EVALUATING THE EFFECTS OF A STORYLINE INSTRUCTIONAL APPROACH ON BIOLOGY STUDENT PERFORMANCE AND ATTITUDES

by

Shane Michael Cullian

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 2022

DEDICATION

This paper is dedicated to the numerous science teachers throughout my educational journey that have given me an appreciation for the natural world. These individuals have taught me to think critically, to be objective and empathetic in my approach to teaching learning, and to understand that science teachers function as both educators and stewards to this incredible planet. The teachers to credit for this profound impact include Jody Nesheim, Robert Huffman, Erin Palese, Steven Rick, Jeannine Stantesley, Nicholas Maravalo PhD., Lynn Schadrie, and Kara Rowbotham. I would also like to thank Dr. Jason Crean for all his work and willingness to share his insights regarding storylines with me over the past two years. He remains an inspiration to me as a young science teacher and cannot thank him enough for his help in this process.

I would also like to dedicate this paper to my most loving and discerning critics. My mother Stacey Cullian, my father Mike Cullian, and my brother Connor Cullian, have provided an immense amount of emotional support since the day I applied to the program (always believing their son and brother should have originally chosen to become a Bobcat as an undergrad!) I would also like to thank my partner, Carly O'Brien, for the long nights of proofreading, assisting in transcribing interviews, and for taking time out of her daily life to help see me to the finish line. I could not have accomplished this task without their guidance and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This study was conducted to determine whether a storyline approach to teaching high school biology had an effect on student performance and attitude toward science in comparison to a traditional approach. Three sections of biology and honors biology classes were exposed to traditional science instruction as a non-treatment unit, which preceded a treatment unit that incorporated a storyline approach by using concepts such as anchoring phenomena, collaborative sense-making using science practices, and whole-group discussions directed towards constructing knowledge. Scores from each unit's pre- and post-tests were compared to determine any significant difference between mean and median values, in addition to an analysis of survey data and focus-group interviews. Results indicated that students performed better on the non-treatment pre/post tests and statistical analysis show that these differences are significant. Results also indicated that students performed worse on the first treatment posttest, yet marginally better on the second treatment posttest compared to the pre-test. Statistical analysis show that these differences are both significant as well. Student attitudes toward learning also increased as a result of this study.

CHAPTER ONE

INTRODUCTION AND BACKGROUND

Context of the Study

Throughout my first five years of teaching biology at Williams Bay High School, I have observed incoming freshman struggle with not only understanding core biological ideas, but also lack confidence in their ability to conduct scientific practices and identify crosscutting concepts in their thinking. Their initial perception of biology, in addition to science education in general, has been reduced to a survey of loosely-connected topics in a sequence that makes sense to the teacher, but little-to-no sense through the eyes of students. By viewing science as an immense body of knowledge to memorize that only few can master, many of my students have expressed frustration through exasperated claims such as “I can’t do science,” or “science just isn’t for me.” Like mathematics, the popular thought has been that science as a discipline is inaccessible to the average person. For the past half-decade, it is my opinion that I have directly facilitated this concept by teaching biology by way of guided notes, vocabulary quizzes, and assessments that ask students to regurgitate memorized facts. Coupled with “cookbook” labs and a curriculum that was not truly three-dimensional or aligned to the *Next Generation Science Standards* (NGSS), these circumstances have led me to question whether the status quo of my instructional methods were truly effective in terms of student understanding and how students perceive themselves in the science classroom.

Since enrolling in MSSE, I have become more familiar with the NGSS, the nature of inquiry, and how the 21st-century science teacher should be less of a lecturer and instead adopt the role of a facilitator. The incorporation of Classroom Assessment

Techniques (CATs) into my teaching, aligning my lessons to the standards, and other methods I have acquired over the past two years has encouraged me to make an instructional shift to not only benefit student learning outcomes, but hopefully move the needle closer to best-practice on my behalf. During the 2020-2021 school year, I piloted a unit centered around a storyline approach to teaching biology. A storyline is a coherent unit, or sequence of lessons, in which engagement is driven by student-generated questions about phenomena (Reiser et al., 2016). The questions are organized into themes using a Driving Question Board, which students then use as structure for investigations into the phenomenon (Figure 1).

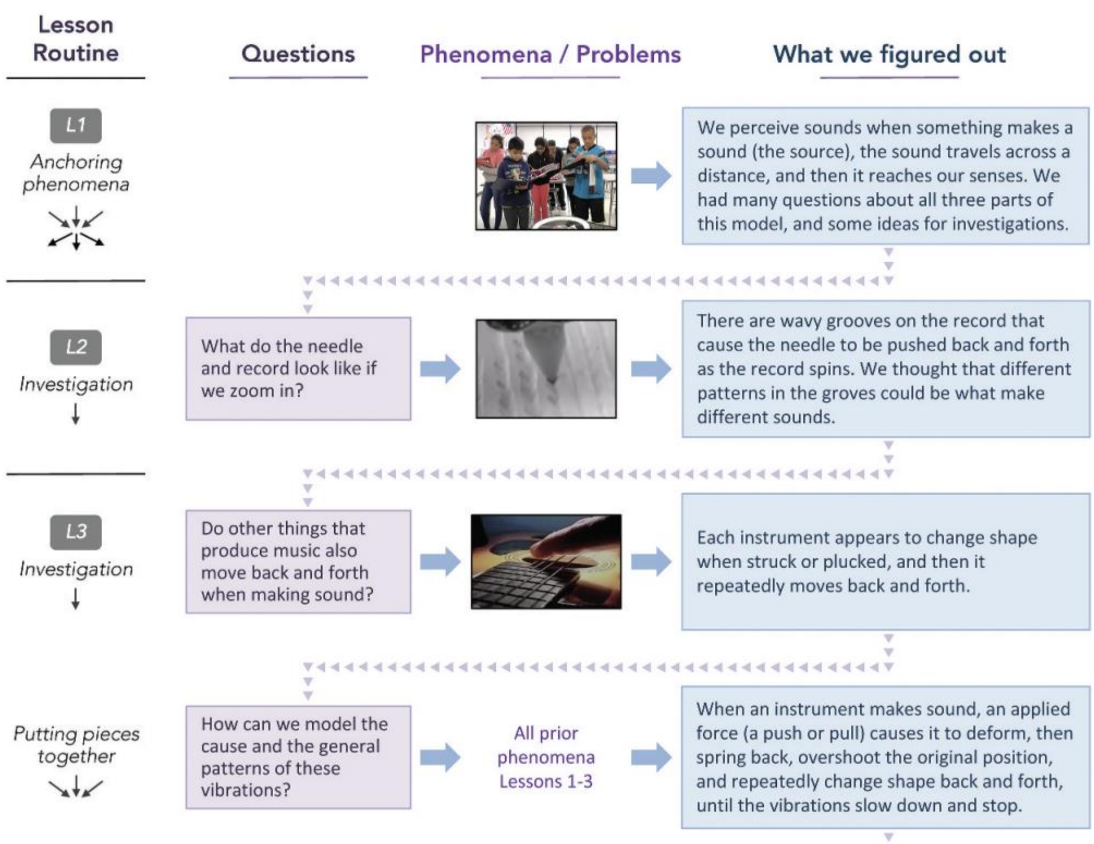


Figure 1. An example of sense-making using an anchoring phenomenon in the initial stages of a storyline (Reiser et al., 2021, p. 819).

This process can lead to independent inquiry investigations, however the Illinois Science Teaching Association's storyline curriculum that I've piloted has lessons in place for the teacher to use that address common student questions. The teacher's role is to act as both a facilitator and co-contributor within this process, encouraging students to "figure things out" by engaging in collaborative sense-making, knowledge-building, and constructing explanations within the classroom community.

Storylines promote coherence on behalf of the students, as opposed to solely through the perspective of the teacher or disciplinary expert (Reiser et al., 2021). Students are tasked with discussing their findings from the previous class period and prioritizing new questions as they drive the learning, as opposed to the previous teacher-driven approach. This concept of epistemic agency on behalf of the student is directly in line with having students engage in Science & Engineering Practices of the NGSS through the context of disciplinary core ideas. Over time and through sustained engagement, student knowledge and skills build on what they discover and figure out in the science classroom. Above all else, storylines beg the question "what did *we* figure out so far?" as opposed to the traditional "what did *you* learn today?" This promotes a collaborative learning environment between peers.

An example of this collaboration can be found in the Africa Storyline from the ISTA's website, where a phenomenon from Kruger National Park prompts collaborative student questioning to kick-start investigations and sense-making activities. Aligned with specific NGSS performance tasks and standards, students figure out how organisms interact in a complex ecosystem, obtain energy for daily needs, and how these needs impact their environment (Figure 2). Daily sense-making through discussion, formative assessments, and group work culminate in an end-of-unit summative that is aligned to multiple performance tasks. Continuing into the next unit, prior knowledge and skills are spiraled into the curriculum where students continuously revisit what they've figured out in a real-world context. The use of multiple storyline units allows teachers to reuse crosscutting concepts, core ideas and science practices introduced in previous units to new contexts, allowing for the transfer of the same skills across various topics.

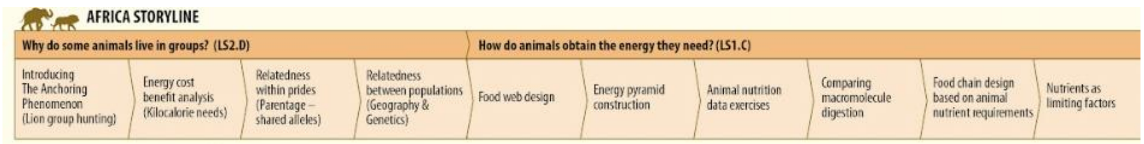


Figure 2. Partial outline of an Illinois Science Teaching Association storyline (ISTA Working Group).

Based on the previously-mentioned example, there is a noticeable difference between traditional instructional methods and the storyline approach. Whole-group discussions centered around terms introduced at the phenomenon and lesson-level replace guided notes and isolated unit-specific vocabulary instruction. Labs with a predetermined outcome are swapped with investigations that are contextually aligned to the anchoring phenomenon, leading to a richer understanding by asking students to explain how it connects to the greater storyline. Summative assessments replace standardized tests where students are asked to make claims and provide evidence using data from previous lessons. Experiencing this difference in real-time has changed how I view my students' role in the classroom and my own responsibility as their science teacher. It provides an opportunity to not only change my practice, but to communicate the results of this shift with others who are also seeking to change their teaching and interested in learning more about this type of three-dimensional instruction.

This study was conducted at Williams Bay High School (WBHS), a fringe-rural public high school located in Williams Bay, WI. Williams Bay is a village in Walworth County, WI, located roughly 50 miles Southwest of Milwaukee, WI and 87 miles Northwest of Chicago, IL. Williams Bay High School is the only high school within the K-12 Williams Bay School District with an enrollment of 229 students. 94.3% of the population identifies as white/Caucasian, 28% of the student body is economically disadvantaged, 29% of students participate in the free and reduced lunch program and 54% of the student body are female compared to 46% male. Williams Bay High School was ranked 75th in Wisconsin based on academic achievement, graduation rate, and AP Exam scores (US News & World Report, 2022).

WBHS offers honors and general biology to all 9th graders, which also serves as a prerequisite for graduation. The courses act as an introduction to the high school science experience and emphasize a "minds-on" conceptual approach in addition to hands-on kinesthetic labs. The pace of the honors course and the depth of content-related detail are the two differentiating factors that separate a general from honors biology experience. This study incorporated my two general and one honors biology sections, with population sizes of 13, 28 and 13 students, respectively, during 1st and 2nd semesters of the 2021-2022 school year. Even though all sections engaged in a treatment and non-treatment unit, only 47 students attempted all testing components and participated in at least one survey. 26 of these students were female and 21 were male; one student had an Individualized Education Plan and one student had a 504 Plan. The non-treatment unit lasted four weeks from October 25, 2021 to November 19, 2021, whereas the treatment unit also lasted four weeks from February 7th to March 4th.

Focus Question

My focus question was, What is the effect of a storyline approach to high school biology compared to a traditional method of instruction and learning?

My sub-questions include the following:

1. How do storylines compare to the traditional approach in terms of increasing the level of content understanding among students?
2. Will student attitudes of both biology and self-efficacy in science practices differ between the storyline and traditional approaches?
3. How has the storyline instructional approach impacted my views of teaching science?

CHAPTER TWO

CONCEPTUAL FRAMEWORK

Science Education's Three-Dimensional Shift

Since the publishing of the National Research Council's *Framework for K-12 Science Education* (NRC, 2012), the focus of teaching science in the United States has shifted from *what* students are being taught to *how* they are learning. This shift has been stimulated by the Next Generation Science Standards and tasks students with engaging in science & engineering practices, using crosscutting concepts, and developing disciplinary core ideas throughout their primary and secondary science education (NGSS Lead States, 2013). The integration of the three dimensions is intended to give students the chance to explain phenomena, solve problems and build knowledge within a specific context (Krajcik et al., 2014). An NGSS-aligned approach prompts students to interact with phenomena by "doing science", as opposed to solely memorizing facts about the world. According to Krajcik et al. (2014), teachers and districts must design their instruction to help students build understanding of the NGSS performance expectations by incorporating the three dimensions in their lessons. Performance expectations are statements that describe student proficiency in science and "form standards by blending science and engineering practices, DCIs, and crosscutting concepts", specifying how students should make use of the science content knowledge (Krajcik et al. 2014, p. 161). This approach varies drastically from the conventional *sage-on-the-stage*, where students are simply asked to memorize and absorb facts about scientific knowledge. When the three dimensions are bundled together, students show that they not only *know* science concepts, but that they can apply their understanding with how they learn it through the practices in a specific context (NGSS Lead States, 2013f, p. 1).

This vision of a 21st century science classroom can be attained in a variety of ways, one of which includes addressing a necessary instructional shift to not only meet the new cognitive demands of the standards, but in turn to meet the needs of all students (NGSS Lead States, 2013d, p.1). Entirely new curriculums centered around inquiry have been developed with opportunities for students to “make-sense” of phenomena, whereas some districts have altered previous teaching practices promote epistemic agency such developing essential questions, big ideas, long-term problems or projects, and content storylines (Sikorski & Hammer, 2017). Integrating the practices, concepts and core ideas of the NGSS are at the front and center of these new approaches and there are multiple avenues that a single science department, or even individual teacher, can attempt to align their instruction with the standards. Storylines take the standards and bundle them in a way that emphasizes coherence on behalf of the learner by making learning more meaningful (Zivic et al., 2018).

Storylines and Student Coherence

Traditionally, the idea of instructional coherence has always made sense through the lens of an expert or disciplinary perspective (Reiser et al., 2016). However, what may seem straightforward and apparent to the teacher may not be the same for their students. Recent studies have expressed the need for coherence on behalf of the student to promote sense-making, knowledge-building, and overall epistemic agency. According to Reiser et al. (2016), storylines support this type of classroom culture and “require that teachers partner with students in developing and managing the trajectory” of their investigations and activities. Resier et al. (2016) argue that this deliberate approach provides substantial meaning from the student perspective because they have some degree of ownership on where they are taking their learning. As co-constructors of knowledge in partnership with the teacher, students take ownership by asking “what have *WE* figured out so far?” as a coherent way of making sense of the world around them.

Storylines are championed by a handful of theoretical topics, some stemming from over 50 years of research. Inquiry-based instruction, universal design for learning, and epistemic agency are all learner-centrist in nature. However, there is one learning model that provides an underpinning unlike any other to support this approach: constructivism. According to Yager (1991), the constructivist learning model is not simply to know about the world, but instead is “to know how to make.” When students have truly acquired an understanding about a topic or concept, they will be able to explain what they know and how they know to others. Yager (1991) implies that knowledge is not an “objective representation of an observer-independent world,” rather it is a series of “conceptual structures that epistemic agents consider viable.” This process lies at the heart of the storyline approach, where students are tasked with making sense of phenomena through investigation, explanation, and other science practices. Yager (1991) further argues that within a constructivist science classroom, a casual observer will notice a handful of features that may or may not be significant shifts from conventional practices. These procedures include emphases on “using student questions and ideas to guide lessons…promoting collaboration…using student thinking, experiences, and interests to drive lessons” and refraining from viewing science as something that exists to be mastered (Yager, 1991). The idea of constructivist learning is almost identical to the three-dimensional shift occurring in modern science education. Aligning with constructivist beliefs, storylines emphasize the coherence on behalf of the student and providing a way for kids to interpret the natural world through their own explanations. The constructivist learning model supports action research that investigates the effects of a storyline approach, driving at the core of how the behaviorist and constructivist mindsets are at an inflection point in the conversation of a relevant 21st century science education.

Effectiveness of Storylines

While methodologies comparing a storyline to a traditional teaching approach are difficult to find in the literature, there are some of articles that provide insight to this comparison. The work of Reiser et al. (2016) was previously highlighted in terms of the promotion of coherence through three dimensional storylines. Almost a decade earlier, Reiser was part of a team investigating the effect of coherence as a design principle for constructing, implementing, and evaluating a middle school physical science curriculum (Shwartz et al., 2008). Schwartz et al. designed and implemented an inquiry and problem-based curriculum referred to as IQWST within 14 urban, suburban, and rural Michigan classrooms, consisting of a diverse total population of 308 students (2008). The purpose of their study was to measure the effectiveness of a curriculum that emphasized coherence related to intra-unit and interunit learning goals. There is no mention of storylines within their study, but their data collection methods are worth mentioning. Prior to implementing the curriculum, Schwartz et al. developed pretests and posttests that included “multiple-choice and open-ended items” with “content validity and alignment” that were verified by external reviewers (2008). Results from the tests were presented within the article in table form and included scores as well as the effect size and respective p-values. The researchers also collected anecdotal feedback from the classroom teachers, but did not specify the collection methods or how the quotes were obtained. Schwartz et al. could determine whether students were recalling specific information, using reasoning to explain their thoughts, and exhibiting higher-order thinking skills as they were immersed in the IQWST curriculum (2008). Even though the NGSS had yet to be constructed, these students were engaging in practices like a storyline and immersed in a constructivist environment.

If both student and teacher perceptions are participating in an active learning environment based on co-constructed knowledge, administering brief surveys after lessons and units can also yield insight to latent student thoughts that are not easily detectable using pre-tests and post-tests. Surveys can yield insight into how student perceive their teacher, themselves, and even their feelings towards their learning outcomes in various environments.

CHAPTER THREE

METHODOLOGY

The purpose of this study was to determine if the storyline approach to teaching biology would influence student performance and attitudes toward science in comparison to a traditional, teacher-centered approach. The subjects of this study were 47 high school general and honors biology students. The research methodology for this project received an exemption by Montana State University's Institutional Review Board and compliance for working with human subjects was maintained (Appendix A).

Treatment and Non-Treatment

This study incorporated the adoption of one of the ISTA's biology storyline units and adapting the instructional materials to fit the constraints of a four-week treatment unit. The study also included traditional non-storyline instructional methods that were incorporated into a four-week non-treatment unit. A portion of the ISTA's Homeostasis Storyline served as the treatment for this study, which centers around the theme of ecosystem homeostasis and how populations are maintained within the environment. The non-treatment traditional unit also focuses on ecology; however, it introduces students to the movement of energy and matter in the environment. Lessons, activities and assessments were administered during both 90-minute block periods and 45-minute regular periods for each unit.

The non-treatment unit incorporated traditional formative assessments that were teacher-driven. The subsequent lessons following the pre-test began by posting bell-ringers or DO-NOW's on the board. Content-specific questions that were related to the day's topic were posted and students were tasked with answering these individually.

Think-pair-share discussions with elbow partners followed to discuss possible answers, culminating in a brief teacher-led discussion of the answers and time for questions, if any arose during the conversations. Students were then provided a formal lesson objective sheet with important vocabulary and guided notes that coincide with PowerPoint lectures for the day's topic. Notes were taken for 10-15 minutes at a time, followed by guided-inquiry activities with pre-determined outcomes and worksheets related to the day's topic. Vocabulary quizzes were assigned at the end of each week to assess students' ability to memorize key terms within the unit. An anchoring phenomenon was absent during this unit. The teacher determined the content, lesson objectives, and science practices that were implemented and assessed. Students were taught content at a predetermined speed and the sequence of activities, such as the order of POGILs, card-sorting activities, and the use of online formatives such as Quizlet was never discussed in terms of sense-making on behalf of the students. Collaboration and group work was not the focal point of the unit, nor was the use of discussion to "make sense" of what students were engaging with during this time. At the end of the unit, the post-test was administered, as well as survey 1.

In direct contrast, the treatment unit incorporated activities that encouraged epistemic sense-making opportunities on behalf of the students. Following the pre-test, an anchoring phenomenon video clip was shown to the class and prompted questions from the students directly. Students were asked to record their questions and, while working in pre-determined small groups of 4-5, asked to compile and rank their group's questions on order of importance to understanding the phenomenon. From there, the class organized their group's questions on a Driving Question Board in a process facilitated by the teacher. Once the questions were organized based on themes through class discussion, the students were provided with guided inquiry activities that aligned with the most important questions, allowing students to pursue answers to their questions directly. At the beginning of each subsequent day in the treatment, a student-led discussion around the driving question “What Have We Figured Out So Far?” was facilitated before moving onto the next investigation or activity. Once each class came to a consensus of what they figured out in the previous lesson, they were provided with a guiding question or two to be answered in that day’s lesson. This process of questioning, attempting an activity, formatively assessing group learning, and then discussing findings the following day repeated itself for four weeks. Students were then asked to explain the phenomenon using the science practices specific to the unit within their groups at the culmination of the treatment. This discussed preceded the implementation of post-test 1, which was the same pre-test administered at the beginning of the treatment, followed by surveys 1 and 2 and focus group student interviews. Many storyline units extend longer than the four-week timeline, to which the same post-test, titled post-test 2 in the data analysis section, was also administered six weeks following conclusion of the treatment unit. This was done in attempt to measure the transfer of science practices and content knowledge over the course of a longer period, even though the treatment instruction had ended weeks prior. A second post-test was not administered six weeks following the non-treatment unit since traditional units are typically shorter than a storyline unit. Table 1 shows a comparison between the individual components and implementation strategies of the treatment and non-treatment units.

Table 1. Treatment and non-treatment instruction and assessment components.

Non-Treatment Unit	Treatment Unit

• Administer unit pre-test. • Directly address students about the lessons objectives, lesson calendar, and when to expect certain assessments. • Lecture using PowerPoint and guided fill-in-the-blank notes. • Teach lessons and activities with pre-determined outcomes and no coherence between activities. • Assign formative assessments that include vocabulary quizzes, lab write-ups, and worksheets as homework. • Formally review with class using pre-constructed study guide, summative breakdown, and class Quizlet. • Administer unit post-test (summative)	• Administer unit pre-test. • Present anchoring or supporting phenomena • Students ask, prioritize, and organize questions pertaining to phenomena • Teacher uses questions to drive instruction through inquiry activities • Teacher engages students in daily discussion of "What Did We Figure Out So Far" as both formative assessment and check for understanding • Students continue asking questions at the end of each activity/investigation and navigate storyline via coherence • Students construct study guide based on what they figured out • Administer unit post-test (summative) • Administer Survey 1 and 2, conduct focus group interviews • Administer unit post-test 6-weeks after completed treatment

Regarding the timing for each unit, the length was given consideration prior to implementation. The Homeostasis Storyline typically takes 8 weeks to implement, whereas the topics of matter and energy within a larger ecology unit was traditionally covered in 2-3 weeks. To ensure that students spent an equal amount of time engaging in the core ideas, practices and standards for each approach, both the treatment and non-treatment units were implemented over a four-week timeframe. A single performance assessment and single standard was chosen per unit (Table 2). Performance assessments were incorporated into each units' pre/post-tests.

Table 2. Performance expectations & standards used for treatment and non-treatment.

Unit Name (Length)	Topic of Study	Performance Expectation(s)	Standard	Science & Engineering Practice(s)
Non-Treatment Unit (4 weeks)	Ecology – Energy and Matter	HS-LS2-4	Use mathematical representations to support claims for the cycling of matter and flow of energy among organisms in an ecosystem.	Using Mathematics and Computational Thinking
Treatment Unit (4 weeks)	Ecology – Ecosystem Homeostasis & Keystone Species	HS-LS2-6	Evaluate the claims, evidence, and reasoning that the complex interactions in ecosystems maintain relatively consistent numbers and types of organisms in stable conditions, but changing conditions may result in a new ecosystem.	Evaluate the claims, evidence, and/or reasoning behind currently accepted explanations or solutions to determine the merits of arguments.

Quantitative Data Collection Strategies

Both units incorporated instruments that are commonplace for collecting data in action research studies. A pre/post-test was administered at the beginning and end of each unit, corresponding to the type of instructional and learning approach that followed. The non-treatment pre/post-test was administered on paper and was adapted from pre-existing curriculum materials from Howard Hughes Medical Institute’s BioInteractive Food Webs activity. It was comprised of multiple choice, short-answer, model construction and one claim-evidence-reasoning section (Appendix B). Students were given a 45-minute class period to complete the pre/post-test and could use a calculator to calculate percentages at each trophic level. Students were tasked with creating food chains and webs using a word bank of animals, place said models in an energy pyramid, calculate the amount of energy that moved up the pyramid, and make a claim that distinguishes the different movements of matter and energy in an ecosystem by providing evidence from the student-created food chain and web models.

The treatment unit pre/post-test was adapted from the ISTA Storyline Group's assessment and administered using Google Forms and was comprised of multiple choice, short-answer, and multiple claim-evidence-reasoning questions (Appendix C). Students were given a 45-minute class period to complete the pre/post-test and could use a calculator to calculate percentages of different wolf populations over time. The 19 questions asked students to interpret figures that displayed changes in animal populations over time, identifying the carrying capacity of specific populations, and making claims about the ecological relationships between organisms using specific quantitative evidence from the figures to support said claims. The post-test was administered twice, once at the end of the treatment unit, and once six weeks after completion of the treatment unit to assess if students retained content knowledge and the ability to conduct science practices over a longer period.

All pre/post-test data were compiled and organized using Google Sheets to conduct descriptive statistics and analyze the results graphically. Means, median, mode, standard deviation and standard error were calculated for each unit and test, in addition to effect size and Wilcoxon-signed rank tests given the non-normal distribution of the data. The website Plotly was used to generate box & whisker plots of the pre/post-test results. Effect size for each test was calculated by subtracting the average pre-test score from the average post-test score, and then dividing that value by the standard deviation of the pooled pre-and post-test scores. Box & whisker plots and bar charts were created for each test. Individual bars and boxes were differentiated by the type of biology class, whether the students identified as male or female regardless of biology class type, and whether the student was considered a “top performer” on the pre-test by scoring higher than the mean.

Qualitative Data Collection Strategies

Both treatment and non-treatment units incorporated the use of surveys that were administered at the end of each unit. Survey 1 was administered at the end of both units and was comprised of Likert scale items and free response questions that asked students to elaborate on their thoughts about the instruction and approach to learning (Appendix E). Colleagues in the Williams Bay Science Department examined survey 1 for any inconsistencies and to ensure that survey questions addressed the specific sub-questions for the study.

Survey 2, which was adapted from the ISTA’s professional development materials, was administered at the end of the treatment unit and included a science practices self-efficacy scale and Likert items that referenced their attitudes before and after the treatment unit (Appendix F). The ISTA refers to it as a “Then-Post” survey given the use of the terms BEFORE and NOW in Likert items, which guarantees a consistent number of respondents answering each question in a single setting as opposed to administering a survey on two separate dates, potentially not garnering the same number of responses. All survey responses were collected via Google Forms and converted to spreadsheets using Google Sheets to conduct descriptive statistics and analyze the results graphically. Free response items were coded and organized into themes to identify any relevant patterns.

Focus group interviews were conducted for each class section and a convenience sampling method was used to determine willing participants. Included at the end of survey 1 was a question that asked if the student would be comfortable participating in an audio-recorded focus group interview. Responses to interview questions were transcribed using Google Docs and coded based on themes that were identified in the responses. Non-Likert items from both surveys and student interviews were analyzed by identifying patterns or trends in student attitudes towards the treatment. Likert data from each survey was analyzed using distribution frequencies in the form of histograms to determine normal vs. non-normal distributions. The Then-Post Survey (survey 2) Likert items were compared using stacked bar charts and a Wilcoxon-ranked test was used to compare the mean values of each response.

A teacher journal was also kept noting the daily qualitative observations within classroom discussions and overall discourse. These instruments were part of a triangulation matrix that relied on collecting both quantitative and qualitative data. Table 3 displays the different data sources aligned to specific focus and sub-questions.

Table 3. Data triangulation matrix of research questions and data sources.

Research Questions	Data Source		
	1	**2**	**3**
What is the effect of a storyline approach to high school biology compared to a traditional method of instruction and learning?	Treatment and Non-Treatment Pre/Post-Test Scores	Survey 2 (Then-Post Self-Efficacy Scale) Responses	Survey 1 (Likert items) Responses
How do storylines compare to the traditional approach in terms of increasing the level of content understanding (core ideas) among students?	Treatment Pre-and Post-Test Scores	Non-Treatment Pre- and Post-Test Scores	DO-NOW's and "What Have We Figured Out So Far" Discussions
Will student attitudes of both biology class and their science skills differ between the storyline and traditional approaches?	Survey 2 (Then-Post Self-Efficacy Scale) Responses	Survey 1 (Likert items) Responses	Focus Group Interviews
How has this three-dimensional, constructivist method impacted my approach to teaching science?	Reflective Teacher Journal	Survey 1 (Free Response Questions)	Focus Group Interviews

CHAPTER FOUR

DATA ANALYSIS

Student Performance

Utilizing pre-and post-tests to measure student performance attempted to help answer my primary research question and first sub-question. Student achievement of core content knowledge was measured by mean and median values of treatment and non-treatment pre/post-tests. Effect size was calculated by subtracting the average pre-test score from the average post-test score, and then dividing that value by the standard deviation of the pooled pre-and post-test scores. Graphical analysis was conducted using box & whisker plots and bar charts for each test. Comparisons were drawn between the type of biology class, whether the students identified as male or female regardless of biology class type, and whether the student was considered a "top performer" on the pre-test by scoring higher than the mean, regardless of biology class type.

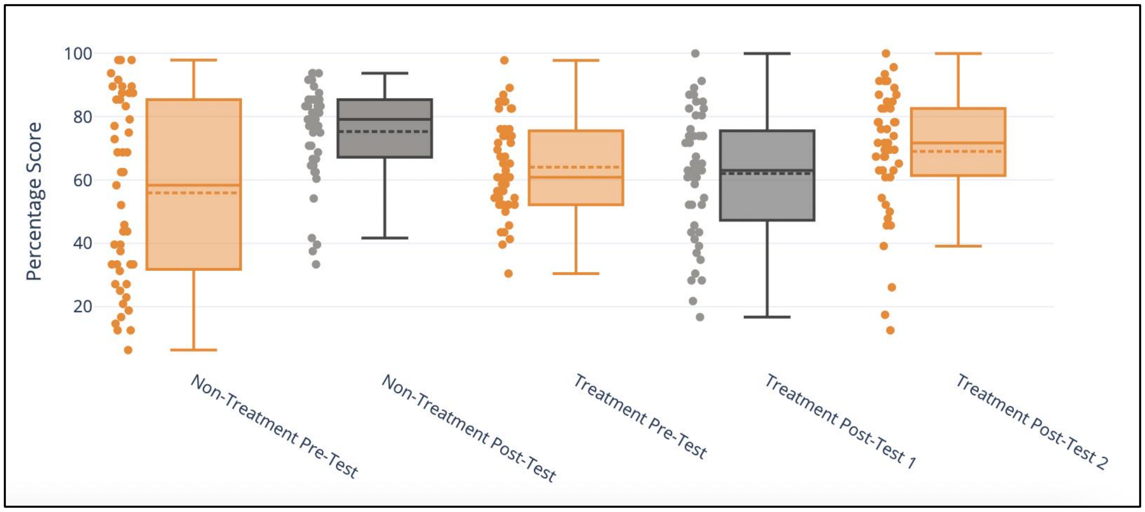


Figure 3. Pre-and post-test scores for both non-treatment and treatment units, (*N*=47).

When comparing the scores of the non-treatment pre-test to the non-treatment post-test for both honors and general biology classes, there was an increase in the median value from 58% to 79% (Figure 3). Altogether, the post-test for the non-treatment exhibited the highest overall mean and median scores, as well as the smallest range of values, of all pre/post-tests regardless of unit type. The dashed line in Figure 3 indicated the mean value increased from 56% to 75%, with 70% of the students showing an increase in their overall test score from pre-test to post-test. An effect size of 0.78 indicates the non-treatment instruction had a medium effect on student achievement. Scores from the pre/post-tests were compared using a Wilcoxon-signed rank test given the non-normal distribution. The test statistic of 139 was considerably lower than the critical value of 378, indicating a significant difference between the pre-test and post-test and that the median change was non-zero. Student achievement can be further compared by class type and gender to their respective pre/post-test scores, as seen in Figure 4. On average, female mean scores were higher on both pre-and post-tests, however males recorded slightly larger gains (20%) compared to females (18%). This could potentially be attributed to the 11 low-performing males recording greater gains on average (41%) compared to their 12 female counterparts (37%).

Honors students performed better than general biology students, however their margin of improvement was less than half of the 22% increase in mean value exhibited by the biology students. Forty-six percent of the top-performers (n=11) on the pre-test scored equal to or more than their pre-test score, whereas all 24 low-performers who scored below the mean value on the pre-test increased their score on the post-test.

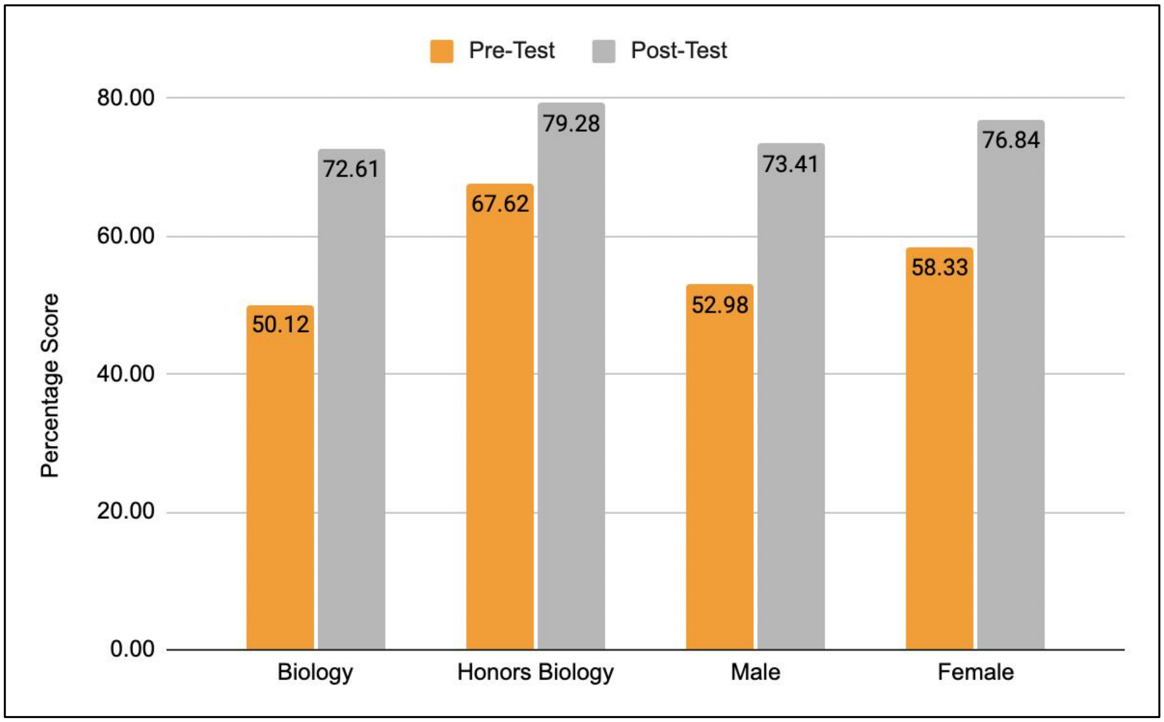


Figure 4. Comparison of non-treatment pre-and post-test scores by class type and gender, (*N*=47).

Unlike the non-treatment test results, the treatment pre-and post-tests can be interpreted differently given that the post-test was administered twice, once at the end of the unit and once after multiple weeks had passed. Referring to Figure 3, the test scores for the pre-test and post-test 1 show a decrease in mean value by 1.95% for all students, however the median value increased from 58% to 60%. Effect size for these results was determined to be -0.11, indicating that the treatment had no effect on student performance on the first post-test. A Wilcoxon-signed rank test was conducted to indicate significance given the non-normal distribution, with the test statistic 347 being lower than the critical value of 378, indicating the difference between scores was significant.

Figure 5 also breaks down student performance by class type and gender. Both male and females scored lower on average on post-test 1 compared to the pre-test. This mean decrease was also present across both class types, however the disparity between the honors scores was much more noticeable than the general biology sections. Between all class types, 22 students scored at or above the mean value of the pre-test and seven of these top performers (32%) scored equal to or more than their pre-test score on post-test 1. Thirteen of the twenty-five low-performing students (52%) increased their score on post-test 1, which was really encouraging to note given the overall class performances. A low-performing student commented on Survey 1 stating that she struggled initially with the treatment unit because she prefers "having a [study] guide and getting a starting point," as opposed to starting with a figurative blank slate. Another student elaborated on her thoughts by saying "it's hard to talk about something that you don't know how to explain." The treatment pre/post-tests involve more explanation and CER-related questions than the non-treatment pre/post-tests, which could have contributed to these frustrations and results.

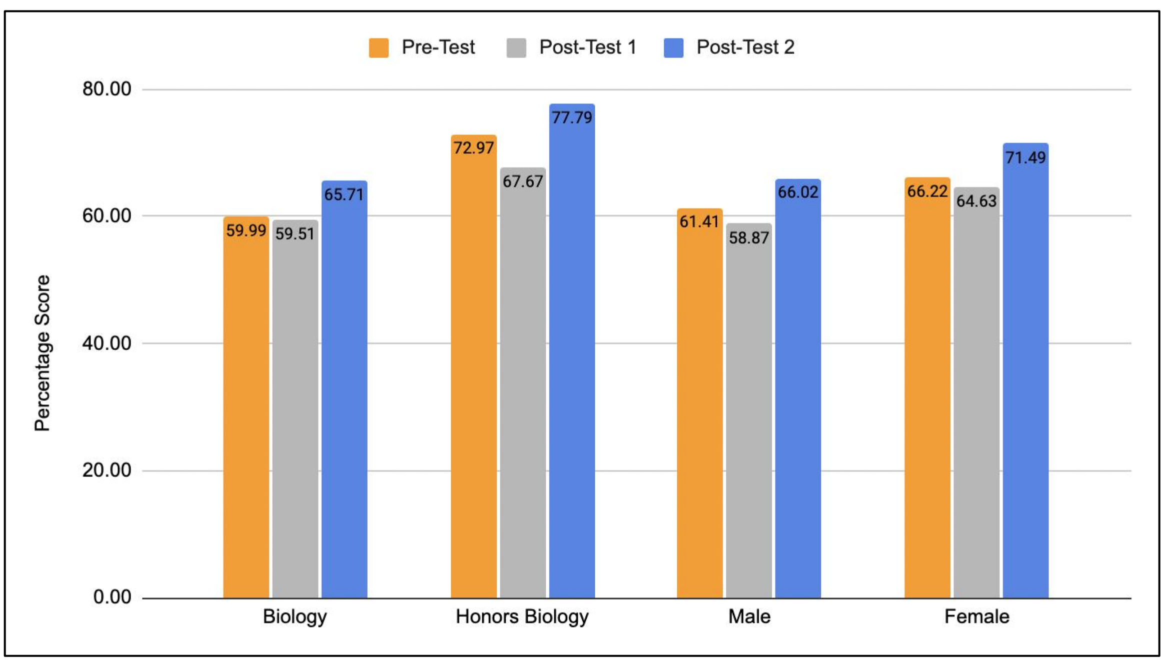


Figure 5. Comparison of treatment pre-test and post-test 1 and 2 scores by class type and gender, (*N*=47).

Just as noteworthy as the first post-test scores were the results comparing the treatment pre-test to post-test 2. Overall, Figure 3 shows students improving from the pre-test to post-test 2 as the mean values increased from 61% to 67%, while the median values increased from 58% to 69%. This resulted in a positive effect size of 0.27, which was welcome to see compared to the negative pre-test/post-test 1 value of -0.11. This indicates the treatment unit had a small effect on student performance on the second post-test. A Wilcoxon-signed rank test was also completed, with the test statistic 282 being lower than the critical value of 378, indicating the difference between scores was significant. Figure 5 shows that females on average scored higher compared to males, however their average gains were not as large as the male students. Honors students scored higher on each test on average compared to biology students, however the biology students improved from pre-test to post-test 2 by a greater margin than the honors students, with the mean value increasing by almost 6%. Of the top performing students on the treatment pre-test for all classes, 15 of the 22 (68%) scored the same or greater on post-test 2 compared to the pre-test. Eighty-percent (n=25) of the low-performing students, regardless of class, scored below the mean value on the pre-test increased their score on post-test 2.

Student Attitudes

Qualitative data was collected using Surveys 1 and 2 (Appendices E and F), in addition to focus group interview responses upon immediate completion of the treatment unit. Non-Likert items from both surveys were analyzed in the same fashion as student interviews, specifically for identifying patterns or trends in student attitudes towards the treatment. Likert data for both surveys were analyzed using distribution frequencies, specifically in the form of histograms to determine normal vs. non-normal distributions. The Then-Post Survey Likert items were compared using stacked bar charts. A Wilcoxon-ranked test was used to compare the mean values of each response.

Seventy-nine percent of students surveyed after the treatment either agreed or strongly agreed with the statement that they enjoyed biology, which was marginal increase of almost 2% compared to Survey 1's results after the non-treatment. One student supported their choice by saying that their "friends are in this class and I like this style of teaching." Another stated that they also "like working in groups and the assignments are easy." Responses from Survey 1 also show 83% of students agreeing or strongly agreeing that the sequence of lessons making sense to them after the treatment, which is an increase from 71% post-non-treatment (Figure 6). Twenty-nine percent of students either strongly agreed or agreed that they were confused during biology class post-treatment, which was a decrease from the 40% of students who were confused after the non-treatment. One student supported their choice by saying "it's placed in an order that isn't hectic at all," referring to the activities in the storyline and time allotted for each lesson. Another stated that the storyline "gives [the students] an example of how things are supposed to go and why they work the way they do."

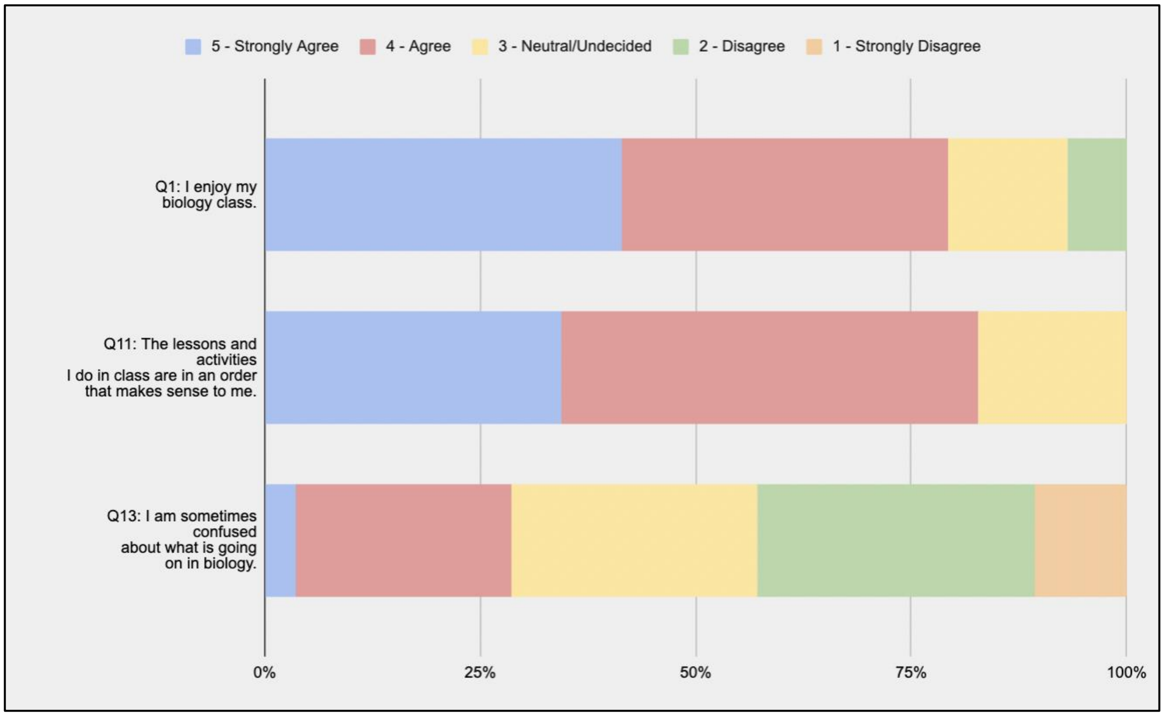


Figure 6. Confidence and enjoyment of biology results from survey 1 post-treatment, (*n*=29).

Part of the results from Survey 2 include a self-efficacy scale in student abilities to conduct scientific practices. Forty-nine percent of students stated they felt much more confident in their ability to plan and carry out investigations after completing the treatment unit, which is slightly larger than the second-highest confidence score of 45% towards constructing explanations (Figure 7). Using mathematics and computational thinking had the lowest mean value of 1.94 out of all of science practices. Overall, students exhibited more confidence in the three practices that were assessed in the pre/post-tests for both the treatment and non-treatment, regardless of test score.

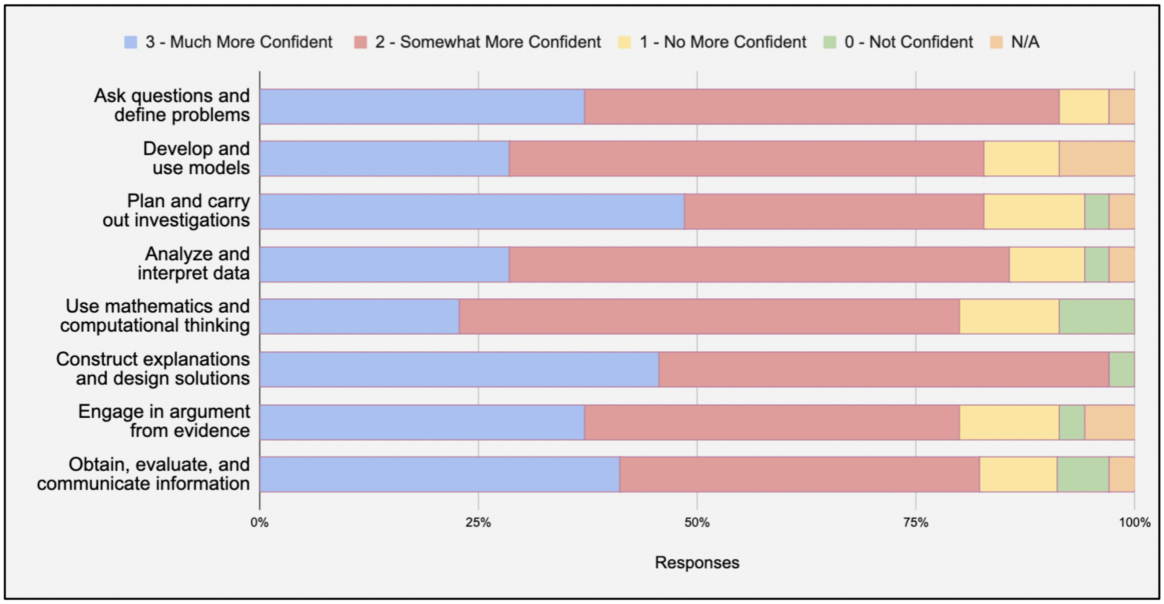


Figure 7. Science practices self-efficacy scale responses post-treatment from survey 2, (*n*=35).

Survey 2 also provided a comparison in student attitude, perception of individual understanding, and willingness to work in groups before and after the treatment unit (Figure 8). Almost 59% of students agreed or strongly agreed that they were more willing to work in a group upon experiencing the treatment unit, which was an increase from 29% who made the same statements regarding their non-treatment experience. A Wilcoxon-signed rank test of results showed sufficient evidence to suggest there is a difference in responses (z=136, p=28).

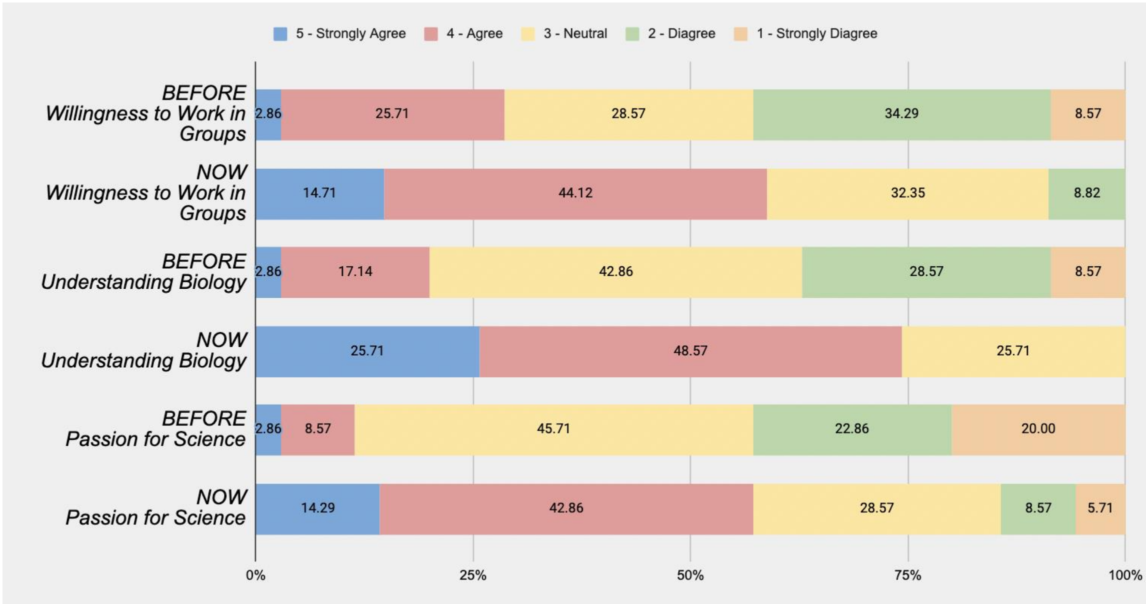


Figure 8. Student attitude towards group work, content understanding and passion for science post-treatment (*n*=35).

Another encouraging result from Survey 2 shows an increase in student perception of their general understanding of biology over time. Pre/post-tests are used to determine if students have an actual understanding of core ideas and applying science practices, however this speaks more to biology making sense to the individual student based on whether they felt it was coherent to them. Seventy-four percent of students agreed or strongly agreed that they understood biology better upon completing the storyline compared to twenty percent of students before attempting the treatment unit. The results of the Wilcoxon-signed rank test calculated a test statistic of 20, which was significantly less than the critical value of 136. While post-test 1 did not exhibit improvement in student performance compared to post-test 2, the initial perception of student understanding was greater immediately following the treatment unit. As one student claimed when asked in the interview regarding the day-by-day approach to figuring things out in a storyline, everyone is “taking it one part at a time which helps because it gives you time to memorize it, it gives you time to learn it, and it just repeats itself until you finish it."

Survey 2 also shows an increase in student passion for science over time. Initially, 11% of students agreed or strongly agreed that they had a passion for science prior to the treatment unit, which increased to 57% after completing the storyline. A Wilcoxon-signed rank test was used to determine significance and the test statistic of 37 was less than the critical value of 136, indicating a difference in responses. To coincide with the increase in passion, a 31% decrease in responses that either disagreed or strongly disagreed with being passionate for science occurred after the treatment unit.

Summary of Results

Key takeaways from quantitative data include the significant increase in student performance on the non-treatment pre/post-test, the initial decrease from the pre-test to post-test 1 of the treatment unit, as well as the eventual increase when comparing the treatment pre-test to post-test 2. Findings from qualitative data include the overall positive attitudes towards biology class, the increase level of confidence in self efficacy towards science practices after the treatment unit, and the positive attitude shift towards collaborative sense-making, passion for science, and understanding biological concepts upon the completion of the treatment unit.

CLAIM EVIDENCE AND REASONING

Claims From the Study

The primary focus of my action research was to determine whether a storyline approach is an effective instructional model in the biology classroom. Measured through understanding core ideas, attitudes towards biology and science in general, and self-confidence regarding the use of science practices, this capstone yielded insight to make a handful of claims regarding the contrast in teaching approaches.

With regards to the first two questions on the effect of a storyline compared to traditional instruction, the results of the non-treatment pre/post-test indicate that students learn effectively in a traditional science classroom. Not only did the mean and median test scores increase on average by the end of the unit for all classes, but further statistical analysis indicates that traditional methods were more effective than the storyline approach (Effect size of 0.78). A possible explanation of the results is students' familiarity with the teacher-directed style of learning in the science classroom. One of the themes I identified within the interview response is the individual nature of past science classes and the emphasis put on rote memorization of concepts, the use of guided notes to explain concepts, and the independent nature of the classes overall. For example, one general biology student stated in their interview that "other science classes were more packets, 'go sit in a corner and do a packet,'" acknowledging the sit-and-get nature of taking notes and being assessed on content. If students are prompted to treat science class as an individual activity, it should come as no surprise that, on average, they perform well in a testing environment after being exposed to direct instruction with clear learning objectives.

Another potential explanation that supports this claim lies in the change in scores from the treatment post-test 1 and the treatment post-test 2. The results imply that learning content and mastering science practices could take significant time to develop in the student population. The initial decrease in overall score from the pre-test (61%) to post-test 1 (59%) was disheartening to witness, given how positive student interview and survey responses had been to the storyline approach. Prior to administering the second post-test, it was frustrating to find positive themes from student responses, yet no manifestation of these opinions within the post-test 1 results. My third research question sought to investigate the effect a storyline had on student attitudes, which can be clearly seen in the responses to survey 1, survey 2 and the focus group interview responses. Students thoroughly enjoyed the collaborative aspect of the treatment, some interviewees going so far to say that “different opinions helps me put my opinions together” and that the storyline made “us work with kids where we usually wouldn’t work with. That prepares us for the real world when we have to get jobs and work with people that we don’t necessarily like, but we have to play nice in the sandbox.” These interview responses, as well as Likert data from Survey 2 showing an increase in willingness to work together from 29% to 59%, supports the positive sentiment resulting from the student sense-making approach When students offer candid responses to the collaborative nature of storylines, it is hard to ignore responses such as this: “when we were first in groups, that’s when I first met (student name) and now we are, like good friends.” This qualitative evidence supports a positive effect on student attitude towards science, which eventually prompted the administration of a second post-test. Post-test 2 results exhibited an increase in overall scores for all classes and genders. It cannot be confirmed that the treatment had a statistically significant effect on the second post-test, however the improvement across all groups and sections by almost 5% suggest that small gains were made in content knowledge and science practices from the beginning of the treatment to the weeks that followed the unit's conclusion.

A potential explanation of the post-test results could be the time that allotted for each unit. Since a storyline can often last longer than four weeks, it should be considered that students could use more time to spiral information, have repeated attempt for the same science practices, and have ample opportunities to explain phenomena with their peers. If the storyline is shortened, this could potentially impact student performance, regardless of how positively the instructional practices are received. Another factor that could explain the initial disparity in scores is the nature of the assessment compared to the instruction provided. Transitioning from direct instruction to having a voice in classroom discussion is another theme whose manifestation could impact test scores. One student who was interviewed felt strongly that being able to ask questions and discuss ideas is "important because, like personally for me, I struggle a lot with getting a hold of subjects quickly". Another student brought up the power of making sense in a small group, saying

> like if one person gets it wrong [in the group] then like you can just help them to understand it better. Like obviously it's going to be scary if you do get it wrong and you are worried that people are going to make fun of you but it just kind of helps you relearn the thing and make sure you know better.

The notion of being able to discuss one's constructs and individual engagement with the phenomenon, all with a certain level of comfort, could translate to a level of understanding that cannot be synthesized into an individual written assessment. At least, perhaps, not a post-test presented in such a short time that these core ideas and concepts have not had enough time to gain a foothold in student understanding.

Not every student enjoyed collaborating to make sense of phenomena, especially those like a student in my 3rd hour who finds it "frustrating when people don't pull their weight." This sentiment, coupled with similar survey data suggests that a percentage of the population prefers to make sense of material individually. However, when almost 69% (*n*=24) of the students who attempted Survey 2 indicate they preferred a storyline over traditional methods, it is safe to say that this approach has been received positively by most the population.

Value of the Study and Consideration for Future Research

Even though qualitative data speak to positive student attitudes towards the storyline approach like collaborative sense-making and student-driven inquiry, this study could benefit heavily from improving some of the methodological aspects. By shortening the length of the storyline and not having the opportunity to extend the treatment beyond a month, I feel that it provided more questions than answers. The nature of a storyline lies in providing multiple instances for students to engage with anchoring phenomena and spiraling practices and concepts through different activities at various times within a longer unit. The shorter I made the unit, the less opportunities students had to make sense of what they're figuring out. Implementing an entire traditional unit, such as a unit solely focused on the topic of cells, and comparing it to a complete storyline line unit on a multiple biological core ideas may yield more conclusive results that can be more easily compared. To echo the notion of spiraling, a storyline focused on bundling performance expectations and standards to create a coherent experience provides students with a richer experience than simply covering one topic at a time. Fully immersing a class with a longer treatment unit may provide clearer insight to these differences and the effect it has on student learning and attitude. From a logistical standpoint, using two classes as comparisons in the future, one for the storyline treatment and the other for traditional methods, could also glean insight to my research questions. Even though the sample size would shrink significantly per group, there are non-parametric tests that can be used in those instances. There is also the potential to examine specific science practices that I am assessing more closely to see if I can incorporate the same ones into both a traditional and storyline unit to measure change in student performance.

In terms of assessments and tests, instead of making small changes to previously-created assessments, I think it would benefit the future study to design my own assessment that incorporates different science practices and questions of varying levels of difficulty. This could provide clearer insight to student performance and if a storyline can help low-performing students improve their science practices. Ideally, these adjustments would allow me to develop instruments that more accurately measure student coherence. Also, providing pre/post-tests using the same medium is also an area that could use more consistency, as opposed to administering one on paper and the other on Google Forms.

Impact of Research on the Author

Attempting the capstone process and administering both treatment and non-treatment units had a significant impact on my professional experience as a science teacher. The planning, adjusting, improvising, and reflecting that went into this project is without doubt the most challenging, yet rewarding experience I have had in the classroom. Speaking to the teacher's role, experiencing the instructional shift from deliverer-of-information to discussion-facilitator opened my eyes to the importance of student voice and asking questions in the classroom. Students may not know all the minute details about a topic prior to starting class. However, if they are provided with an equitable amount of resources in the form of anchoring phenomena, a group of peers willing to inquire, and a supportive environment that fosters agency and sense-making, students can figure out many core ideas through three-dimensional instruction. Shifting the work to the students and requiring them to engage in discussion, ask questions and evaluate data with the teacher as a facilitator was a shock-to-the-system for an impressionable 6th year teacher. Providing an anchoring phenomenon gave every student in each of the three sections a chance to participate in the discussion since there were on an equitable playing field with one another. If something didn't make sense at the time, through repeated questioning and discussion with their group, regardless if groups 'played nice in the sandbox', most could eventually figure out and verbally explain the core ideas within the treatment unit. I intend on continuing this practice of student-driven questioning and discussion by incorporating phenomena at the beginning of each biology and honors biology unit in the years to come.

In the days that followed the treatment unit, it felt awkward to revert to the "sage on the stage" style of instruction and tell students what they needed to learn and when they needed to learn it. Guided notes on protein synthesis felt forced and alien since the dynamic learning environment had been stripped away to reveal a quieter lecture period. It took 4 school days for the regular biology classes to request that we move away from DO-NOWs and back to What Have We Figured Out So Far discussions, which prompted me to change my practice and incorporate more collaborative sense-making in small groups. The honors biology section still preferred guided notes and lecture-based teaching, however I would like to adopt the collaborative aspects of a storyline curriculum to match the phenomena-driven lessons for my regular biology students. Carrying over these practices, in addition to the continued use of CERs, will hopefully shift the perspective away from *science as a body of knowledge* to *what can we figure out using science*?

Although the treatment test scores did not show a significant improvement compared to the non-treatment, the qualitative data is clear regarding how a storyline can positively impact student learning. Fostering and environment that emphasizes sense-making and an overall comfort level of being able to ask questions and investigate phenomena has moved the needle in terms of how I view science education. The result is no longer whether I cover content within a given timeframe, rather it is about the students' ability to make sense of what they're learning and apply what they've figured out in multiple ways. This paradigm shift coincides with a societal need for science education reform, where students need to engage in the three dimensions of the NGSS as opposed to relying on direct instruction, memorizing facts, and listening to the teacher tell them what they need to know. Storylines are not the only way of achieving that goal, however they are a unique way to usher in this new age of science education. The more students find value in this approach, the more meaningful they find science in their everyday lives.

REFERENCES CITED

About the Working Group. (2022). Illinois Science Teaching Association - Storylining working group. (n.d.). https://ilscience.org/Storyline

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, 116*(39), 19251-19257.

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*(2), 157-175.

Lipsitz, K., Cisterna, D., & Hanuscin, D. (2017). What's the story? *Science and Children, 55*(4), 76-80.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

NGSS Lead States. (2013). *Next generation science standards: For states, by states.* The National Academy Press.

NGSS Lead States. (2013a). APPENDIX A—Conceptual shifts in the Next Generation Science Standards. *Next Generation Science Standards: For States, by States*.

NGSS Lead States. (2013d). APPENDIX D—All standards, all students: Making the Next Generation Science Standards accessible to all students. *Next Generation Science Standards: For States, by States*.

NGSS Lead States. (2013 f). APPENDIX F—Science and Engineering Practices in the NGSS. *Next Generation Science Standards: For States, by States*.

Reiser, B. J., Fumagalli, M., Novak, M., & Shelton, T. (2016). Using storylines to design or adapt curriculum and instruction to make it three-dimensional. Paper presented at the NSTA National Conference On Science Education, Nashville, TN. https://www.academia.edu/24083676/Using_Storylines_to_Design_or_Adapt_Curriculum_and_Instruction_to_Make_It_Three-Dimensional

Reiser, B. J., Novak, M., & McGill, T. A. (2017). Coherence from the students' perspective: Why the vision of the framework for K-12 science requires more than simply "combining" three dimensions of science learning. In *Board on Science Education Workshop "Instructional Materials for the Next Generation Science Standards*.

Reiser, B. J., Novak, M., & McGill, T. A. & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education, 32*(7), 805-829.

Shwartz, Y., Weizman, A., Fortus, D., Krajcik, J., & Reiser, B. (2008). The IQWST experience: Using coherence as a design principle for a middle school science curriculum. *The Elementary School Journal*, *109*(2), 199-219.

Sikorski, T. R., & Hammer, D. (2017). Looking for coherence in science curriculum. *Science Education*, *101*(6), 929-943.

US News & World Report. (2022) *Williams Bay High in Williams Bay, WI - US News Best high* ... US News Best High Schools. https://www.usnews.com/education/best-high-schools/wisconsin/districts/williams-bay-school-district/williams-bay-high-21983

Yager, R. E. (1991). The constructivist learning model. *The Science Teacher*, *58*(6), 52.

Zivic, A., Smith, J. F., Reiser, B., Edwards, K., Novak, M., & McGill, T. (2018). Negotiating epistemic agency and target learning goals: Supporting coherence from the students' perspective. International Society of the Learning Sciences, Inc.[ISLS].

APPENDICES

APPENDIX A:

IRB EXEMPTION

IRB Submission for MSSE 509 (Capstone Requirement)

Beiswanger, Kelly <kelly.beiswanger@montana.edu> Tue, Oct 26, 2021 at 9:45 AM
To: Shane Cullian <shanecullian@gmail.com>
Cc: "Woolbaugh, Walter" <walter.woolbaugh@montana.edu>

Hello Shane,

Thank you for your application. This email acknowledges receipt of the request for IRB Review and serves as the Approval Letter for your research. Your new **IRB Exempt Protocol # is SC102621-EX**.

Study Title: **The Effect of a Storyline Approach to Teaching Biology on Student Achievement, Student Perception, and Teacher Attitude**

As the PI, it is your responsibility to facilitate subject understanding by informing subjects of all aspects of the project, providing an opportunity to ask questions, and describing risks and benefits of participation. Submit any new changes to the research protocol to the IRB via Amendment Form prior to implementing.

The research described in your submission is exempt from the requirement of additional review by the Institutional Review Board in accordance with 45 CFR 690.104(d). The specific paragraph which applies to your research is:

> (1) Research, conducted in established or commonly accepted educational settings, that specifically involves normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction. This includes most research on regular and special education instructional strategies, and research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Thank you,

Kelly Beiswanger

IRB Administrator & Program Manager

Office of Research Compliance

Hamilton Hall 114

Montana State University

kelly.beiswanger@montana.edu

406-994-4706

https://www.montana.edu/orc/irb

APPENDIX B:

NON-TREATMENT PRE-TEST/POST-TEST

Food Webs, Chains and Energy Pyramids Pre-Test/ Post-Test

Please read and acknowledge the following academic integrity statement,

"For this assessment, I have not received assistance from, worked with, or given assistance to another student taking this assessment. My work will not be plagiarized and all responses will be my own. I understand my school's policy on academic honesty and will abide by it during this and future assessments."

☐ Please check the box to the left when you are done reading.

Overview

Your task is to identify producers and consumers in the savanna ecosystem of Gorongosa National Park. Using trophic relationships below, you will create a food chain to show the flow of energy in that system. This food chain will help you construct a pyramid of energy that mathematically represents the flow of energy within that system. Lastly, you will construct a more complex model by depicting multiple relationships in a food web and make claims about the three models you've created. Follow the instructions provided in Parts 1-3.

Trophic Relationships in Gorongosa National Park:

a. Leaves from **small trees/shrubs** are eaten by **elephants, kudu (antelope), hare (rabbits),** and **impala (antelope).**
b. **Tall grasses** are eaten by **water buffalo** and **zebras.**
c. **Short grasses, roots and berries** are eaten by **warthogs, impala, insects, shrews, wildebeest,** and **hare.**
d. **Tall trees** are knocked down by elephants for food.
e. **Insects** are eaten by shrews and warthogs**.**
f. **Elephants** have no natural predators.
g. Water buffalo, kudu, impala, zebra, warthogs and wildebeest are eaten by **lions.** Lions have no natural predators.
h. Hare and shrews are eaten by **serval (wild cat)** and **African Fish Eagles.**
i. **Vultures** eat carrion (dead animals) and have no natural predators.

Part 1: Creating a Food Chain

Using the trophic relationships above, ***label*** the organisms as either "producers" or "consumers". Each of the labeled names are counted as one species of organism (20 total).

Q1: Circle the number of producers you have.

A. 2 B. 4 C. 6 D. 8

Q2: Circle the number of consumers you have.

A. 12 B. 14 C. 16 D. 18

A **food chain** is a model that identifies the feeding relationships and the flow of energy in an ecosystem. Select a producer and a consumer, then use them to help you answer the next question.

Q3: ***Circle*** the model below (A or B) that correctly shows the flow of energy.

A. ________________ → ________________
Consumer Producer

B. ________________ → ________________
Producer Consumer

Q4: ***Justify*** why you chose A or B as the correct model.

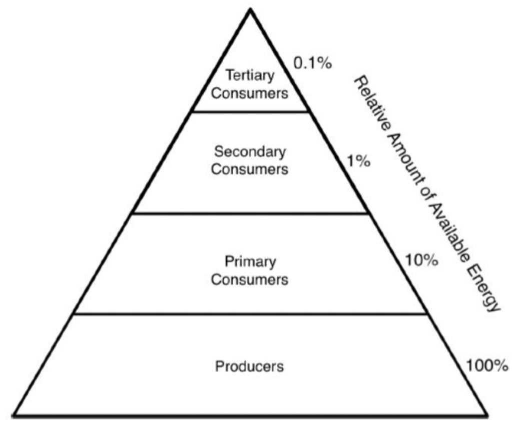
Q5: ***Select*** four organisms to create a food chain, starting with a producer. ***Label*** the **trophic level** of each organism in your food chain as follows:

- producer
- primary consumer
- secondary consumer
- tertiary consumer.

Draw your food chain below using species names you chose. Connect each species with arrows.

Part 2: Available Energy at Each Trophic Level

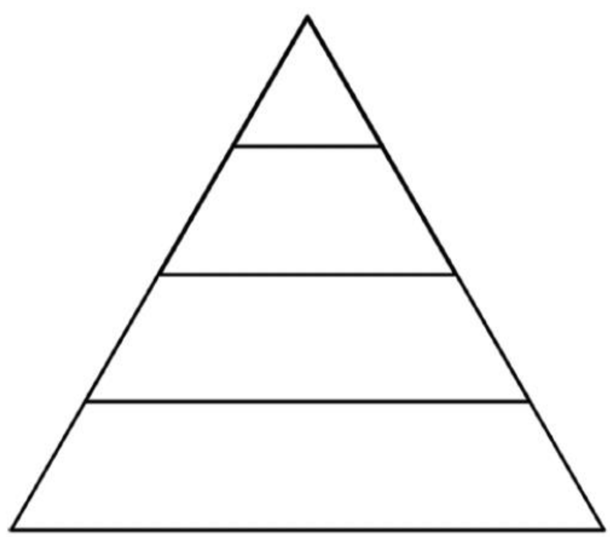
Figure 1. Only a small fraction of energy available at any trophic level is transferred to the next trophic level. The overwhelming majority of the energy is needed by organisms at each trophic level for living, growing, and reproducing. This relationship is shown in the energy pyramid to the right. It suggests that for any food chain, the primary producer trophic level has the most energy and the top trophic level has the least.



Q6: Based on your observations of the energy pyramid above, what is the amount of energy that is typically transferred from trophic level to trophic level? ***Circle*** your answer below.

a. 1% b. 5% c. 10% d. 20%

Q7: ***Place*** the organisms from your original food chain in the blank pyramid below. Using the percentage you circled for **Q6**, ***calculate and record*** the amount of energy available at each level if your producer level had 3,000,000 kilocalories of energy/area.



Q8: In one or two sentences, ***describe*** how the available energy may affect the population sizes of organisms at different trophic levels.

Part 3: Creating a Food Web

Food chains are simple models that show only a single set of energy-transfer relationships, but most organisms obtain energy from many different sources and, in turn, may provide energy to several different consumers. A **food web** illustrates all these interactions and is a more accurate model of how energy moves through an ecological community.

Q9: Starting with your original food chain (**Q5**), add another plant and four more animals to construct a food web that shows how energy flows from producers through primary consumers, secondary consumers, tertiary consumers, and possibly a quaternary consumer. You can have more than one arrow connecting each organism. ***Draw*** a version of your food web below.

Q10: In two to three sentences, ***describe*** any patterns you notice in the relationships between trophic levels.

Complete the CER below using your food chain (**Q5**), web (**Q9**), and energy pyramid (**Q7**) as evidence.

The Driving Question: ***How do energy and matter move differently through an ecosystem?***	
Your Evidence: *(from data and observations)*	**Science Concepts:** *(Circle all that apply)*
	Patterns **Cause & Effect** **Scale, Proportion, & Quantity** **Systems & System Models** **Energy & Matter** **Structure & Function** **Stability & Change**
Your Claim: *(Your claim should answer the driving question)*	
Your Reasoning: *(How do your evidence and the science concepts support your claim?)*	

APPENDIX C:

TREATMENT PRE-TEST/POST-TEST

Treatment Unit Pre/Post-Test

The respondent's email (**null**) was recorded on submission of this form.
* Required

1. Email *

Acknowledgement

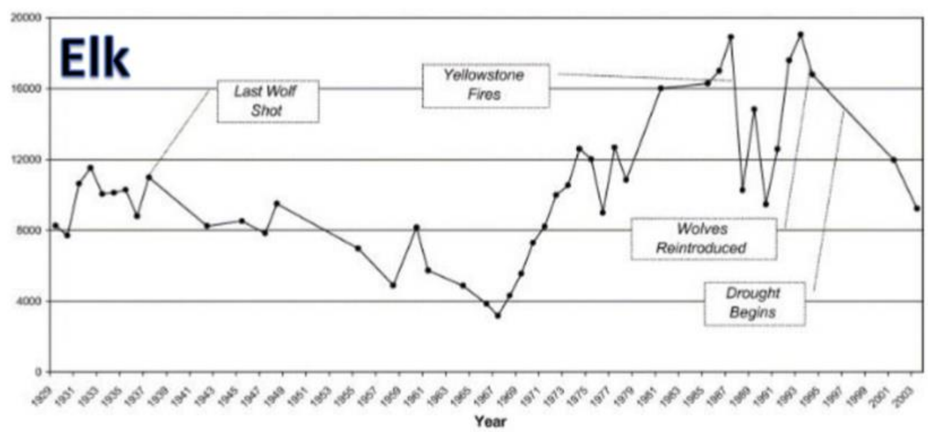
2. Please read and acknowledge the following academic integrity statement, "For this assessment I have not received assistance from, worked with, or given assistance to another student taking this assessment. My work will not be plagiarized and all responses will be my own. I understand my school's policy on academic honesty and will abide by it during this and future assessments." Please check the box below when you are done reading.

Check all that apply.

☐ I have read and acknowledged the academic integrity statement.

Part I - Carrying Capacity

Although wolf packs once roamed from the Arctic tundra to Mexico hunting elk and other prey species, loss of habitat and extermination programs by humans led to their demise throughout mc of the United States by early in the 1900s. In 1973, the US Fish and Wildlife Service listed the northe Rocky Mountain wolf (Canis lupus) as an endangered species and designated greater Yellowstone National Park as one of three recovery areas. From 1995 to 1997, 41 wild wolves from Canada and northwest Montana were released in Yellowstone National Park. As expected, wolves from the growing population dispersed to establish territories outside the park where they are less protecte from human-caused mortalities. The park helps ensure the species' long-term viability in Greater Yellowstone and has provided a place for research on how wolves may affect many aspects of the ecosystem. The graph below is the population census data for the elk in Yellowstone National Park



3. Q1: After 1967, the elk population began to surge. Approximately HOW LONG (years) after the extermination of wolves did it take to reach this point? Your answer should be estimated to within ± 1 year. * 1 p

4. Q2: Assume the carrying capacity of elk in this ecosystem is just below 20,000 individuals. Based on the graph above, which of the following is a BIOTIC factor that is impacting the elk population? * 1 p

Mark only one oval.

- Drought
- Wolves

5. Q3: Explain your choice for Q2 by DESCRIBING THE EFFECT it has on the elk population. * 1 p

6. Q4: Assume the carrying capacity of elk in this ecosystem is just below 20,000 individuals. Based on the graph above, which of the following is an ABIOTIC factor that is impacting the elk population? * 1 p

Mark only one oval.

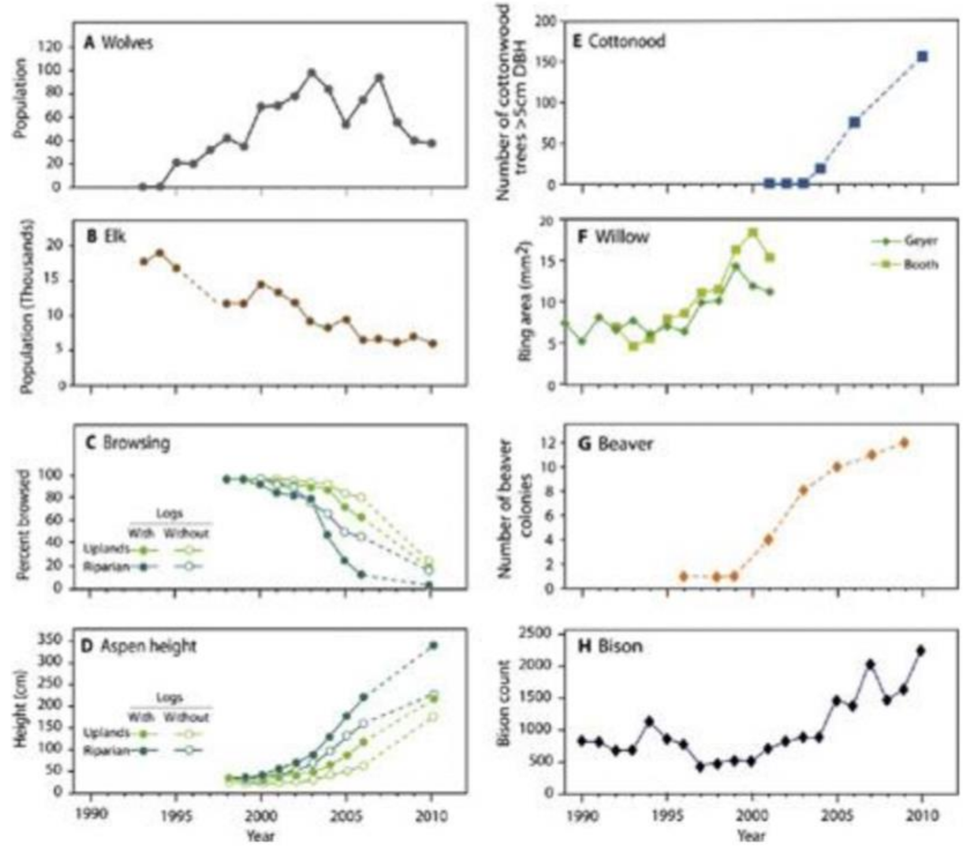
- Drought
- Wolves

7. Q5: Explain your choice for Q4 by DESCRIBING THE EFFECT it has on the elk population. * 1 p

8. Q6: CLAIM - Based on the elk population data from approximately 1967 to 1987, what effect do you think the elk population had on the populations of producers? Write your COMPLETE claim that answers this scientific question below. * 1 p

Part II - Ecosystem Homeostasis

The graphs that follow represent the species and behaviors in the wolf ecosystem in Yellowstone National Park. Three tree species (aspen, cottonwood, and willow) are included so you can identify patterns and cause and effect relationships with other species.



9. Q7: What do we call species like wolves who have impact on so many other species, whether the effects are direct or indirect? * 1 p

Mark only one oval.

- Impact species
- Indicator species
- Keystone species
- Relevant species

10. Q8: How do the data points from 2005 specifically illustrate the relationship between elk and wolves? EXPLAIN using quantitative data. * 1 p

____

____

____

____

____

11. Q9: CLAIM - Of the prey species included in the graphs, the wolves preferred prey item is _________ . * 1 p

Mark only one oval.

- Beaver
- Bison
- Elk
- Willow

12. Q10: EVIDENCE - Cite evidence from the graphs to support the claim made in Q9. * 1 p

13. Q11: Browsing is a foraging behavior where animals consume leaves and twigs from trees. What is the relationship between the elk and browsing data provided? * 1 p

Mark only one oval.

- Elk are browsers and eat leaves from tree branches
- Elk are grazers and eat primarily grasses
- Elk are hunters and consume live prey

14. Q12: CLAIM - The INCREASED presence of browsers like elk in an ecosystem impacts the growth of the three (3) tree species by _______________ . * 1 p

Mark only one oval.

- decreasing the growth of the three tree species'.
- increasing the growth of the three tree species'.

15. Q13: Cite evidence from the graphs to support your claim in Q12. * 1 p

16. Q14: Using quantitative data from the graphs as evidence, EXPLAIN the effect the reintroduction of wolves has on the carrying capacity of the ELK, BISON, and BEAVER populations. * 3 pc

17. Q15: EXPLAIN how the data at the beginning of Part II supports or refutes the claim you created at the end of Part I. (You might have to click back to Part I to review your answer to Q6). * 1 p

Scientists used molecular genetic techniques to analyze DNA sequences from 150 wolves in Yellowstone National Park, which covers parts of Wyoming, Montana and Idaho. They found that a novel mutated gene in dogs, known as the K locus, is responsible for black coat color and was transferred to wolves through mating with domesticated dogs. The inherited black-coat allele results in an enhanced immunity against a highly contagious viral disease called "distemper". The table below lists the numbers of each wolf color phenotype in the national park during different tir periods.

Time Period	Gray Population Numbers	Black Population Numbers	White Population Numbers	Total Population Numbers	% Gray	% Black	% White
1958	36	44	0				
1969	31	26	0				
1977	59	70	3				
1988	39	115	3				
1997	42	101	1				
2005	35	97	0				
2016	12	56	0				

18. Q16: Which of the following represents the total population numbers of wolves in 1997, 2005 and 2016? * 1 p

Mark only one oval.

- 111, 122, 88
- 120, 188, 92
- 144, 132, 68

19. Q17: Which of the following represents the % gray, black and white phenotypes of wolves in 1997? * 1 p

Mark only one oval.

- 15%, 47%, 38%
- 29%, 70%, ~1%
- 35%, 64%, ~1%

20. Q18: Which of the following represents the % gray, black and white phenotypes of wolves in 2016? * 1 p

Mark only one oval.

- 15%, 47%, 38%
- 17.6%, 82.4%, 0%
- 22%%, 68%, 10%

21. Q19: How did the wolf population change over time? DESCRIBE the trend for each phenotype in the population using evidence from the data. * 3 pc

APPENDIX D:

FOCUS GROUP INTERVIEW QUESTIONS

Focus Group Interview Questions:

1. How has school been going this year? What classes do you really enjoy?
    a. Has learning about biology been fun this year? (N=10)
        i. If yes, why?
        ii. If no, why not?
2. Speaking specifically to biology, what has been your most memorable experience in class so far?
    a. What specifically made that experience/lesson/activity memorable? (N=10)
    b. Did you have similar experiences in your past science classes? (N=10)
3. Has this past unit in biology been different from your other science classes?
    a. If not, how has it been like your previous science classes? (N=3)
    b. If it has, can you give an example of this difference. (N=9)
4. How important is it for you to have a say in a science classroom? E.g. being able to ask questions and have those questions drive the next lesson(s).
    a. Do you feel like you're able to ask questions freely and make claims without judgement? (N=10)
    b. Do you like being able to ask questions freely and figure things out with your group? (N=10)
    c. Do you feel like everyone in class feels the same way as you do about your comfort level speaking up? (N=10)
5. If you had your choice, would you rather be given all the information that is covered in a unit up front or would you prefer to discover new information and figure things out as you progress throughout a unit?
    a. Why so? (N=10)
    b. What are the pros and cons to that approach? (N=8)
    c. Do you feel that everyone would prefer to learn in that fashion? (N=10)
6. How do you think this past unit's style of teaching and learning will prepare you for other science classes, or science in the real world?
    a. How do you think the concepts and practices of science will help you in the future?

APPENDIX E:

SURVEY 1

Survey I (Unit Comparison Likert & Non-Likert Items)

Participation in this research survey is voluntary. Participation or non-participation will not affect a student's grade or class standing in any way. This survey will take 10-15 minutes to complete.

The respondent's email (**null**) was recorded on submission of this form.

1. Email *

Part I:

Directions:

Using the scale below, indicate the extent to which you agree with the following statements.

1 = Strongly Disagree, 2 = Disagree, 3 = Neutral/ Undecided, 4 = Agree, 5 = Strongly Agree

2. I enjoy my biology class.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

3. Please explain your reasoning behind your choice for the question above:

4. Biology is a challenging class for me.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

5. Please explain your reasoning behind your choice for the question above:

6. Discussing my ideas within a small group helps me learn.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

7. Please explain your reasoning behind your choice for the question above:

8. Having the teacher directly tell me what I need to know helps me learn.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

9. I like biology class when I have to figure out an answer through problem-solving.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

10. I like biology class when I am told exactly how to solve a problem.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

11. Please explain your reasoning behind your choice for the question above:

____

____

____

____

____

12. The lessons and activities I do in class are in an order that makes sense to me.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

13. Please explain your reasoning behind your choice for the question above:

____

____

____

____

____

14. I am sometimes confused about what is going on in biology.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

15. My teacher encourages sense-making and "figuring out" discussions within my group.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

16. I prefer when my teacher tells us why the things we are learning about are important.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

17. Please explain your reasoning behind your choice to the question above.

18. I always pay attention in biology.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

19. I feel confident in my abilities to conduct scientific practices when I am learning in biology.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

20. I enjoy explaining what we are learning to people in our group.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

21. The activities, discussions and questions we investigate help me prepare for summative assessments.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

22. Please explain your reasoning behind your choice for the question above.

Part II

Directions: Using the scale below, indicate the usefulness to which you found the following instructional methods.

1 = Extremely Not Useful, 3 = Neutral/ Undecided, 5 = Very Useful

23. I think that Driving Question Jamboards are

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

24. I think that What Have We Figured Out So Far Discussions are

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

25. I think that DO-NOW's are

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

26. I think that being provided a unit sheet in advance is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

27. I think that knowing what the objectives of the unit are before we start learning is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful						Very Useful

28. I think that figuring out information and ideas day-by-day is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful						Very Useful

29. I think that taking guided notes are

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful						Very Useful

30. I think that working in small groups is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful						Very Useful

31. I think working individually is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

32. I think that learning about phenomena (e.g. otters and ecosystem homeostasis) is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

33. I think that learning about separate/ individual topics (food chains & energy pyramids) is

Mark only one oval.

	1	2	3	4	5	
Extremely Not Useful	◯	◯	◯	◯	◯	Very Useful

Part III

Student interview interest.

34. Would you be interested in participating in an audio-recorded interview regarding storylining and traditional instruction in biology via Google Meet?

Mark only one oval.

◯ Yes

◯ No

APPENDIX F:

SURVEY 2

Survey II (Then-Post Treatment Attitudes and Self-Efficacy Scale)

Participation in this research survey is voluntary. Participation or non-participation will not affect a student's grade or class standing in any way. This survey will take 10-15 minutes to complete. Adapted from the Illinois Science Teachers Association's NGSS Then-Post Conference Survey.

The respondent's email (**null**) was recorded on submission of this form.

1. Email *

Previous Experiences

2. How familiar with learning science via a storyline were you BEFORE taking this class?

Mark only one oval.

	1	2	3	4	5	
Very unfamiliar						Very familiar

3. How familiar with learning science via a storyline are you NOW after experiencing it in this class?

Mark only one oval.

	1	2	3	4	5	
Very unfamiliar						Very familiar

Science Practices Self-Efficacy Scale

4. As a result of participating in this storyline, please rate your confident are you in your ability to complete these tasks.

Mark only one oval per row.

	0 (Not Confident)	1 (No More Confident)	2 (Somewhat More Confident)	3 (Much more confident)	N/A
Ask questions and define problems					
Develop and use models					
Plan and carry out investigations					
Analyze and interpret data					
Use mathematics and computational thinking					
Construct explanations and design solutions					
Engage in argument from evidence					
Obtain, evaluate, and communicate information					

Agreement/Disagreement

5. Now, after experiencing biology through a storyline, I believe this method of learning is an improvement to my science education.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

6. Before experiencing biology through a storyline, I believe this method of learning is an improvement to my science education.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

7. Now, after experiencing biology through a storyline, I prefer this style of teaching and learning science over traditional methods.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

8. Before experiencing biology through a storyline, I preferred this style of teaching and learning science over traditional methods.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

9. Now, after experiencing biology through a storyline, I have a deep understanding of scientific concepts, practices and core ideas.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

10. Before experiencing biology through a storyline, I have a deep understanding of scientific concepts, practices and core ideas.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

11. Now, after experiencing biology through a storyline, I feel passionate about science.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

12. Before experiencing biology through a storyline, I felt passionate about science.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

13. Now, after experiencing biology through a storyline, biology is easier to understand.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

14. Before experiencing biology through a storyline, biology was easier to understand.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

15. Now, after experiencing biology through a storyline, I am more wiling to work with others and make sense of phenomena.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

16. Before experiencing biology through a storyline, I was more wiling to work with others and make sense of phenomena.

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

Attitude & Impression

17. Now, after experiencing biology through a storyline, my attitude towards science is

Mark only one oval.

	1	2	3	4	5	
Very Negative	⬭	⬭	⬭	⬭	⬭	Very Positive

18. Before experiencing biology through a storyline, my attitude towards science was

Mark only one oval.

	1	2	3	4	5	
Very Negative	⬭	⬭	⬭	⬭	⬭	Very Positive

19. Please rate the degree to which the following factors act as barriers to your ability to learn science effectively:

Mark only one oval per row.

	Not a barrier	A minor barrier	A major barrier	A repressive barrier
My science knowledge	⬭	⬭	⬭	⬭
My feelings about science	⬭	⬭	⬭	⬭
My science teacher(s)	⬭	⬭	⬭	⬭
Available time for learning science	⬭	⬭	⬭	⬭
Availability of resources and materials	⬭	⬭	⬭	⬭