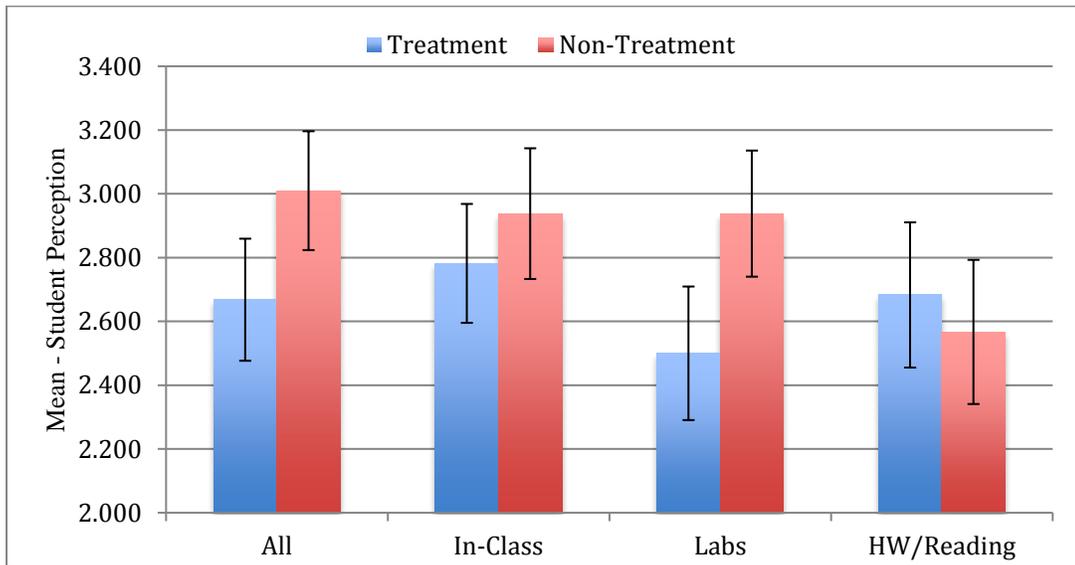


Figure 7. Student perception of instructional alignment – treatment units v. non-treatment units, ($N=104$).

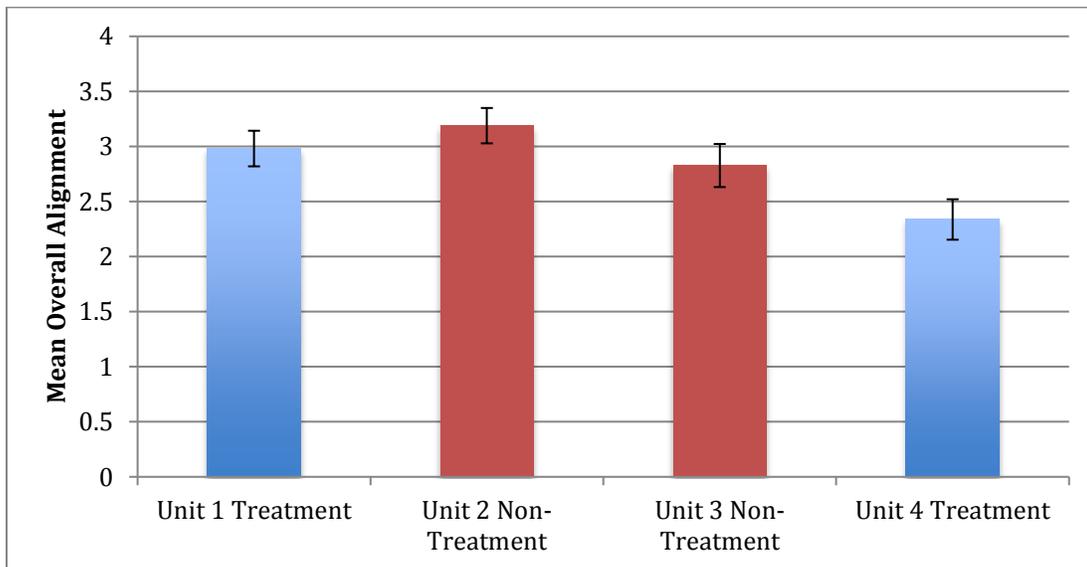


Figure 8. Student perception of overall alignment – by unit, ($N=104$).

The questionnaire data reveals some interesting trends that are contradictory to the assessment results. First, students generally felt the instruction from the non-treatment units was better aligned with the assessments than were the treatment units. This was true as an overview of all instruction forms, with the mean score for the treatment at 2.668 ± 0.191 CI compared with 3.010 ± 0.186 CI for the non-treatment. The

only sub-category where students felt the treatment unit was better aligned with assessment was for homework and reading assignments – treatment = 2.683 ± 0.228 CI and non-treatment = 2.567 ± 0.226 CI. In considering the distribution of responses, by category, the researcher noted that students felt the laboratory activities from the treatment units were the least effective instruction for their learning. This was the opposite of the starting assumption that laboratory activities were the most effective instructional approach. Based on the questionnaire, in three of the four instructional units the laboratory activity was most frequently identified as the *least* effective instructional event for student learning. These results are summarized in Table 3.

Table 3.

Student Questionnaire – Most/Least Effective Instructional Activities (N= 104)

	Most Effective			Least Effective		
	Lab	Lecture	Other	Lab	Lecture	Other
Treatment	13	41	24	43	4	11
Non-Treatment	28	43	26	20	14	10

Note: These are open-ended response items, so the numbers do not match $N=104$.

This data suggests several interesting points. First, the overall identification of what is most and least effective seems to be based less on the treatment/non-treatment and more on individual student preference. For each of the instructional units, there were students who identified the laboratory work as most helpful and students who identified the laboratory work as least helpful. While the overall numbers suggest the lab activities to be less helpful, analysis of individual responses leads to the idea that this may be more about personal learning preference and the student's ability to tolerate the ambiguity and challenge inherent in an open-ended laboratory activity. From student

interviews, some quotes support the idea of laboratory instruction as useful. From a cohort 1 student, “The labs force me to do all the calculations and give a thorough write up of my understanding so it all synthesizes. Also, doing something in real life helps.” A cohort 3 student noted, “Labs really showed me how photosynthesis and cellular respiration work. I was able to see the outcomes with my own eyes.” Other student quotes suggest a different picture. A top cohort 1 student said, “I found it difficult to connect the main points of the solar cell lab to what we were studying in the unit.” From the questionnaire a cohort 2 respondent added, “We didn't even really know what the lab was about, we just followed the instructions and everyone was confused.” And finally, from a cohort 3 questionnaire response, “ I hate labs. They are confusing and I have trouble with a lack of instruction. We should use class time to complete these as we are walked through them for total comprehension.” The design of the questionnaire may also play a role in this response pattern, where the students were asked what instruction was ‘least effective’ for their learning. Some students responded that the laboratories were helpful, but still less effective than other activities. This pattern of response was consistent in the interviews, as well. Also, in the non-treatment unit on energy concepts students identified the laboratory activity as the *most* useful for their learning. This laboratory has been taught for a number of years and is well developed in that regard. The two treatment units involved labs that were newer to the curriculum and, in addition, had been significantly re-designed as part of the NGSS alignment process. Consequently, these lab activities varied in how well they connected to the curriculum, how challenging they were for the students and in their overall level of engagement. It is possible that this interpretation of how useful laboratory based instruction is to learning

is mostly linked to the individual lab/experiment that the students are conducting. It may also be that these students, not having a long experience with NGSS and inquiry based instruction, are simply reacting to their – and the teacher’s – unfamiliarity with this type of work. These issues may be rectified with targeted revisions to the activities and increasing sophistication, from both teacher and student, with this type of work.

A second observation is that many students, contrary to most prevailing theory, feel the lectures are effective for their overall learning. This was the case in three of the four instructional units, with the exception being the second non-treatment unit where the lab was identified as most effective for learning. Regarding the treatment unit on photosynthesis a cohort 2 student noted, “In the lectures, what you wanted us to know became clear and I didn’t spend time studying things that weren’t important.” Similarly a cohort 1 student, following the non-treatment unit on agriculture, stated: “It’s easier to understand the material if Mr. Samet explains the topic with a PowerPoint before we do the unit or lab because my brain can connect the bigger picture of the unit.” Finally, and perhaps most interesting, from a struggling cohort 3 student: “The lectures helped me most because I learn better with visual representations while someone talks about them.” Among the instructional alignments made for the treatment units was an effort to include the SEP into direct instruction. This included using modeling in lectures on photosynthesis and solar energy, as well as the use of argumentation to defend model revisions. Did these instructional changes influence the perception of student learning from lectures or is lecture/direct instruction an effective tool in the appropriate situation? Alternatively, is this favorable perception of direct instruction an artifact of generally motivated honors students who simply want the “right” answer?

Considering student perception of alignment for the individual units, as shown in Figure 7, is also interesting. It is important to remember that the student perception is based on alignment between instruction and learning. Student perception of learning is largely associated with their assessment performance. This, in turn, largely focused their attention on DCI learning. They were not assessing alignment with the NGSS nor were they specifically addressing learning of the SEP. In that light, their perception of better alignment in the non-treatment units may be a result of those units having been taught, in their current form, for several years. The treatment units both involved instructional revisions being implemented for the first time and, as such, may have been exhibited aspects that would be revised – with improved alignment – in subsequent years.

There are also some contradictory results in student perception of assessment difficulty. These results are displayed in Figure 9 and Figure 10.

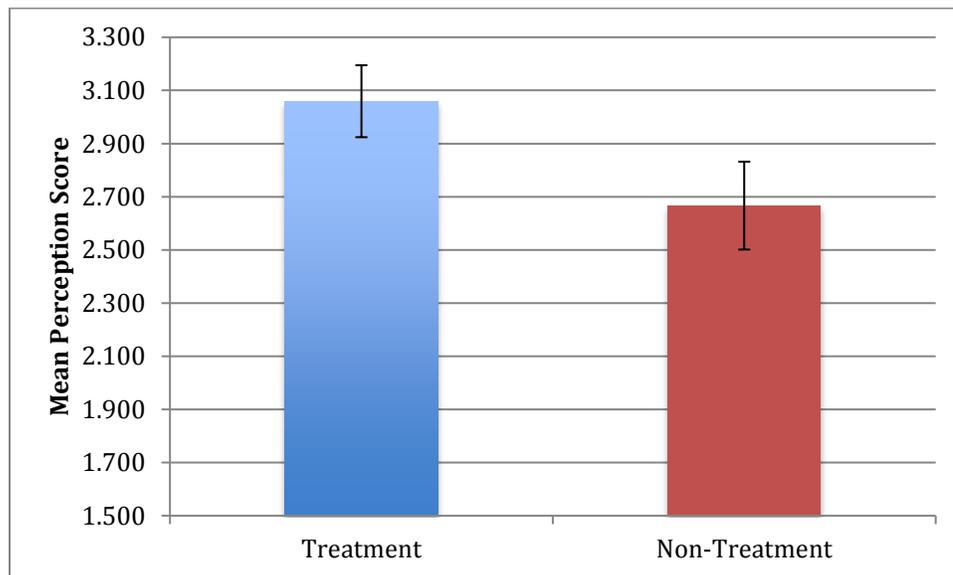


Figure 9. Student perception of assessment difficulty, ($N=104$).
Note: Higher value indicates *greater* perceived difficulty.

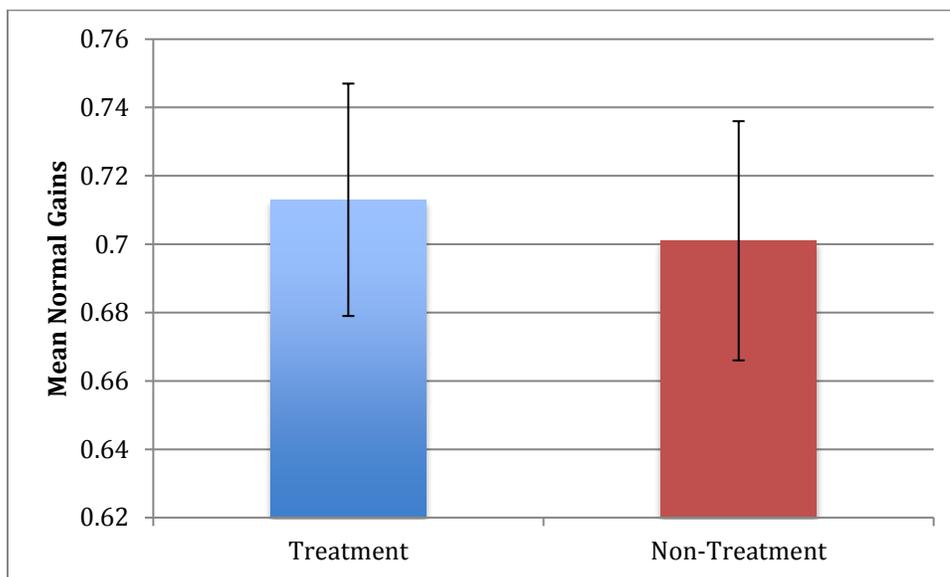


Figure 10. Normalized gains – pre/post assessments, ($N=104$).
 Note: Higher value indicates *greater* learning.

Collectively, the treatment unit assessments were perceived as more difficult than the non-treatment unit assessments even though the treatment units showed slightly better normalized gains. The mean perception of assessment difficulty for the treatment units was 3.056 ± 0.135 CI compared to 2.667 ± 0.165 CI for the non-treatment units, while the mean normalized gain for the treatment units was 0.713 ± 0.034 CI for the treatment units and 0.701 ± 0.035 CI for the non-treatment units. The data suggests that students showed larger learning gains even when measured by more difficult assessment instruments. This seems to be an argument supporting the efficacy of the NGSS aligned instruction.

INTERPRETATION AND CONCLUSION

The primary research question states: What is the impact on student learning of revising curriculum and instructional activities to align with the NGSS model of three-dimensional learning? There are two secondary questions. The first states: What is the

impact of NGSS-aligned revisions on student learning, as measured by acquisition and application of the disciplinary core ideas? The second states: What is the impact of NGSS-aligned revisions on student learning, as measured by application of the science and engineering practices.

Acquisition of Disciplinary Core Ideas

Student learning of disciplinary content at the acquisition level, as measured by normalized gains between pre- and post-assessments, was consistently greater in the treatment units. Additionally, learning gains in this area were greater in a cohort of low performing students than in two comparison cohorts of higher performing students. To further support these conclusions, t-Tests comparing the differences in overall normalized gains for DCI acquisition were determined to be statistically significant ($p=0.0014$), as were the normal gains for the middle ($p=0.0142$) and lower performing ($p=0.0065$) student cohorts.

These learning gains make sense in light of instructional revisions to the treatment units. First, there were changes to both lectures and in-class activities based on the assessment boundary and clarification statements from the NGSS performance expectations. These were curricular changes that provided a narrower and more focused approach to the disciplinary content of photosynthesis/cell respiration and energy resources. Instructional strategies were also modified to align with the NGSS. Direct instruction activities were shifted from a standard lecture format to more interactive strategies based on the SEP. In the photosynthesis and cell respiration unit, some lecture content was eliminated based on the DCI assessment boundary. The lectures were also organized into smaller chunks that were interspersed with concept mapping, a form of

modeling. The concept maps were subsequently shared using a discourse protocol that applied the SEP of constructing explanations and argument from evidence. A similar process occurred in the treatment unit on energy resources. Based on the clarification statement, the students studied fewer types of energy resources – with a narrowed focus comparing electrical generation from solar photovoltaic and fossil fuels. The classroom activities again had a focus on the practices of modeling and argument from evidence. When studying fossil fuels, students worked to research and then build representational models of a resource use cycle for different fossil fuel types. They followed this with an evidence based argument – presented both orally and in writing – about the costs and benefits of fossil fuel use based on their resource use model. In considering solar energy, their model building centered on the energy transformations that occur in a solar photovoltaic panel. Student feedback, from both surveys and interviews, was very positive about the modeling work. Modeling activities were also identified as positive and productive in the researcher’s teaching journal entries. Additionally, to the researcher’s surprise, students reported that the lectures in both of these units were most effective to their learning. Through the interview process, it became clear that they viewed the modeling work as part of the lectures, which further contributed to the strong perception of value.

The laboratory activities in each treatment unit were also revised to better align with both the disciplinary content statements and the science and engineering practices. The photosynthesis and cell respiration lab was revised around the practice of mathematical and computational thinking. The students did not design this experiment, but were required to plan the analysis of raw data to effectively measure rates of

photosynthesis and cell respiration. The lab on solar energy was an open-inquiry lab, where the students were responsible for the design of both experiment and data analysis. They also applied their conclusions from this lab to further inform their cost–benefit analysis of solar energy.

In the treatment units, both classroom activities and laboratory activities were modified with a purposeful integration of DCI and SEP. This three-dimensional design may have been the catalyst behind the larger, and statistically significant, gains in student learning of disciplinary content at the acquisition level. Additional research is certainly suggested on the role of these revisions relative to the larger learning gains of the lower performing student cohort. What aspects of the SEP are most helpful to learning overall and in this low performing group specifically. Conversely, it is also important to consider whether these revisions have a positive, negative or neutral effect on the high performing group.

Application of Disciplinary Core Ideas

In considering the effect of NGSS-aligned instruction on student learning of disciplinary content at the application level, the results show similarity to the acquisition level results. This lends further support to the idea that these revisions based on NGSS alignment have validity. The normal gain results from the application questions on the pre- and post-tests show slightly more learning – albeit not statistically significant – in the treatment units. Interestingly, the normal gains score for application level learning was the highest of the three categories analyzed, indicating the largest learning gains in this area. The cohort analysis from the normalized gains repeats the trends seen in the acquisition level analysis, where cohort 3 showed the greatest learning improvement in

the treatment units. Cohort 2 showed a very slight increase in the treatment units, while cohort 1 showed a very small decrease in the same units. The results from the DCI application prompts in the lab reports mirror the pre- and post-test results. Cohort 3 again showed the greatest difference in average rubric scores between the treatment and non-treatment units. Cohort 2 had a smaller difference, while cohort 1 actually showed a very slight difference in favor of the non-treatment unit. The pattern of results in these two measures of application level learning lends additional support regarding the overall efficacy of this instructional alignment approach.

The improved alignment of all instructional activities to the SEP should, according to the concept of three-dimensional learning, help the students learn and apply the disciplinary content at deeper levels. Overall, the data seems to reflect this. This is perhaps even more striking when considering the student questionnaire data that showed clearly the laboratory activities were perceived to be least effective in supporting their learning. On a finer level of analysis, three of the four labs were perceived least effective for learning while the lab from the non-treatment unit on energy concepts was most effective. Two of the ineffective labs were newly revised/aligned to NGSS and being taught in that format for the first time. It is likely that the normal process of reflection and revision will further improve the effectiveness of those activities, resulting in even greater access to three-dimensional learning. It is also possible that the lab activities really are less effective for overall student learning. An interesting question for further consideration would take a deeper look into the student perception of labs as ineffective instruction. Is the problem with these specific labs or with laboratory activities in general? Is the problem a lack of connection between the disciplinary content and

laboratory activity? Is the problem with the overall level of complexity in an inquiry-based lab? Accurate answers to these questions should help make more effective revisions to the laboratory activities that are so important to the NGSS model and, the researcher would argue, to effective science instruction.

Application of Science and Engineering Practices

The second research sub-question states: What are the impacts of NGSS-aligned revisions on student learning, as measured by application of the science and engineering practices? This question was measured using the laboratory reports associated with each unit. Student results for the overall rubric score and for the combined scores from related prompts were compared between treatment and non-treatment units. It is worth noting that the researcher originally hoped to measure both acquisition and application of the SEP in this classroom research study. However, the author found it very difficult to construct an instrument that adequately measured acquisition-level learning of these practices and skills. This is, perhaps, because their utility is almost exclusively tied to their application and, as a result, evaluation of learning is best measured in that context.

The pattern of student learning results on SEP application was very similar to the learning patterns reported for the application and acquisition of the DCI. In comparing overall results of the laboratory reports – averaged rubric scores on eight individual prompts – the treatment units had a higher average score than the non-treatment, but the difference was not statistically significant with a $p=0.2418$. Additionally, analysis of the overall lab report score by learning cohort reveals the same patterns from the DCI analysis, that the cohort 3 students showed the largest difference between treatment and non-treatment units. This continues to support the idea, consistent with the DCI

interpretations, that this version of aligned instruction is helpful to the learning of lower performing students. What aspects of the three-dimensional model are most effective in producing these results? Is the three-dimensional integration most important or do the individual components bring differential value of their own? Given that the cohort 3 learners may be slower learners, what is the cumulative effect of the SEP over time?

The individual lab report prompts were grouped into two categories and the scores were averaged for additional analysis. The first cluster of prompts focused on student application of the following practices: plan and carry out investigations; analyze and interpret data; mathematical and computational thinking. This cluster had the highest overall score and the non-treatment units outperformed the treatment units by a very slight, statistically insignificant, margin. These aspects of the practices have historically been a focal point of the entire science curriculum in this school district, which may explain both the high level of these scores and the lack of discernible difference between treatment and non-treatment units. The second cluster of prompts focused on these practices: ask questions; construct explanations; argue from evidence. In this data set, the treatment units showed higher scores than the non-treatment. This category showed the largest difference between treatment and non-treatment of the three categories, approaching statistical significance with a $p=0.0962$. The researcher suggests that the instructional revisions – to both direct instruction and lab activities – played a role in this difference. For example, the students utilized the argument practice recursively in the final treatment unit on energy resources. Following each class activity and the lab, they were asked to develop arguments both for and against the use of each energy source that was studied. In the class activities, these arguments were created collaboratively and

defended orally, with subsequent peer feedback. At the end of the unit, these revised, re-worked arguments were presented in writing as a summative assessment. This recursive practice, so heavily focused on the final treatment unit, likely played a role in the stronger learning results exhibited by the combined treatment units.

There is an additional aspect that may also impact the interpretation of the SEP application results. Unlike the DCI, which is mostly specific to a given instructional unit, the SEP are recursive across units (and, as the NGSS is fully implemented, through a student's entire academic career). While revisions to explicitly include and align the SEP occurred in the treatment units, students received some instruction and practice around the SEP in *all* of the units. The Tamalpais district's science curriculum has always valued and included laboratory work as an instructional cornerstone. The students have used aspects of the practices, although not all as described in the NGSS, in many parts of the district curriculum. As a result it was difficult to evaluate the effects of learning that were related specifically to the SEP revisions implemented in the treatment units. It is possible that student learning related to the SEP continuously improved with the ongoing opportunity to use those practices in each of the units, explaining why the treatment units showed greater overall gains in application when compared with the non-treatment units. The difference in the explanation/argument prompts may be explained by the new and explicit focus on these skills, while the lack of difference in the planning/data analysis prompts is a result of the ongoing focus on those skills in the existing curriculum.

Instructional Strategies

Laboratory-Based Instructional Activities

The development of effective NGSS-aligned instruction requires the careful integration of content (DCI), practices (SEP) and concepts (CCC). This integration happens through planned instructional activities and they, as such, are crucial to the model's efficacy. In the data analysis process, the researcher was struck by a mismatch between his perception and student perception of the different instructional strategies and their value to learning. From both questionnaire and interview responses, students consistently rated laboratory work as less effective and lectures/direct instruction as more valuable than the instructor had assumed.

A couple of interpretations are possible from this data. First, the value of laboratory work cannot be assumed, as the researcher has done. The primary assumption was that engaging in the practices – whether the SEP or older Nature of Science concepts – would improve learning in any context. From the student responses, it seems that this connection between engaging in practices and learning disciplinary content is not automatic and did not, necessarily, exist in the NGSS-aligned laboratory activities. In that revision process, the researcher took pains to align the laboratory activity to specific aspects of the DCI and SEP. However, the researcher did not purposefully align *between* the two areas and this lack of alignment was a major theme in student critiques of instructional value. This connection must be made explicitly and a mechanism for this may lie in the cross cutting concepts (CCC). These important connecting ideas were identified in the planning and alignment process, but were not used to their full advantage, nor were they specifically addressed in this research project. For example, the CCC of energy and matter would have been a useful tool in connecting both of the

treatment unit lab activity back to the content – photosynthesis/cell respiration in the first unit and energy resources in the last unit.

A second interpretation of this data reminds the instructor to not mistake engagement for learning. In the interviews, students generally reported liking the laboratory work. Additionally, the researcher's teaching journal specifically and repeatedly notes the level of student engagement during these activities. The open-ended laboratory on solar energy, in its newly revised format, was especially notable for its engagement/interest level. However, this was the lab that the student responses especially noted was difficult to connect back to the primary content of the unit. Based on both research and personal experience, engagement is certainly useful for learning but is not necessarily a good indicator of it. In the design of three-dimensional instruction, we should strive for student engagement but focus on and measure student learning. This would be an interesting area of future research.

A final conclusion in this area is how the level of difficulty and expectations impacts student learning and student perception of learning. The move to inquiry-based lab work – less directive, more open-ended – means students must think creatively and be open to ambiguous results. These are not skills that come easily to tenth grade honors students. Much of their experience with science instruction, and school in general, has been very different. As a result, they are not practiced or experienced in these skills. It is worth considering that their low perception of the laboratory work may be related to this disconnect between past and current practice. Since the overall scope of the data does suggest slightly greater learning – even from newly implemented labs with problematic

alignment – it may be that the low student perception of instructional value is the result of more challenging work.

Lectures/Direct Instructional Activities

Student responses indicated a higher value on this lecture-based instruction than the researcher expected. In the treatment units, these lectures were revised to incorporate aspects of the SEP, making the activity as a whole more interactive. Students indicated, in the interviews, that they considered these activities as part of the lectures and this likely added value to their perception. The instructor, being somewhat “old-school,” has long felt that lectures are a necessary instructional strategy and one that can be effective when employed thoughtfully and strategically. It is, of course, only one tool in the instructional toolbox and can be very ineffective when deployed indiscriminately. The student learning results, coupled with their overall perception, suggests the researcher’s preliminary assumptions have value. An alternative explanation could be that these study participants, generally motivated honors students, are most interested in getting the “right” answer and they find the lectures are most effective to this goal. Additional research into the qualities and circumstances of effective lectures would be useful and interesting.

In-Class Instructional Activities

The role of revised in-class activities in learning was an area of agreement between instructor and student perception of value. These were instructional activities that were not lectures or labs and were designed to provide students with opportunities for meaning making. In both treatment units, these involved the use of modeling, constructing explanations and arguing from evidence to process disciplinary content. In

some instances, these strategies were woven into lecture-based instruction. The instructor consistently noted high value for these specific activities and the student perceptions – from questionnaire and interviews – was consistent with those observations. This data, in conjunction with the improved learning in those treatment units, suggests the value of this model of three-dimensional planning and instruction.

Assessment Quality and Difficulty Level

Another aspect to consider in these comparisons between treatment and non-treatment units is the validity and consistency of the assessment instruments. While the researcher devoted significant attention to their development, this is always a potential weakness in such comparative studies. The pre- and post-tests were built primarily with assessment items that had been used in the past. In this context, those items had been subjected to multiple evaluations for their overall efficacy. Over time they have demonstrated both validity and reliability. The researcher and his colleague individually evaluated each item for alignment and difficulty level. Additionally, the students provided an overall evaluation of the test difficulty in the post-assessment questionnaire.

In the teacher evaluation process, the results were generally consistent. The goal was an overall average close to 3.0 on a four-point rubric for difficulty. These results were within 0.25 for each of the assessments. In the second, non-treatment, unit the first evaluation of the test resulted in averages of 2.62 and 2.35, a difficulty level lower than preferred. This resulted in a re-evaluation and revision of that test, substituting three higher-level questions for more basic acquisition questions. The second evaluation resulted in average scores of 3.05 and 2.88, right in the preferred overall range. In spite of the general agreement between the teachers, the student assessment of difficulty level

was less consistent than was hoped. In particular, based on average questionnaire scores, the students felt the second non-treatment unit test was easier than the instructors' evaluation and the last treatment unit test was more difficult compared to the instructors' interpretation. The first and third unit assessments were much closer in agreement between teacher and student perception of difficulty. Additionally, the magnitude of the student's perceived difficulty level was big enough, ranging from 2.34 for the first non-treatment unit to 3.37 for the second treatment unit, to call into question the overall consistency of these instruments.

It is also interesting, as noted previously, that the students felt the treatment unit tests were more difficult even though their normalized gains from those tests showed more learning. This suggests a couple of possible interpretations. If the student perception is accurate and the tests are reliable and valid, then it is possible that the magnitude of the normalized gain difference is actually underestimated. If the tests are lacking in overall consistency, as suggested by the range of student perception scores, than any solid conclusions become problematic. An improved, more robust test development protocol would be helpful in this regard. Additionally, a larger, more comprehensive test of the content knowledge would help to improve the overall validity of the pre- and post-assessments.

External Factors

The interpretation of results from this classroom research project was influenced by some important factors outside of the research design. With only two treatment and two non-treatment units, small, unavoidable variations in unit implementation may have a significant effect on the overall results. In this study, the final instructional unit may

have been affected in this manner. From the teacher's journal, there were no fewer than five important interruptions to the instructional sequence. These included: a school holiday; the instructor missing school for a workshop; a minimum day for teacher meetings; two days devoted to the completion of student's independent research project. The result of this was to stretch out the overall instructional sequence while students, at the same time, had a divided focus while completing an independent research project. This pushed the end of instruction to a day that is normally the science department testing day, which in turn pushed the post-test back by almost a week. The students wound up testing on instruction that was further removed than in any of the other three units. The teacher's observations on this were backed up by student feedback from the questionnaire and from interviews. It is possible that a more compact instructional sequence, comparable with the other units, would have improved the post-test results. This, in turn, would have also increased the overall differences between treatment and non-treatment units. This is not offered as an excuse but, rather, as an observation on the difficulty of controlling all of the variables in a real classroom contained in a real educational system. There were, undoubtedly, many other small variations in the instructional delivery that the instructor did not note but that may have been otherwise important to the overall learning results. Maintaining consistency in all aspects of the study – especially in the relative difficulty level of content, instruction and assessments – is a very difficult challenge. It is, of course, one that educators deal with on a continuous basis and one that the researcher has worked on over an entire career.

There is a final problem to consider with this research model. Much of the existing curriculum – non-treatment units, as taught in prior years – was already aligned, albeit

not explicitly and to varying degrees, with the NGSS. Existing gaps between prior instruction and the NGSS were carefully identified, but calling the units ‘aligned’ and ‘nonaligned’ is not entirely accurate. In evaluating learning based on those units, it remains difficult to parse cause and effect based solely on alignment to the NGSS.

Overall Student Learning

The primary research questions states: What is the impact on student learning of revising curriculum and instructional activities to align with the NGSS model of three-dimensional learning? There is a consistent pattern to the learning results that suggests a real relationship between this type of NGSS-aligned instruction and student learning. Although only some of the results demonstrate statistical significance, there is a pattern of improved performance in the treatment groups receiving aligned instruction when compared to groups receiving non-aligned instruction. This pattern is repeated in the different learning categories – DCI acquisition, DCI application and SEP application. There is another potentially important pattern showing the biggest learning gains in treatment units occurring in the lowest performing group of students. This result is also consistent across all the learning categories. When considering the possible flaws in the study, for example inconsistency in instructional implementation and in assessment difficulty, it is possible that these results have under-estimated the treatment effects. The results are also supported by student feedback on the positive aspects of some revised instruction. The data suggests, to the researcher, that this type of three-dimensional alignment deserves wider implementation and further study.

VALUE

The process and results of this classroom research project have produced some useful lessons for the researcher. The big takeaway, based on the consistent pattern in learning results, is that this type of NGSS-aligned, three-dimensional instruction has potential value. This suggests continued use of the model, while more detailed analyses of this research also points to areas for improving the model.

It is the researcher's intention to continue three-dimensional instruction in his classroom. While it is possible to measure the learning in each of the dimensional categories – DCI, SEP and CCC, it is difficult to parse the individual effects of the dimensions. Nonetheless, there seems to be a real synergy amongst the dimensions that supports the further use of the model. While this project did not deal explicitly with the cross cutting concepts, the researcher recognizes this dimension as critical to linking practices with content, as well as one content topic to the next. In future work, the CCC will receive significantly more focus. The researcher will also be responsible for organizing and directing the NGSS implementation in his school and district. These lessons from the research project will be featured prominently in that work.

One of the weak links indicated by the data involves the effectiveness of laboratory-based instruction, explicitly in how the students make connections between labs and the disciplinary content. Going forward, this will be a focus of the researcher's design work. There are two specific aspects that merit greater consideration. The first, as noted above, is a more explicit use of the cross cutting concepts to connect lab practices with content. The second – and this relates to all instruction, not only lab-based – is an improved and more consistent use of 'explaining phenomenon' as an organizer for

instructional strategies. The idea of using three-dimensional instruction to explain phenomena is a central tenet of the NGSS, but one the researcher was late in recognizing. In the process of this research, it became clear that interesting and/or important phenomena should serve as specific examples for the performance expectations. In a similar fashion, using the idea of explaining phenomenon can be an important connection between laboratory activities and disciplinary content. The author will use cross cutting concepts and explaining phenomena in future laboratory-based instruction.

Another area of focus in designing laboratory activities, particularly across the whole curriculum, is the cumulative effect of students using the science and engineering practices. Unlike the DCI, which are typically specific to a given instructional unit, the SEP are recursive through a single course, a sequence of courses or across the student's K-12 experience. The NGSS specifies a broad learning progression for each of the practices (and each of the cross cutting concepts, as well). Consequently, the articulated use of the SEP over time is an issue to consider in planning three-dimensional instruction. The knowledge of where students are in the SEP continuum is crucial in designing laboratory activities that both take advantage of their prior knowledge/skills and allow them to appropriately continue that skill development. Understanding this requires larger conversations to assess and ensure both horizontal and vertical articulation about development of the SEP. If the three-dimensional model is accurate, students with higher-level SEP skills should be at an advantage in learning the disciplinary content at higher levels. This articulation will be incorporated into the researcher's design strategy.

The research project has also shed light on the importance of consistency in assessments. Implementing the NGSS across a school or district requires a conceptual change on the part of staff. In this teacher's experience, considering student work and student learning data can facilitate these sometime difficult discussions. Accurate and useful comparisons to drive these conversations require data from good instruments. To this end, a more careful and more purposeful alignment between learning goals and assessment tools must be a focus going forward. This is always recognized as a necessity, but this project served to reinforce the concept and remind of its importance.

The final takeaway from the research was more of a reminder but a very important one, particularly in the design of three-dimensional units. From the student questionnaire and student interviews, the instructor was reminded of the differences between individual students. This came to light especially when students reported on specific activities that were most effective and least effective to their learning. The same laboratory, or the same lecture, was interpreted as most helpful by some students and least helpful by others, reinforcing the challenge teachers face each day in each class. When relying heavily – or exclusively – on fewer strategies, teachers run the risk of consistently not reaching some students. Since reaching *all* students is the primary goal, instructional design must account for these differences. It seems that designing three-dimensional instruction is difficult *without* purposefully including a variety of instructional strategies. In this manner, the NGSS-aligned three-dimensional model supports what we know about designing effective instruction for all students and, as a result, deserves further exploration.

REFERENCES CITED

- Achieve. (2012). *Next generation science standards*.
<http://www.nextgenerationscience.org>
- California Department of Education (13 March, 2015). *Enrollment by Ethnicity 2014-15*. Retrieved April 17, 2016 from
<http://dq.cde.ca.gov/dataquest/Enrollment/EthnicEnr.aspx?cType=ALL&cGender=B&cYear=2014-15&Level=School&cSelect=Redwood+High--Tamalpais+Union+High--2165482-2132587&cChoice=SchEnrEth>
- Catley, K., Lehrer, R., & Reiser, B. (2005). Tracing a proposed learning progression for developing understanding of evolution. Paper Commissioned for the Committee on Test Design for K-12 Science Achievement Center for Education, National Research Center. Retrieved March 25, 2015 from:
<https://northwestern.academia.edu/BrianReiser>
- Cobern, W.W., Schuster, D., Adams, B., Applegate, B., Skjold, B., Undreiu, A, Loving, C.C., & Gobert, J.D. (2010) Experimental comparison of inquiry and direct instruction in science. *Research in Science & Technological Education*, 28:1, 81-96, DOI: 10.1080/02635140903513599
- Duncan, R. G. & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, 46(6), 606-609. doi: 10.1002/tea.20316
- Education Data Partnership (n.d). *Redwood High School Summary*. Retrieved April 17, 2016 from <http://www.ed-data.org/school/Marin/Tamalpais-Union-High/Redwood-High>
- Ewing, M. (2015). EQuIP-ped for success: a rubric to help implement the next generation science standards. *The Science Teacher*, 82(1), 53. Retrieved March 15, 2015 from
http://go.galegroup.com.proxybz.lib.montana.edu/ps/i.do?id=GALE%7CA396767732&v=2.1&u=mtlib_1_1123&it=r&p=AONE&sw=w&asid=ec4ae3705a00fdf5eda011a9f3d66c94
- Kloser, Matthew (2014). Identifying a core set of science teaching practices: A Delphi expert panel approach. *Journal of Research in Science Teaching*, 51(3), 1185-1217. doi:10.1002/tea.21171
- Krajcik, J., McNeill, K.L., Reiser, B.J. (2008). Learning-goals driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*. 92(1), 1-32. doi: 10.1002/sce.20240

- Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the Next Generation Science Standards. *Journal of Science Teacher Education*, 25, 157-175. doi:10.1007/s10972-014-9383-2
- Krathwohl, D.H. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212-218. Retrieved March 21, 2015 from: http://www.unco.edu/cetl/sir/stating_outcome/documents/Krathwohl.pdf
- Lederman, J.S., Lederman, N.G., Bartos, S.S., Bartels, S.L., Meyer, A.A., & Schwartz, R.S. (2014). Meaningful assessment of learners' understandings about scientific inquiry – The View About Scientific Inquiry (VASI) Questionnaire. *Journal of Research in Science Teaching*, 51(1), 65-83. doi:10.1002/tea.21125
- McNeill, K.L. & Krajcik, J. (2008). Scientific Explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, 45(1), 53-78. doi:10.1002/tea.20201.
- McNeill, K.L., Pimentel, D.S., & Strauss, E.G. (2013). The impact of high school science teachers' beliefs, curricular enactments and experience on student learning during an inquiry-based urban ecology curriculum. *International Journal of Science Education*. 35, 2608-2644. doi:10.1090/095000693.2011.618193.
- Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, 25, 177-196. doi:10.1007/s10972-014-9384-1
- Pellegrino, J.W., Wilson, M.R., Koenig, J.A., & Beatty, A.S. (Eds.) (2014). *Developing Assessments for the Next Generation Science Standards*. Washington, D.C.: National Academies Press.
- Schneider, R.M., Krajcik, J., Marx, R.W., & Soloway, E. (2002) Performance of students in Project-based science classrooms on a national measure of science achievement. *Journal of Research in Science Teaching*. 39(5), 410-422. doi: 10.1002/tea.10029
- Schwarz, C.V., Reiser, B.J., Davis, E.A., Kenyon, L., Acher, A., Fortus, D., Shwartz, Y., Hug, B. & Krajcik, J. (2009) Developing a leaning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46, 632-654. doi: 10.1002/tea.20311
- Wiggins, G. & McTighe, J. (2005). *Understanding by Design, 2nd ed.* Alexandria, VA: Association for Supervision and Curriculum Development.

Wong, D., Pugh, K., & Dewey Ideas Group at Michigan State University (2001). Learning Science: A Deweyan Perspective. *Journal of Research in Science Teaching*, 38 (3), 317-336. Retrieved March 15, 2015 from: <https://www.msu.edu/~dwong/publications/JRST-WongDeweyPersp.PDF>

APPENDICES

APPENDIX A

INSTRUCTIONAL PLAN AND ALIGNMENT INSTRUMENT

TREATMENT UNIT – PHOTOSYNTHESIS AND CELL RESPIRATION

<p><i>Performance Expectation(s):</i> <u>HS-LS1-5.</u> Use a model to illustrate how photosynthesis transforms light energy into stored chemical energy. <u>HS-LS1-7.</u> Use a model to illustrate that cellular respiration is a chemical process whereby the bonds of food molecules and oxygen molecules are broken and the bonds in new compounds are formed resulting in a net transfer of energy.</p>		
Instructional Activity	NGSS Alignment	Evidence
Photosynthesis Reading Analysis	DCI: LS1.A; LS1.C	Textbook reading activity with questions significantly revised to be more open-ended and aligned to practices.
	SEP: AQ; CE; OI	
Photosynthesis Concept Mapping	DCI: LS1.A; LS1.C	New activity designed to align with the practice of developing models in the context of direct instruction.
	SEP: DUM; OI	
PS, CR and Primary Productivity Lab	DCI: LS1.A; LS1.C; LS2.B	Re-designed experiment using mathematical thinking to evaluate experimental connection between cell respiration and photosynthesis. Analysis questions were also revised to focus on argumentation.
	SEP: PCI; AID; MCT; AE	
Cell Respiration Reading Analysis	DCI: LS1.A; LS1.C	Textbook reading activity with questions significantly revised to be more open-ended and aligned to practices
	SEP: AQ; CE; OI	
Cell Respiration Direct Instruction	DCI: LS1.A; LS1.C	Lecture was revised to include a modeling activity similar to the photosynthesis concept mapping.
	SEP: DUM, OI	
<p><i>Cross-Cutting Concepts:</i> <u>Energy and Matter</u> – This concept is central to both photosynthesis and cell respiration; <u>Systems and System Models</u> – Constructing process models for both photosynthesis and cell respiration is an essential aspect of this instructional sequence</p>		
<p><i>Alignment Adjustments:</i></p> <ul style="list-style-type: none"> • DCI components are obvious and unchanged from prior instruction, but significant revisions to incorporate SEPs in all instructional activities. • Focus on building and revising models in the direct instruction components. • Experiment revised to include a focus on mathematical thinking and using that thinking to argue from evidence 		

APPENDIX B

INSTRUCTIONAL PLAN AND ALIGNMENT INSTRUMENT

NON-TREATMENT UNIT – AGRICULTURE AND FOOD RESOURCES

<i>Performance Expectation(s):</i> <u>HS-ESS3-6</u> : Use a computational representation to illustrate the relationships among Earth systems and how those relationships are being modified due to human activity. <u>HS-LS2-7</u> : Design, evaluate and refine a solution for reducing the impacts of human activities on the environment and biodiversity.		
Instructional Activity	NGSS Alignment	Evidence
Human Food Production Reading Analysis	<i>DCI</i> : LS2.A; LS2.B; ESS3.C	Analysis of background reading for content information, focusing on farm as an ecosystem and the role of inputs – irrigation, fertilizer, pesticides – on that system.
	<i>SEP</i> : OI	
Agricultural Practices Jigsaw Activity	<i>DCI</i> : LS2.A; LS2.B; ESS3.C	Research-based activity, students share information on agricultural inputs to develop common model of an agricultural ecosystem.
	<i>SEP</i> : OI; DUM	
Fertilizers, Nitrogen Cycle and Food Production Direct Instruction	<i>DCI</i> : LS2.A; LS2.B; ESS3.C	Lecture and related reading on the pros and cons of fertilizer use, including its role in the global nitrogen cycle. This direct instruction activity does not include any significant work in the practices.
	<i>SEP</i> : None	
Light, Fertilizer, Primary Productivity Lab	<i>DCI</i> : LS2.A; LS2.B; ESS3.C	Students conduct, but do <i>not</i> design, a class wide experiment examining the effects of changing light and fertilizer levels on grass. Emphasis is on data analysis, and drawing conclusions. Students make connections between experimental data, primary productivity and human impacts related to agriculture.
	<i>SEP</i> : PCI; AID AE	
<i>Cross-Cutting Concepts:</i> <u>Energy and Matter</u> – This is the primary focus for this instructional sequence and is represented in each activity <u>Systems and System Models</u> – This is the primary focus of the jigsaw activity, but is less prevalent in the remaining activities.		
<i>Alignment Gaps:</i> <ul style="list-style-type: none"> • Students are not held to details of the performance expectations. For instance, they are not asked to construct a computation representation in the first performance. • The direct instruction phase is not explicitly linked to any of the practices. • The laboratory activity is prescribed for the purpose of conducting a class-wide experiment. It could be revised to be more student directed and, as a result, better reflect the ‘plan and carry out investigation’ practice. 		

APPENDIX C
INSTRUCTIONAL PLAN AND ALIGNMENT INSTRUMENT
TREATMENT UNIT – ENERGY RESOURCES

<i>Performance Expectation(s):</i>		
<u>HS-PS3-3</u> : Design, build, and refine a device that works within given constraints to convert one form of energy into another form.		
<u>HS-ESS3-2</u> : Evaluate design solutions for developing, managing and utilizing energy resources based on cost-benefit ratios.		
<u>HS-ESS3-4</u> : Evaluate or refine a technological solution that reduces impacts of human activities on natural systems.		
Instructional Activity	NGSS Alignment	Evidence
Fossil Fuels Jigsaw Research	<i>DCI</i> : ESS3.A; ESS3.C; ETS2.B	Background research on role of fossil fuels in providing energy resources, evaluated using the resource use cycle. Presentations include evaluating costs and benefits of each fossil fuel in this context.
	<i>SEP</i> : OI; AQ; CE	
Fossil Fuels Driving Questions: Self and Peer Evaluations	<i>DCI</i> : ESS3.A; ESS3.C; ETS2.B	Students use data and explanations from jigsaw activity – above – to argue their opinions/conclusion about the use of fossil fuels for human energy needs.
	<i>SEP</i> : AE	
Renewable Energy Reading and Video Analysis	<i>DCI</i> : ESS3.A; ESS3.C; ETS2.B	Students use background resources to evaluate the science and application of solar photovoltaic energy, including developing model of solar cell function.
	<i>SEP</i> : OI; DUM; AQ	
Solar Energy Lab	<i>DCI</i> : ESS3.A; ESS3.C; ETS2.B	Students design and implement an experiment to evaluate factors influencing efficiency of solar photovoltaic cells. They use this information to evaluate using solar PV for their home energy needs.
	<i>SEP</i> : PCI; AID; MCT; AE; OI	
Renewable Resources Driving Questions: Self and Peer Evaluation	<i>DCI</i> : ESS3.A; ESS3.C; ETS2.B	Students use data from experiment and explanations from background research to argue their opinions/conclusion on the use of renewable sources for human energy needs.
	<i>SEP</i> : AE	
<i>Cross-Cutting Concepts:</i>		
<u>Energy and Matter</u> – This is the primary focus of this instructional sequence.		
<u>Systems and System Models</u> – The research focus is on a systems approach to both fossil and renewable fuels. The experiment is focused on improving efficiency of a solar system.		
<u>Cause and Effect</u> – Interpretation of the experimental data is focused on this relationship.		
<u>Scale, Proportion and Quantity</u> – Focus on measuring energy output of solar cells and their efficiencies.		
<i>Alignment Adjustments:</i>		

- All instructional activities were revised to focus on ‘argument from evidence’ regarding the cost-benefit analysis of energy resources
- Background research components focus on using modeling to better understand the evidence to build individual arguments.

APPENDIX D

INSTRUCTIONAL PLAN AND ALIGNMENT INSTRUMENT

NON-TREATMENT UNIT – ENERGY CONCEPTS

<p><i>Performance Expectation(s):</i> <u>HS-PS3-1</u>: Create a computational model to calculate the change in the energy of one component in a system when the change in energy of the other component(s) and energy flows in and out of the system are known. <u>HS-PS3-4</u>: Plan and conduct an investigation to provide evidence that the transfer of thermal energy when two components of different temperature are combined within a closed system results in a more uniform energy distribution among the components in the system (second law of thermodynamics).</p>		
Instructional Activity	NGSS Alignment	Evidence
Global Resource Analysis	<i>DCI</i> : ESS3.A; ESS3.C	Students engage in an introductory, inquiry-based activity exploring human energy use, analyzing provided data to evaluate data patterns and pose questions for further study.
	<i>SEP</i> : AQ; AID	
Energy Forms, Transformations and Efficiency Reading Analysis + Practice	<i>DCI</i> : PS3.A; PS3.B	Students do background reading, including directed practice in calculating energy conversions and efficiencies. The calculations are largely procedural and require limited mathematical thinking.
	<i>SEP</i> : MCT	
Energy Forms, Transformations and Efficiency Direct Instruction	<i>DCI</i> : PS3.A; PS3.B	This activity is direct instruction, conducted primarily to confirm and reinforce the background reading. There is no real use of the practices beyond checking calculations from the previous activity.
	<i>SEP</i> : None	
Generating Electricity Direct Instruction	<i>DCI</i> : PS3.A; PS3.B	Direct instruction using generation of electrical energy as an application of the principles from the previous lecture.
	<i>SEP</i> : None	
Energy Transformations Lab	<i>DCI</i> : PS3.A; PS3.B; ETS1.A; ETS1.B	Students design and conduct a class-wide experiment to optimize efficiency in an electrical to heat energy system. They must analyze the ensuing data and draw connections from the experiment to larger science concepts.
	<i>SEP</i> : PCI; AID; MCT; AE	
<p><i>Cross-Cutting Concepts:</i> <u>Energy and Matter</u> – Primary focus of this instructional sequence and the overall unit <u>Systems and System Models</u> – Experiment focused on improving system efficiency, requiring a systems approach <u>Cause and Effect</u> – Interpretation of experimental data is focused on this relationship <u>Scale, Proportion and Quantity</u> – Focus on measuring energy and calculating energy transformations and energy efficiencies</p>		
<p><i>Alignment Gaps:</i></p> <ul style="list-style-type: none"> • Instructional sequence was reasonably aligned to the performance expectations. • Experiment is aligned to DCI and SEP and includes aspects of engineering design. 		

- Introductory activity is misaligned to the PS DCI, but is aligned to the SEP
- Direct instruction activities are aligned to the DCIs, but do not include any SEPs.

APPENDIX E

PHOTOSYNTHESIS/CELL RESPIRATION PRE-TEST

1. Photosynthesis uses sunlight to convert water and carbon dioxide into
 - a. oxygen
 - b. high energy sugars
 - c. oxygen and high-energy sugars
 - d. ATP and oxygen

2. The balanced chemical equation for photosynthesis is
 - a. $(6)\text{CO}_2 + (6)\text{O}_2 + \text{light energy} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + (6)\text{H}_2\text{O}$
 - b. $\text{C}_6\text{H}_{12}\text{O}_6 + (6)\text{O}_2 + \text{light energy} \rightarrow (6)\text{H}_2\text{O} + (6)\text{CO}_2$
 - c. $(6)\text{H}_2\text{O} + (6)\text{O}_2 + \text{light energy} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + (6)\text{CO}_2$
 - d. $(6)\text{H}_2\text{O} + (6)\text{CO}_2 + \text{light energy} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + (6)\text{O}_2$

3. Which of the following is not a product of the light-dependent reactions?
 - a. oxygen gas
 - b. PGAL
 - c. NADPH
 - d. ATP

4. The Calvin cycle is another name for the
 - a. light-independent reactions
 - b. light-dependent reactions
 - c. photosynthesis
 - d. electron transport chain

5. Which of the following is a product of the Calvin cycle?
 - a. oxygen gas
 - b. high-energy sugars
 - c. ATP
 - d. carbon dioxide gas

6. The light independent reactions of photosynthesis can occur in a test tube that is supplied with
 - a. ATP and NADPH
 - b. ATP and NADP+
 - c. ADP and NADPH
 - d. ADP and NADP+

7. When chlorophyll molecules absorb light energy, the electrons in those molecules
 - a. lose most of their energy as heat
 - b. lose most of their energy as light
 - c. are raised to a higher energy level
 - d. fall to a lower energy level

8. Each of the following is true of photosynthesis except that
 - a. O_2 is produced by the light dependent reactions
 - b. H_2O is broken down in the light independent reactions
 - c. CO_2 is utilized by the light independent reactions
 - d. ATP is produced by the light dependent reactions

9. The light independent reactions of photosynthesis use the energy stored in the
 - a. light dependent reactions to make glucose
 - b. light dependent reactions to make carbon dioxide
 - c. light independent reactions to make glucose
 - d. light independent reactions to make carbon dioxide

10. What are the reactants in the overall equation for cellular respiration?
 - a. oxygen and lactic acid
 - b. carbon dioxide and water
 - c. glucose and oxygen
 - d. water and glucose

11. Cellular respiration uses one molecule of glucose to produce
 - a. 2 ATP molecules
 - b. 34 ATP molecules
 - c. 36 ATP molecules
 - d. 38 ATP molecules

12. How are cellular respiration and photosynthesis almost opposite processes?
 - a. photosynthesis releases energy, and cell respiration stores energy
 - b. photosynthesis removes CO₂ from the atmosphere, cell respiration puts CO₂ into the atmosphere
 - c. photosynthesis removes O₂ from the atmosphere, cell respiration puts O₂ into the atmosphere
 - d. photosynthesis produces ATP and cell respiration consumes ATP

13. Photosynthesis is to chloroplasts as cell respiration is to
 - a. chloroplasts
 - b. cytoplasm
 - c. mitochondria
 - d. nuclei

14. Unlike photosynthesis, cell respiration occurs in
 - a. animal cells only
 - b. plant cells only
 - c. all prokaryotic cells
 - d. all eukaryotic cells

15. The products of photosynthesis are the
 - a. products of cell respiration
 - b. reactants of cell respiration
 - c. products of glycolysis
 - d. reactants of the Krebs cycle

Short Answer Questions.

1. Briefly describe photosynthesis and its role in energy transfer from sunlight to food.
2. Briefly describe aerobic cellular respiration and its role in energy transfer from food to the energy used by cells.
3. Compare and contrast aerobic cellular respiration and photosynthesis.
4. An Integrated Science Honors student at Redwood eats a banana during class. What is the role of photosynthesis and aerobic cellular respiration in that activity – for both the banana and the student?

APPENDIX F

ASSESSMENT ALIGNMENT INSTRUMENT

TREATMENT UNIT – PHOTOSYNTHESIS AND CELL RESPIRATION

Test Item	DCI Alignment	Conceptual Level	Difficulty Level
1. Photosynthesis uses sunlight to convert water and carbon dioxide into	LS1.C	MM	3
2. The balanced chemical equation for photosynthesis is	LS1.C	AC	1
3. Which of the following is not a product of the light-dependent reactions?	LS1.C	AC	2
4. The Calvin cycle is another name for	LS1.C	AC	1
5. Which of the following is a product of the Calvin cycle?	LS1.C	AC	3
6. The light independent reactions of photosynthesis can occur in a test tube that is supplied with	LS1.C	MM	3
7. When chlorophyll molecules absorb light energy, the electrons in those molecules	LS1.A	AC	3
8. Each of the following is true of photosynthesis except that	LS1.C	MM	3
9. The light independent reactions of photosynthesis use the energy stored in the	LS1.C	AC	3
10. What are the reactants in the overall equation for cellular respiration?	LS1.C	AC	2
11. Cellular respiration uses one molecule of glucose to produce	LS1.C	AC	2
12. How are cellular respiration and photosynthesis almost opposite processes?	LS1.C LS2.B	MM	3
13. Photosynthesis is to chloroplasts as cell respiration is to	LS1.A LS1.C	MM	2
14. Unlike photosynthesis, cell respiration occurs in	LS1.A LS1.C	MM	3
15. The products of photosynthesis are the	LS1.C LS2.B	MM	3
SA 1. Briefly describe photosynthesis and its role in energy transfer from sunlight to food.	LS1.C LS2.B	AP	4
SA 2. Briefly describe aerobic cellular respiration and its role in energy transfer from food to the energy used by cells.	LS1.C LS2.B	AP	4
SA 3. Compare and contrast aerobic cellular respiration and photosynthesis.	LS1.C	MM	3
SA 4. An Integrated Science Honors student at Redwood eats a banana during class. What is the role of photosynthesis and aerobic cellular respiration in that activity – for both the banana and the student?	LS1.C	AP	4

APPENDIX G

STUDENT LABORATORY REPORT AND SCORING RUBRIC

EFFECT OF LIGHT AND FERTILIZER ON PRIMARY PRODUCTIVITY LAB

The Effect of Light and Fertilizer on Primary Productivity
Integrated Science 3 Honors

Name: _____ Per: _____

Introduction

Use information from this lab handout, your textbook and other research materials to write an Introduction to the experiment. Your writing should be both concise and precise, while following the standard conventions of written English. It must include appropriate reference citations. This section should convey the following information:

- Background information about both IV and DV
- Relationships between IV and DV
- Relationships between your experiment and prior research
- Importance/ relevance of your experiment
- Expected results of your experiment

Data Analysis

1. Include your experimental design outline and laboratory notebook with this assignment.
2. Conduct the appropriate descriptive statistics for the class data. Construct a data table to display your results.
3. Construct an appropriate graph of the class data. The graph must allow you to evaluate your hypothesis.

Conclusions

Your conclusions will be presented as a Discussion section from our standard laboratory report format. Your writing should be both concise and precise, while following the standard conventions of written English. It should convey the following information (this is almost identical to the analysis questions from the Photosynthesis, Cell Respiration and Primary Productivity lab):

- What happened in your experiment? Use qualitative observations (lab notes/notebook), quantitative data and relevant statistics to explain whether your hypothesis was supported, refuted or inconclusive?
- Why did these results occur? Explain how light, fertilizer and net primary productivity are related to your experimental data– both qualitative and quantitative.
- Why did these results occur? Explain how the experimental design and experimental procedures influenced your results – both qualitative and quantitative.

- What are some implications of this experiment? Discuss the relationship between light, fertilizers, agricultural practices and human food resources.
- What's next? Are there any new questions/experiments that are suggested by your results? Are there any possible improvements/revisions to your experimental design or experimental procedures?

The Effect of Light and Fertilizer on Primary Productivity

Integrated Science 3 Honors

Scoring Rubric

Introduction

Use information from this lab handout, your textbook and other research materials to write an Introduction to the experiment. Your writing should be both concise and precise, while following the standard conventions of written English. It must include appropriate reference citations. This section should convey the following information:

- Background information about both IV and DV
- Relationships between IV and DV
- Relationships between your experiment and prior research
- Importance/ relevance of your experiment
- Expected results of your experiment

SEP: Obtain, Evaluate and Communicate Information

DCI: LS2.B; ESS3.C

- 4= Experiment is supported by accurate, relevant and properly cited background research on purpose, variables and predicted outcomes; Information is clear, accurate, thorough and in the student's own words.
- 3= Experiment is supported by mostly accurate, relevant and cited background research on purpose, variables and predicted outcomes; Information is generally accurate, thorough and in the student's own words.
- 2= Experiment is partially supported by relevant and/or cited background research on purpose, variables and predicted outcomes; Information is explained with partial accuracy, generally in the student's own words
- 1= Experiment is not supported by appropriate background research on purpose, variables and predicted outcomes; Information is generally inaccurate and/or not clearly explained.

Data Analysis

1. Include your experimental design outline and laboratory notebook with this assignment.

SEP: Ask Questions; Plan/Conduct Experiment

- 4= Design outline is thorough and accurate; Hypothesis is clearly worded to predict data trends; Lab Notebook is detailed.
- 3= Design outline is accurate but lacking some detail; Hypothesis is accurate but lacking some clarity; Lab notebook is adequate
- 2= Design outline is has some inaccuracies; Hypothesis lacks clarity and/or accuracy in predicting data trends; Lab notebook is not complete or is lacking in details

1= Design outline is notably flawed or is not complete; Hypothesis is not accurate; Lab notebook is absent

2. Conduct the appropriate descriptive statistics for the class data. Construct a data table to display your results.

SEP: Analyze and Interpret Data; Mathematical and Computational Thinking

4= Descriptive statistics are appropriate, complete and accurate; they are displayed correctly

3= Descriptive statistics are appropriate and generally accurate; they are displayed correctly

2= Descriptive statistics may be inappropriate or inaccurate; there are flaws in the display format

1= Descriptive statistics have major inaccuracies or are incomplete

3. Construct an appropriate graph of the class data. The graph must allow you to evaluate your hypothesis.

SEP: Analyze and Interpret Data; Mathematical and Computational Thinking

4= Graphs are appropriate to the hypothesis and accurate in representing the experimental data; graph is properly formatted and aesthetically pleasing.

3= Graphs are appropriate to the hypothesis but may be lacking in accuracy or aspects of formatting

2= Graphs are not appropriate to the hypothesis and may be lacking in accuracy or formatting

1= Graphs are incomplete or absent

Conclusions

Your conclusions will be presented as a Discussion section from our standard laboratory report format. Your writing should be both concise and precise, while following the standard conventions of written English. It should convey the following information (this is almost identical to the analysis questions from the Photosynthesis, Cell Respiration and Primary Productivity lab):

- What happened in your experiment? Use qualitative observations (lab notes/notebook), quantitative data and relevant statistics to explain whether your hypothesis was supported, refuted or inconclusive?

SEP: Analyze and Interpret Data; Construct Explanations; Argument from Evidence

4= Hypothesis is accurately evaluated; conclusions are thoroughly supported by specific data – both quantitative and qualitative – from the experiment; quantitative data includes reference to appropriate descriptive statistics – both measures of central tendency and variation.

3= Hypothesis is accurately evaluated; conclusions are supported by specific, but possibly incomplete, experimental data

2= Hypothesis is accurately evaluated, but conclusions are lacking appropriate support from experimental data

1= Hypothesis is not completely or accurately evaluated; conclusions are not supported by experimental data

- Why did these results occur? Explain how light, fertilizer and net primary productivity are related to your experimental data– both qualitative and quantitative.

SEP: Construct Explanations; Argument from Evidence

DCI: LS2.B; ESS3.C

4= Experimental results are thoroughly explained; connections are clearly drawn between the experimental results and concepts of light, fertilizers and net primary productivity.

3= Experimental results are explained; connections are made between experimental results and the concepts of light, fertilizers and net primary productivity

2= Experimental results are partially explained; some connections are made between experimental results and the concepts of light, fertilizers and net primary productivity

1= Experimental results are not explained; minimal or no connections are made between experimental results and the concepts of light, fertilizers and net primary productivity

- Why did these results occur? Explain how the experimental design and experimental procedures influenced your results – both qualitative and quantitative.

SEP: Design and Conduct Experiments; Analyze and Interpret Data; Argument from Evidence

- 4= Experimental results are thoroughly explained; connections are clearly drawn between the experimental results – both quantitative and qualitative – and aspects of experimental design and procedures.
- 3= Experimental results are explained; connections are made between the experimental results – both quantitative and qualitative – and aspects of experimental design and/or procedures.
- 2= Experimental results are partially explained; some connections are made between experimental results and aspects of experimental design or procedures.
- 1= Experimental results are not explained; minimal or no connections are made between experimental results and experimental design or procedures
- What are some implications of this experiment? Discuss the relationship between light, fertilizers, agricultural practices and human food resources.

SEP: Construct Explanations; Argument from Evidence

DCI: LS2.B; ESS3.C

4= Relationships between light, fertilizers, agricultural practices and human food resources are accurately and thoroughly explained; writing is both clear and concise.

3= Relationships between light, fertilizers, agricultural practices and human food resources are explained; writing is generally clear but may lack some focus

2= Relationships between light, fertilizers, agricultural practices and human food resources are partially or inaccurately explained; writing lacks some clarity and/or focus

1= Relationships between light, fertilizers, agricultural practices and human food resources are not explained; writing not clear or concise

- What's next? Are there any new questions/experiments that are suggested by your results? Are there any possible improvements/revisions to your experimental design or experimental procedures?

SEP: Plan and Conduct Experiments; Construct Explanations; Argument from Evidence

4= Analysis of possible errors, improvements and ideas for further study are accurate, insightful and creative; Explanations are exceedingly clear and concise

3 = Analysis of possible error, improvements and ideas for further study are generally accurate and insightful; Explanations are clear and concise

2= Analysis of possible error, improvements and ideas for further study may be present, but are incomplete; Explanations lack clarity

1= Analysis of possible error, improvements and ideas for further study are minimal or missing

APPENDIX H

STUDENT QUESTIONNAIRE – INSTRUCTION/ASSESSMENT ALIGNMENT

Name _____

- Participation is optional.
- Please answer these questions as honestly as possible.
- Circle your answer for each question.

1. Rate the overall difficulty level of this assessment.

4	3	1	1
Difficult			Easy

Think about the work we have done on photosynthesis and cell respiration – lectures, readings and homework, class activities and labs – and how they have contributed to your learning.

2. Rate how well the combination of all these activities prepared you for this assessment.

4	3	1	1
Well Prepared			Not Prepared

3. Rate how well the in-class activities (lectures and other non-lab activities) prepared you for this assessment.

4	3	1	1
Well Prepared			Not Prepared

4. Rate how well the lab activities prepared you for this assessment.

4	3	1	1
Well Prepared			Not Prepared

5. Rate how well the readings and homework prepared you for this assessment.

4	3	1	1
Well Prepared			Not Prepared

6. What was most effective for your learning?

Why was this the case?

7. What was least effective for your learning?

Why was this the case?

APPENDIX I
IRB EXEMPTION



INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 0000165

960 Technology Blvd. Room 127
 c/o Immunology & Infectious Diseases
 Montana State University
 Bozeman, MT 59718
 Telephone: 406-994-6783
 FAX: 406-994-4303
 Email: cherylj@montana.edu

Chair: Mark Quinn
 406-994-5721
 mquinn@montana.edu
Administrator:
 Cheryl Johnson
 406-994-6783
 cherylj@montana.edu

MEMORANDUM

TO: Todd Samet and Walter Woolbaugh
FROM: Mark Quinn *Mark Quinn CJ*
DATE: October 15, 2015
RE: "The Impact of Curricular and Instructional Revisions Aligned to Next Generation Science Standards on Student Learning" [TS101515-EX]

The above research, described in your submission of **October 13, 2015**, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

- (b) (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
- (b) (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.
- (b) (3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if: (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.
- (b) (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available, or if the information is recorded by the investigator in such a manner that the subjects cannot be identified, directly or through identifiers linked to the subjects.
- (b) (5) Research and demonstration projects, which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.
- (b) (6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed, or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the FDA, or approved by the EPA, or the Food Safety and Inspection Service of the USDA.

Although review by the Institutional Review Board is not required for the above research, the Committee will be glad to review it. If you wish a review and committee approval, please submit 3 copies of the usual application form and it will be processed by expedited review.

Looks great.