

USING PRACTICE PERFORMANCE ASSESSMENTS TO  
PREPARE STUDENTS FOR SUMMATIVE PERFORMANCE  
ASSESSMENTS IN GRADE SIX SCIENCE

by  
Neil Trevor Freeman

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

of  
Masters of Science  
in  
Science Education

MONTANA STATE UNIVERSITY  
Bozeman, Montana

July 2020

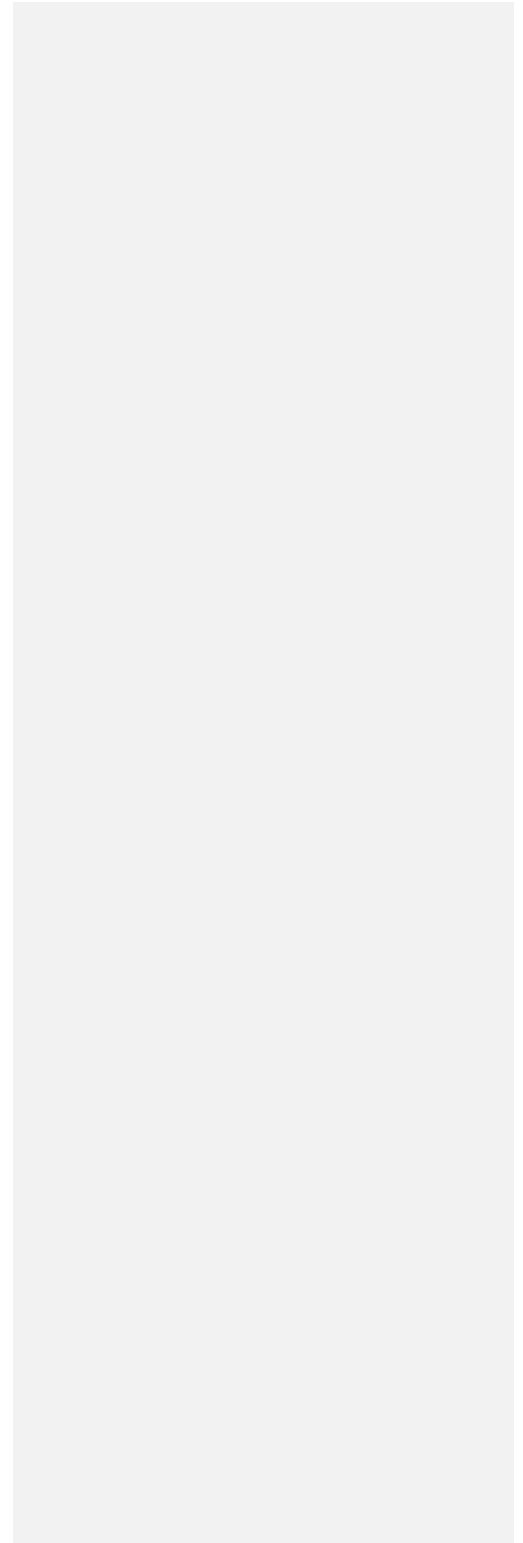
©COPYRIGHT

by

Neil Trevor Freeman

2020

All Rights Reserved

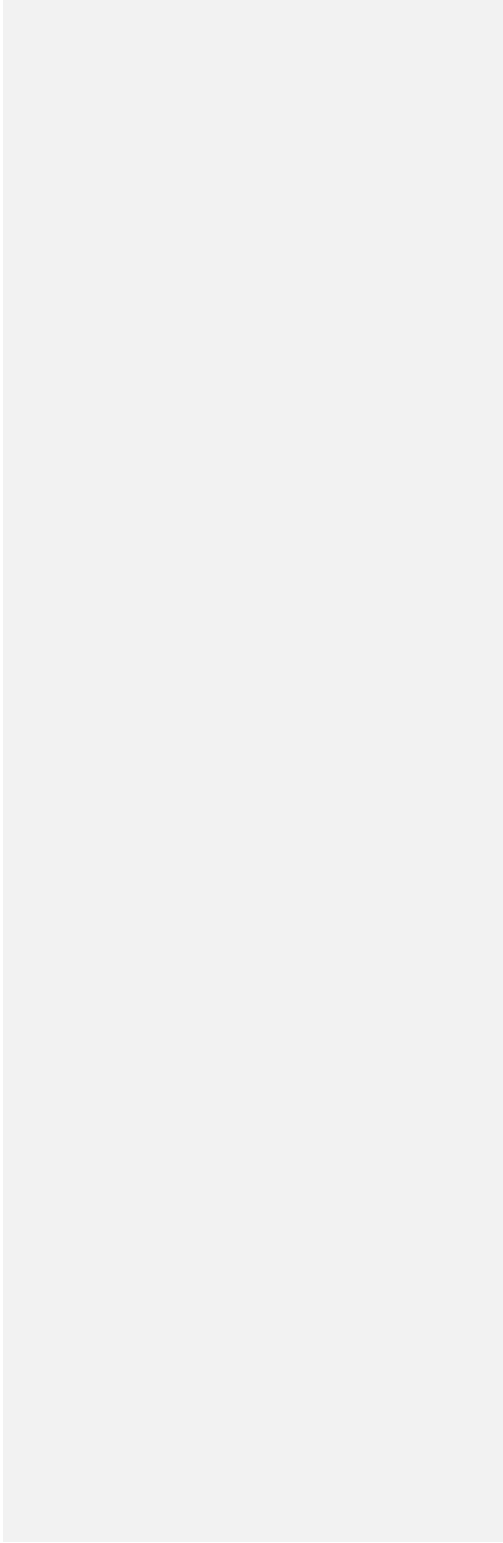


## TABLE OF CONTENTS

1. INTRODUCTION AND BACKGROUND .....	1
Support Team .....	3
Ms. Pamela Jane von Bondungen.....	3
Dr. Erin Leininger .....	3
Ms. Joanna McElhinney .....	3
2. CONCEPTUAL FRAMEWORK.....	3
Connection Between Conceptual Framework and Methodology.....	9
3. METHODOLOGY .....	9
Demographics.....	9
Treatment and Intervention .....	10
Time Frame .....	12
4. DATA AND ANALYSIS.....	14
Student Interviews and Surveys .....	21
5. INTERPRETATION AND CONCLUSION.....	29
Implications on Teaching Practices.....	34
6. VALUE.....	35
REFERENCES CITED .....	38
APPENDICES .....	41
APPENDIX A: Institutional Review Board Exemption .....	42
APPENDIX B: NGSS Performance Expectations for the Assessments .....	44
APPENDIX C: Cells Investigation Practice Performance Assessment .....	47
APPENDIX D: Cells Investigation Summative Performance Assessment.....	50
APPENDIX E: Cell Analogy Practice Performance Assessment .....	53
APPENDIX F: Cell Analogy Summative Performance Assessment .....	58
APPENDIX G: Body Systems Practice Performance Assessment .....	63
APPENDIX H: Body Systems Summative Performance Assessment.....	66
APPENDIX I: Matter Practice Performance Assessment .....	69
APPENDIX J: Matter Summative Performance Assessment .....	72
APPENDIX K: Cells Investigation Google Survey Questions .....	75
APPENDIX L: Cell Analogy Google Survey Questions .....	81
APPENDIX M: Body Systems Google Survey Questions.....	85

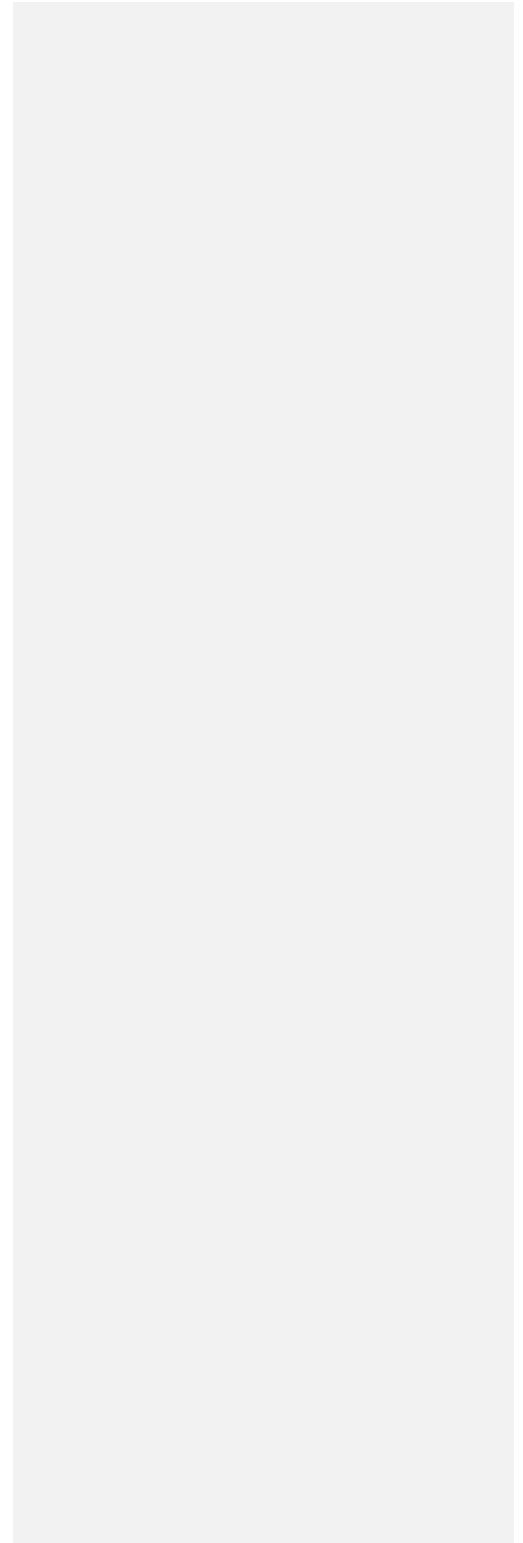
TABLE OF CONTENTS CONTINUED

APPENDIX N: Matter Google Survey Questions.....	90
APPENDIX O: Interview Questions .....	95
APPENDIX P: Strategies for the Final Review Class.....	97



LIST OF TABLES

1. Research Matrix.....	13
2. PPA Versus SPA Scores.....	15



LIST OF FIGURES

1. Cell Analogy PPA and SPA Score Distribution.....	16
2. Cells Investigation PPA and SPA Score Distribution .....	17
3. Body Systems PPA and SPA Score Distribution .....	18
4. Matter PPA and SPA Score Distribution.....	19
5. Post-SPA survey confidence results .....	21
6. Post-SPA survey about students' value of the PPA .....	23
7. Post-SPA survey about students' value of the Final Review Class.....	24
8. Post-SPA survey about students' value of the PPAs being graded but not counting towards their overall grades.....	26
9. Post-SPA survey about students' value of the treatment.....	28

## ABSTRACT

Summative performance assessments (SPAs) are part of most science courses and classes from Elementary School up through University. This study aims to decrease student stress and improve student performance on SPAs by using Practice Performance Assessments (PPAs) followed by Final Review Classes (FRCs) in order to prepare sixth grade students for SPAs. In the pre-treatment phase, each part of a unit was taught to students. The treatment then began with a PPA for that part of the unit which aligned with a Next Generation Science Standards (NGSS) Performance Expectation (PE). The PPAs were graded (although not counting towards students' overall grades) and returned to students during the next class (the FRC) in order to prepare students for the SPA. The class after the FRC, students were given the SPA. Student opinions on each part of the unit, PPA, FRC, and SPA were determined with Google surveys, interviews, and a member feedback session. The results indicated that the treatment (graded PPAs followed by FRC followed by SPA) was valued by students and was effective in improving student performance on SPAs, which aligns with the literature's testing effect.

## INTRODUCTION AND BACKGROUND

For over a century, teachers all round the world have been giving students summative assessments with varying levels of success. Students experience different levels of preparedness, success, and stress for these assessments, which are a fundamental part of education. At the International School Manila (Philippines), summative assessments are less demanding in our Elementary School, which runs up to Grade Four. Grade Five is students' first year of Middle School, and that's when expectations for summative assessments increases. There is another big jump in summative assessment difficulty from Grade Five to Six, and my Action Research was about preparing my sixth-grade students better for summative assessments so that they:

- Feel prepared for assessments so they:
- Are not feeling stressed by them and so they:
- Experience success on them as evidenced by good grades.

Most teachers in our Middle and High School have similar goals, although we all have different methods to prepare students for summative assessments. My Action Research was about using Practice Performance Assessments (PPAs) to prepare students for Summative Performance Assessments (SPAs). Upon completion of my Action Research, I shared my findings with fellow faculty in what our school calls a "Learning Byte". I also shared my findings with my students to develop them metacognitively in the domain of assessment preparation.

To summarize, the problem in my Grade Six science classes is:

- Some students struggle to do well on summative performance assessments (SPAs).

Commented [JAnd1]: I would suggest putting these in the paragraph like this "they: (a) feel prepared..., (b) are not... (c) experience success... Bullets are not generally used in APA



Thus, my primary Action Research (AR) question is:

- How can I use practice performance assessments (PPAs) to prepare students for summative performance assessments (SPAs)?

My secondary AR questions are:

- How does assigning a grade to students PPAs impact their subsequent performance on SPAs?
- How can I use students' results on PPAs to prepare them for SPAs?
- What are students' opinions of PPAs and final review classes (FRCs)?
- How did using PPAs impact me as a teacher?

My Action Research was about preparing students for the increased challenge of Grade Six science performance assessments, some of which bear more similarity to tests while others are more like lab reports. In the past, some of my students did not do very well on non-test type summative assessments, which ultimately reflected on my teaching. At the time, I gave my students practice tests for test type summative assessments, but I wasn't doing the same thing for non-test type summative assessments. Over time, I realized that all summative assessments should be preceded by a practice assessment that prepares students for the summative assessment. During the school year 2018-2019, I learned through student interviews and my literature review that practice assessments should be similar in nature, content, format, and setting to the summative assessment.

Throughout the 2018-2019 school year, I piloted my AR. The students showed improvement from their PPA scores to their SPA scores. Furthermore, the surveys and interviews that I conducted showed that students highly appreciated both the PPAs and the final review classes (FRCs). Thus, I decided that PPAs and FRCs were valuable

Commented [JAnd2]: Update this section to fit the format required for focus questions. See formatting document for what this section should look like.

teaching practices which merited formal AR investigation and documentation during the 2019-2020 school year.

Ultimately, I want my students to feel prepared for summative assessments, and to subsequently do well on them. Most students and teachers want the same, which is why I felt that my topic was worthy of Action Research.

#### Support Team

To support me with this endeavor, these people were part of my support team:

Ms. Pamela Jane (Polly) von Bodungen, Grade 6 science teacher

Polly teaches the other four classes of Grade 6 science at our school. She was selected because I had to consult with her when creating and grading the PPAs and SPAs as our school requires that we give the same SPA to all our students.

Commented [KAT3]: I thought that this helped to clarify why she was selected.

Dr. Erin Leininger, Middle School Curriculum Coordinator

Erin is a respected administrator at our school. She was selected because of her curricular expertise, her experience writing her doctorate, and her love of research. Due to her experience, I used her to help with my survey/interview questions. She was also selected for her propensity to give me honest feedback.

Ms. Joanna McElhinney, English and Social Studies Teacher

My wife proofread many drafts of this paper. She caught mistakes and made astute comments that resulted in further editing.

#### CONCEPTUAL FRAMEWORK

As early as the beginning of the 19th century, teachers have been testing students to check for mastery of what has been taught (U.S. Department of Education, 2008). A

century later, standardized testing was introduced because industry, the military, and the government had become interested in the effectiveness of education (Giordano, 2005). If students did not meet a given benchmark, they may have been held back or retained.

These high stakes standardized tests are still a big part of education today. The assessments in my sixth-grade science classes are not high stakes. However, over 95% of my students will take one or two IB sciences in Grades 11 and 12, which means they will take high stakes exams and performance assessments in six years' time. Getting students to learn concepts, improve their writing, data analysis, and test taking skills were a goal of my teaching and Action Research. This was measured by the students' improvement from the PPAs to the SPAs.

While tests are a significant part of education, students are given a multitude of other formative and summative assessments depending on the subject area. There have been lots of criticisms directed at traditional multiple-choice achievement tests in science (Shavelson, Carey & Webb, 1990). They are limited in assessing students conceptual understanding, problem solving, and science specific skills. (Shavelson & Ruiz-Primo, 1995). These theoretical articles supported the nature of the practice and summative assessments that I gave my sixth-grade science students as mine did not include multiple choice and were focused on higher-order thinking.

Science performance assessments are designed to be two-dimensional (content, skills) or three-dimensional (content, skills, concepts). Some may resemble tests, while others are more performance based. In the last decade, there have been a lot of curricular reforms in teaching, learning, and assessment practices (Alt, 2018). The purpose of these reforms has been to encourage students to determine relationships between concepts, to

interpret data and facts (Hesse et al, 2015), and to metacognitively think about their learning (Lasry, Charles & Whittaker, 2016). It is also to “apply their understanding of new concepts in new situations, employ creative thinking, solve problems, develop scientific skills and construct knowledge, and search for possible solutions to ill-structured questions by inquiry-based discussions” (Alt, 2018, p. 388). These articles supported the higher-order nature of my PPAs and SPAs.

Formative assessment in science (and other subjects) should be used by teachers and students to identify gaps in skills and concepts, which should then guide the planning of subsequent lesson(s) in order to prepare students for summative assessments (Panadero, Jonnson & Strijbos, 2017). No matter what type of assessment, there are many ways that teachers prepare students for assessments. Practice assessments have been found to be very successful in preparing students for summative assessments. The testing effect refers to gains in learning and achievement that occurs when students take a practice test that resembles the format, content, style of questions, and test conditions of the summative test (Adesope, Trevisan & Sundararajan, 2017).

Numerous studies have been done that demonstrate the testing effect. They first became common in the 1960s. In a Stanford University study, college students were given five or ten paired presentation trials (Allen, Mahler & Estes, 1969). Some students were tested on the pairings, while others were not. The long-term retention of those who had been tested was significantly better than those that were not.

A Washington University (in St. Louis) study found that students who studied more, without taking a practice test retained more information initially, but less information in a week, compared to students who studied less, but took a practice test in

place of the missed studying (Roediger & Karpicke, 2006a). The researchers believe that taking a memory test improves long-term memory retention even if it results in less initial retention due to less studying.

A University of California in San Diego study found that testing improved students visuospatial learning of maps more than conventional methods of studying (Carpenter & Pashler, 2007). The fifty students in the treatment group took online tests in place of conventional methods of studying (control group). The final test had the students drawing the maps which they had studied and/or been tested on. The final maps were assessed in four ways, and in all four ways, the treatment group (study/test) did better than the control group (study/study). All these studies provided strong supporting evidence for me having given my students practice assessments for my Action Research.

Transfer-appropriate processing is the idea that performance on a given task will be highest if the characteristics of the learning procedure are like those of the testing procedure (Bransford, Franks, Morris & Stein, 1979). This aligns with my methodology in that the structure, time, place, content, practices, question type of my PPAs matched that of my SPAs. In short, this means I did not assign a lab report as the PPA, and then give a multiple-choice test as the SPA. If the PPA was more of a lab report, then so was the subsequent SPA. If the PPA was a CER (Claim, Evidence, Reasoning) with supporting models then so was my SPA.

Although most of the literature supports the testing effect, one study finds that as the complexity of the learning materials increases, the testing effect significantly decreases (van Gog & Sweller, 2015). Others disagree with this article and argue that the study was flawed (Karpicke & Aue, 2015). Ultimately, the effectiveness of giving

students practice assessments prior to a summative assessment was supported by most of the research, thus I incorporated this strategy into my methodology in order to better prepare students for summative assessments. Similarly, most of these studies graded students on their summative assessments, which is what I did for my methodology.

In 2005, a meta-analysis of many testing effect research studies found that feedback on practice tests was beneficial to learning (Phelps, 2012). Phelps defined feedback as grades and/or comments. It should be noted that some studies have found that feedback has a negligible effect although in these studies, feedback and non-feedback students scored quite high on the practice tests, so that could have been why there wasn't a significant difference between their performance on the practice and summative test (Butler & Roediger, 2007). Most of the literature found that feedback (grades and/or comments) had a positive impact on summative assessment scores and student learning. Some of the methodology in the literature analyzed improvement from practice tests to summative tests, which is what I did for my methodology.

Through a meta-analysis of many studies, it was found that the optimal time between practice and summative tests should be between one and six days (Adesope et al., 2017). This aligns with my methodology, as I gave my SPAs two to five days after the PPAs.

Some testing effect research studies also included qualitative analysis in the form of surveys and interviews about the subject's opinions on practice tests (Phelps, 2012). The responses indicated that the subjects found practice tests beneficial. Accordingly, I surveyed and interviewed my students to determine if their opinions were similar.

The majority of the literature argues that practice tests are beneficial to both student learning and summative test performance, which lent credence to the overall ethos of my AR. Quite simply, giving students practice performance assessments (PPAs) should have been beneficial for their learning and should have resulted in improved summative performance assessments (SPAs) scores, which is why I did just that.

Most of the literature found that feedback (grades and/or comments) had a positive impact on summative assessment scores and student learning, which is why I graded all PPAs in all rounds of data collection. Some of the literature analyzed improvement from practice tests to summative tests, which is what I did for my methodology. The literature argued that testing conditions and question type of practice tests should be like those of the summative tests. Accordingly, I kept these factors the same for the PPAs and SPAs.

My assessments were all open notes. It should be noted that none of the studies in this literature review allowed students to use open notes for either practice or summative tests. This was a fairly significant difference from my AR. My PPAs and SPAs are performance-based and involve applications of content, skills, and concepts. They aren't just assessing student knowledge. The real world is open notes with people being able to access mentor texts/videos when producing their own content, like I have done while writing this paper. The real world involves applications and "performance-based" tasks. This is the big reason that my assessments are open notes. The second reason is that a lot of "testing effect" research has already been carried out on more typical content-based, closed notes "tests", thus my research is a novel take on the "testing effect".

### Connection between Conceptual Framework and Methodology

To summarize, these components of my methodology were supported by the literature:

- Doing a PPA before an SPA is supported by the ideas of (Panadero et al., 2017).
- PPAs were similar in nature, question type, and conditions to those of the SPAs, which is supported by the ideas of (Adesope et al., 2017; Bransford et al., 1979).
- SPAs were one to six days after the PPAs, which is also supported by (Adesope et al., 2017).
- PPAs were graded, which is supported by (Phelps, 2012).
- Students were surveyed and interviewed about the PPAs and SPAs, which is supported by the research methods of (Phelps, 2012).

Commented [KAT4]: an SPA?

Commented [KAT5]: period?

## METHODOLOGY

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). My treatment and methodology were done with all four of my sixth-grade science classes, each of which had between 20 and 22 students.

### Demographics

International School Manila (Philippines) is a private non-profit school catering to wealthy expats. The largest nationalities represented are Filipino, Korean, Japanese, Indian, and American. The rest of the students come predominantly from East and Southeast Asia, Europe, Canada, Australia, and New Zealand. The students come from a high socioeconomic background. They are motivated to do well academically and by the



time they reach Grade 11 and 12, over 90% of them will choose to do the rigorous full IB diploma.

#### Treatment and Instrumentation

Below is the treatment which was implemented four times from September to December 2019. Students who were absent for the PPA, final review class (FRC), and/or the SPA were excluded from that round of data collection.

1. Taught the relevant part of the unit.
2. Gave students the PPA which they had one 70-minute class period to complete. Some students came to finish theirs at lunch or after school.
3. Made notes in my journal during the PPA.
4. Graded the PPAs, while making notes in my journal.
5. Photocopied the graded PPAs.
6. Planned and delivered a 70-minute FRC to account for deficiencies in the PPAs.
  - a. This lesson was the class immediately after the PPAs.
  - b. The FRC began with returning the graded PPAs to students. The graded PPAs did not count towards their overall grade.
  - c. Various strategies were employed during this class, but in all instances, students received an exemplar PPA, which they had to use to further analyze their own deficiencies and then correct their PPA.
7. Recorded FRC reflections in my journal.
8. Gave students the SPA in the class immediately after the FRC, which they had one 70-minute class period to complete. Some students came to finish theirs at

Commented [KAT6]: period?

lunch or after school. As stated in the methodology section, SPAs were similar in nature, question type, and conditions to those of the PPAs.

9. Interviewed eight to thirteen students per class after the second, third, and fourth SPAs, for a total of thirty-one interviews. I attempted to select students that were representative of all my sixth graders. I did my best to select students of different academic abilities (13 high, 11 medium, 7 low), gender (16 boys, 15 girls), ethnicities/nationalities (10) and from all four classes (10, 9, 8, and 4). Two students declined to be interviewed. A few students were interviewed while the other students finished the SPA, while others were interviewed after school or during their English class's library time. The interview was about student perceptions regarding the PPA, FRC, and SPA.
10. Interviews were recorded on my phone. They were later transcribed and coded.
11. Google surveyed (mostly Likert scale) all students (except in round one where I missed one class) about their perceptions regarding the PPA, FRC, and SPA. The survey was done in the class immediately after the SPA. In most instances, this was done before students were interviewed.
12. Graded all SPAs, while recording notes in my journal.
13. Photocopied all SPAs.
14. Stored photocopied SPAs in a locked filing cabinet.
15. Analyzed the data from the PPAs, SPAs, Google surveys, and interviews.
16. Shared findings with students after the fourth SPA in a member feedback session. This session was recorded on my phone for later coding and transcription.

Commented [KAT7]: representative?

Commented [KAT8]: Do you want single or double spaces between sentences?

Time Frame

- Each round of data collection for each class was in the fall semester of 2019.
- I see my students 4 times a week, so Class 1 was the PPA, Class 2 was the FRC, Class 3 was the SPA, and Class 4 was the Google survey.
- PPAs were graded and photocopied between Class 1 and Class 2.
- Interviews took place between 0 – 3 days after students completed the SPA.
- SPAs were graded and photocopied 1 – 2 weeks after students completed them.
- The Cells Investigation round of data collection was from September 12<sup>th</sup> – 24<sup>th</sup>.
- The Cell Analogy round of data collection was from October 3<sup>rd</sup> – 15<sup>th</sup>.
- The Body Systems round of data collection was from November 13<sup>th</sup> – 21<sup>st</sup>.
- The Matter round of data collection was from December 5<sup>th</sup> – 16<sup>th</sup>.
- The member feedback sessions were on December 17<sup>th</sup> and 18<sup>th</sup>.
- Cell Analogy PPAs, SPAs, surveys, interviews, and journal were coded and analyzed from October 26<sup>th</sup> – 31<sup>st</sup>.
- Remaining PPA, SPAs, surveys, interviews, journal, and member feedback sessions were coded and analyzed from April 6<sup>th</sup> to May 4<sup>th</sup>, 2020.

See Appendices C, E, G, and I for the PPAs.

See Appendices D, F, H, and J for the SPAs.

See Appendices K, L, M, and N for the survey questions.

See Appendix O for the interview questions.

Table 1  
*Research Matrix*

Research Question	Data Analysis of PPA and SPA scores	Data Analysis of Journal, Member Feedback	Data Analysis of Interviews and Google Likert surveys
How can I use PPAs to prepare students for SPAs?	Analyzed the differences of the SPAs compared to the PPAs. Used normalized gains to determine if the gains were statistically significant.	Journalled the teaching methods used in the FRC. Correlated these with PPA to SPA gains.	Analyzed quantitative and qualitative data from interviews and surveys to see which methods students found most effective in preparing them for the SPAs.
How does assigning a grade to students PPAs impact their subsequent performance on SPAs?	Analyzed the differences of the SPAs compared to the PPAs. Used normalized gains to determine if the gains were significant.		Asked students during the interviews and surveys if they thought grading their PPAs before the SPAs helped with their learning and performance on the SPAs.
How can I use student results on PPAs to prepare them for SPAs?	Noted and coded deficiencies on PPAs and used them to plan the FRC.	Recorded methods used during the FRC in my journal along with reflections.	Asking students during the interviews and surveys which strategies prepared them most effectively for the SPAs.
What are students' opinions of PPAs and FRCs?		Shared findings with all four classes and noted their comments.	Analyzed Google survey and interview responses.
How did using PPAs impact me as a teacher?	Analyzed gains from the PPAs to SPAs. Higher gains provided me with personal satisfaction thus validating my use of PPAs and FRCs.	Analyzed reflections from teacher journal.	Analyzed student responses on the interviews and Google survey to see if they appreciated the PPAs and the FRCs.

See Appendices C - O for copies of instruments.

Commented [KAT9]: Is there any way of keeping this question all on one page. Not sure if it's necessary, just a pet peeve of mine.

In order to ensure validity of my surveys and interviews, I consulted with Dr. Erin Leininger, who is our Middle School Curriculum Coordinator. Her doctorate also included Likert surveys and interviews, so she was able to help me refine my questions to make sure that they were valid. The validity of the PPAs, SPAs and their grading as well as that of the exemplar was maintained in consultation with Ms. Pamela von Bodungen, who is the other sixth grade science teacher at my school. She helped me make sure that the PPAs and SPAs were aligned with the NGSS Performance Expectations that they were supposed to assess. If students were absent for the PPA, FRC, SPA and/or the Google survey day then they were excluded from that round of data collection which is why my *N* values (73 – 81) differ for each round. I did this because I did not think it was valid to include them if they missed a key component of the treatment.

In order to ensure the reliability of my results, I did four rounds of data collection (PPA, FRC, Google survey, interviews) with all four of my classes, so a total of 84 students. The first exception to this was that the students in one of the classes didn't do the Google survey for the first round of data collection. The second exception to this was that there were no interviews conducted in the first round of data collection and "only" 31 interviews conducted in the last three rounds of data collection.

#### DATA AND ANALYSIS

The results from all four rounds of data collection showed that the treatment was valued by students and was effective in preparing students for the SPAs, as evidenced by the score improvement from the PPAs to the SPA.

Table 2  
PPA versus SPA Scores

Assessment	N	Mean PPA Score	PPA Standard Deviation	Mean SPA Score	SPA Standard Deviation	Average Normalized Gain and Significance from PPA to SPA
Cells Investigation	80	10.2	2.1	12.0	1.9	0.47, Medium
Cell Analogy	73	8.7	2.7	11.2	1.9	0.47, Medium
Body Systems	81	7.5	2.8	10.2	2.4	0.41, Medium
Matter	75	4.5	1.4	5.8	1.1	0.50, Medium

Note. Maximum score = 14, except for Matter Assessment, Max. = 7 (N=80, 73, 81, 75)

The difference from the mean PPAs to the mean SPAs in all four instances indicates a significant improvement. The normalized gain for all was in the medium range. Normalized gain is normally used for pre-intervention versus post-intervention, as in before a unit has been taught to after a unit has been taught. The normalized gain was calculated to determine differences in student knowledge from the PPA to the SPA. Normalized gains of less than 0.3 are considered low gains, while 0.3 to 0.7 are considered medium gains. Finally, normalized gains greater than 0.7 are considered high gains (Hake, 1998).

In my Action Research, students have been taught the relevant part of the unit already when they take the PPA, so in my opinion, a medium gain from the PPA to the SPA a few days later is quite significant. The normalized gains in all four instances is narrow from 0.41 to 0.50, showing that the normalized improvement was consistent in all four instances. This shows that the intervention (PPA that is graded, followed by FRC, following by SPA) was quite successful in preparing students for the SPAs.

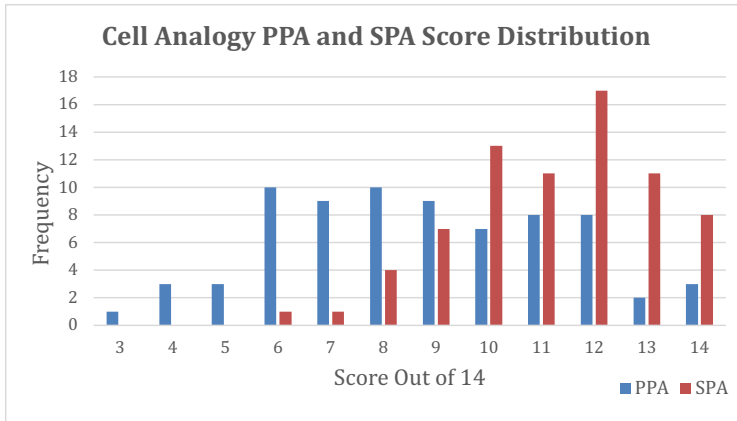


Figure 1. Cell Analogy PPA and SPA Score Distribution, (N=73)

It is quite apparent that students' scores improved significantly from the Cell Analogy PPA to the Cell Analogy SPA. On the PPA, 84% of scores were between 6 and 12 out of 14, with a mean score of 8.7. On the SPA, 92% of scores were between 9 and 14 out of 14, with a mean score of 11.2, which shows both an improvement in scores and a narrower distribution of scores. The latter is supported by Figure 1 and by the fact that the Cell Analogy SPA standard deviation (1.9) was significantly lower than that of the PPA (2.7). Quite simply, this shows that students learned by doing the Cell Analogy PPA, having it graded by me, and then having a final review class, as evidenced by the higher Cell Analogy SPA scores. Those who scored lower on the PPA had more room for improvement and had a greater overall improvement which is why the distribution of scores were less spread out for the SPA versus the PPA. This pattern is echoed in the frequency distribution graphs for the other three rounds of data collection:

Commented [KAT10]: period?

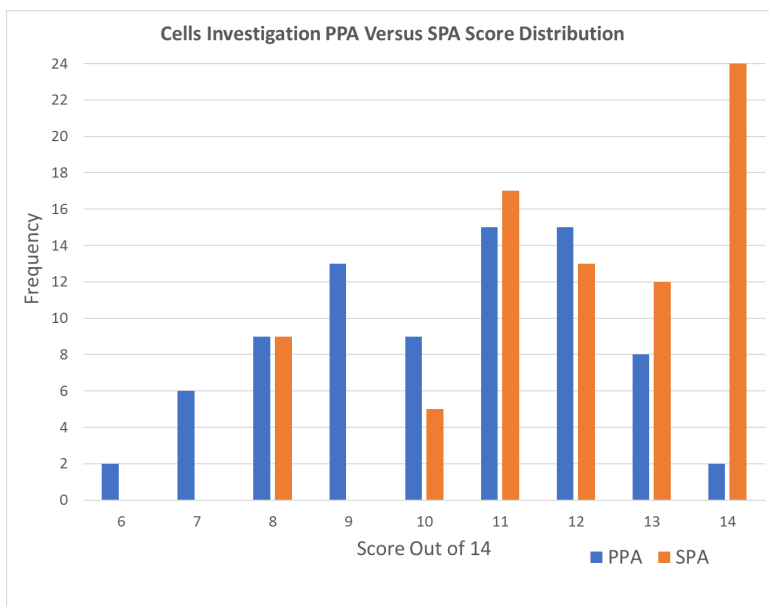


Figure 2. Cells Investigation PPA and SPA Score Distribution, ( $N=80$ )

For the Cells Investigation, 86% of students scored between 8 and 13 on the PPA, while 83% of students scored between 11 and 14 on the SPA. This shows that the scores improved significantly from the Cells Investigation PPA to SPA, which aligns with the mean score improving from 10.2 to 12.0. This also indicates that the spread of scores decreased significantly which is also supported by the PPA standard deviation (2.1) versus that of the SPA (1.9). All of this shows that the treatment (graded Cells Investigation PPA, final review class, Cells Investigation SPA) was effective in preparing students for the SPA.



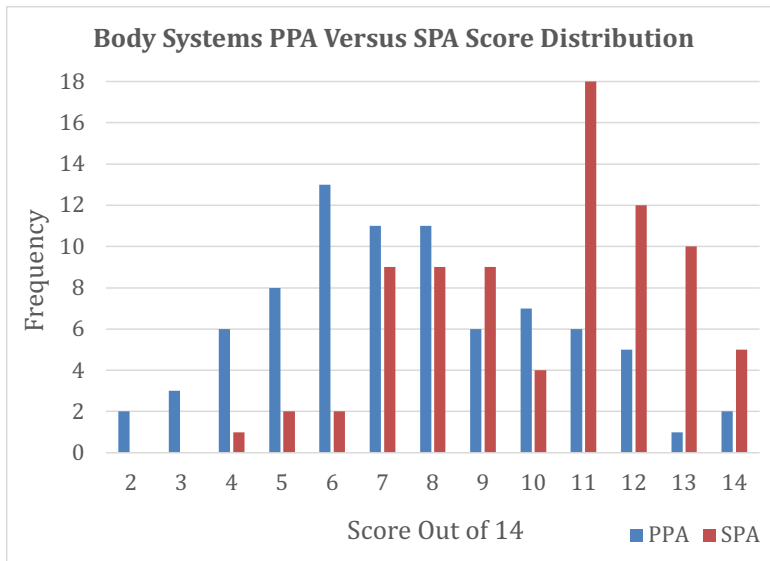


Figure 3. Body Systems PPA and SPA Score Distribution, ( $N=81$ )

In this instance, 90% of students scored between 4 and 12 on the Body Systems PPA, while 94% of students scored between 7 and 14 on the Body Systems SPA. This indicates that the scores improved significantly, which is supported by the mean score improving from 7.5 to 10.2. This also indicates that the spread of scores decreased significantly, which is supported by the PPA standard deviation (1.4) versus that of the SPA (1.1). All of this indicates that the treatment (graded Body Systems PPA, final review class, Body Systems SPA) was effective in preparing students for the SPA.

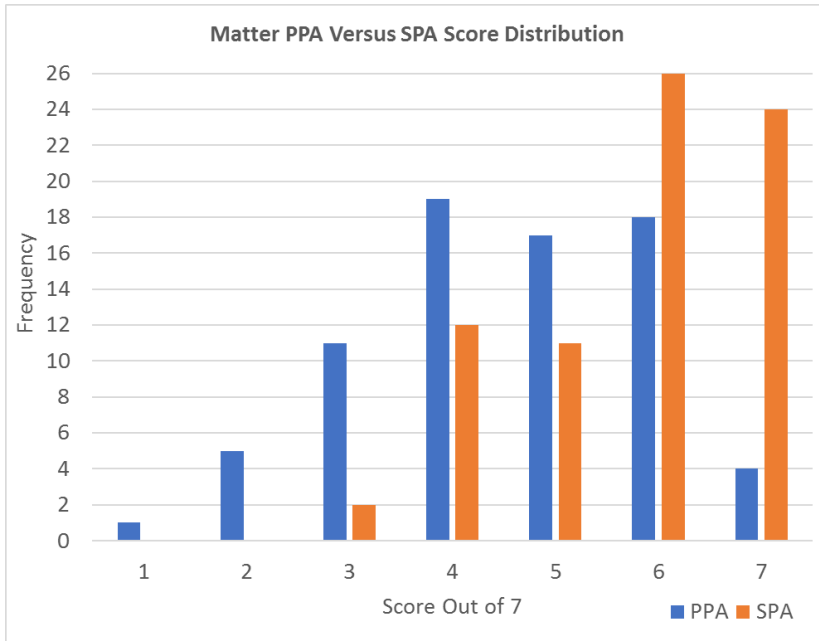


Figure 4. Matter PPA and SPA Score Distribution, ( $N=75$ )

In this instance, 92% of students scored between 2 and 6 on the Matter PPA, while 97% of students scored between 4 and 7 on the Matter SPA. This indicates that the scores improved significantly, which aligns with the mean improving from 4.5 to 5.8. This also shows that the spread of scores decreased significantly which is supported by the PPA standard deviation (2.8) versus that of the SPA (2.4). This is indicative of the effectiveness of the treatment (graded Matter PPA, final review class, Matter SPA).

In all four rounds of data collection, the spread of students' scores decreased from PPA to SPA. The probable reason for this is that lower PPA-scoring students had more room for improvement than the higher PPA-scoring students. Thus, compared to the PPA scores, the SPA scores ended up more clustered together at a higher range due to the

ceiling on the maximum possible score. This is echoed by the visuals of the PPA and SPA score distributions as shown in Figures 1, 2, 3, and 4.

The improvement in scores from PPA to SPA in all four rounds of data collection implies that the results were reliable as there were significant gains in all instances. This also indicates that the treatment (graded PPA then FRC then SPA) was effective. Clearly, the testing effect applies to performance assessments as well. The literature indicates that by doing a practice test that resembles the format, content, style of questions, and test conditions of the summative test, students will improve on a subsequent summative assessment (Adesope et al., 2017), which was supported by the PPA to SPA score improvement. The literature identifies that receiving graded feedback is also beneficial, thus I conclude that grading the PPAs helped students improve for the SPAs. The graded feedback let students know where they stood academically, thereby inspiring them (as privileged, mostly motivated expat students) to learn from their mistakes in skill and content. Seeing an exemplar PPA, enabled students to see what was expected of them, which they emulated in the SPA, causing significant improvement on the SPA. In the next section, I will delve into students' reasons and opinions to see what components of the treatment they found effective.

### Student Interviews and Surveys

Commented [KAT11]: Does this need to be pushed to the next page?

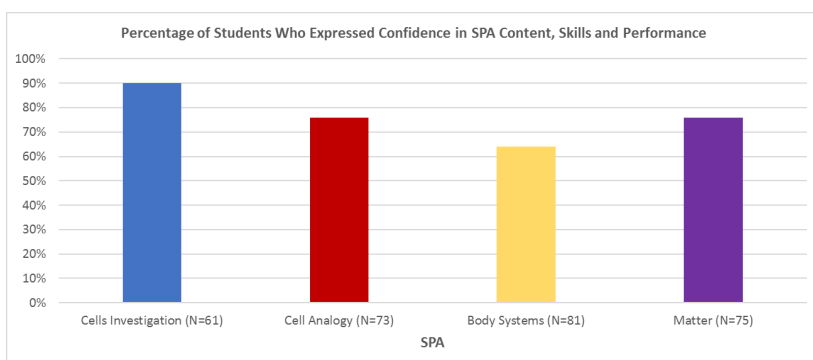


Figure 5. Post-SPA survey confidence results. ( $N=61, 73, 81, 75$ ).

Note: These percentages were the averages of the sum of the students who agreed or strongly agreed with Likert statements pertaining to these areas.

On three out of the four SPAs, students expressed high confidence in the content, skills and their performance on the SPA (75 – 90%), which shows that students felt well prepared for the SPAs. Thirty-one out of my eighty-four students were interviewed once after either the Cell Analogy SPA ( $N=8$ ), the Body Systems SPA ( $N=13$ ) or the Matter SPA ( $N=10$ ). Twenty-nine out of these students (94%) indicated that they felt quite comfortable with the material and skills on the assessment. Students reasons for being comfortable included:

“... , because of the two practice assessments, I learned a lot about how thermal energy affects solids, liquids, and gases”

“Okay and comfortable with it because we discussed it many times already and the things we don’t know are in the practice assessment and there’s an exemplar to show the right answer.”

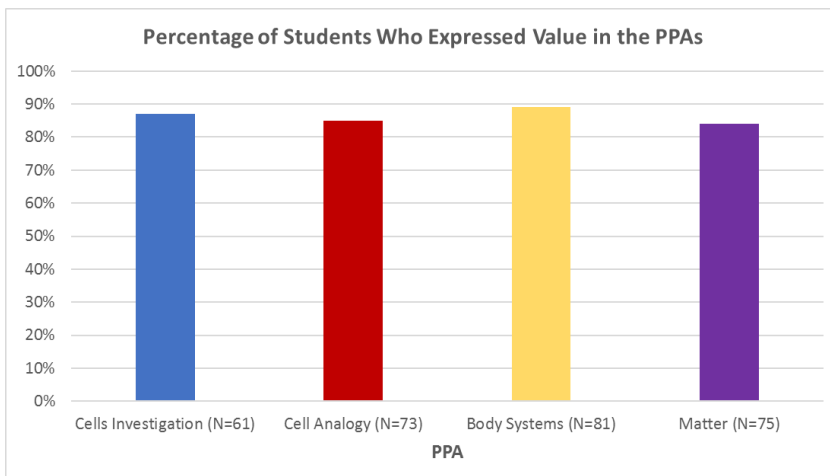
The interview and survey responses align with the SPA distributions of Figures 1 to 4, which show that a total of 78% of students achieved proficiency (10/14 or above, 5/7 or above) on the SPAs. It is safe to conclude that the treatment (PPA that is graded, followed by final review class, followed by SPA) resulted in high student self-confidence levels, which ended up being an accurate portrayal of how they ended up doing on the SPAs.

On a side note, student confidence was notably lower on the Body Systems SPA (64%) with only 46% of students who agreed or strongly agreed with the statement, “I am confident in how I did on the Body Systems final assessment.” Only 60% of students achieved proficiency (10/14 or more) on this SPA, which shows a positive correlation between grades and self-confidence. The Body Systems PPA and SPA was the most difficult performance assessment for my sixth-grade students as it was their first complex CER. These performance assessments were designed to align with MS-LS1-3 which states “Use argument supported by evidence for how the body is a system of interacting subsystems composed of groups of cells.” (NGSS Lead States, 2013). In retrospect, both the content and skill required for this Performance Expectation was a bit much for sixth graders, which is why we have decided to take it out of our unit for next school year.

That said, eleven out of the thirteen students (85%) who were interviewed after the Body Systems SPA, gave positive responses which supported the treatment even though the topic may have been a bit much for them:

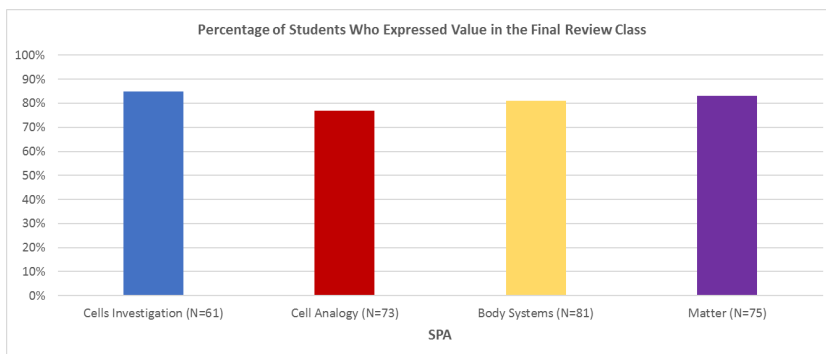
“Comfortable because we went over the body systems after the practice assessment because no one really understood what they do.”

Interviewed students were asked “Did you feel adequately prepared for the summative assessment? Why?” Not surprisingly, 30 out of 31 (97%) students responded favorably because they did well and/or showed improvement from the PPAs to the SPAs. Seven interviewed students (23%) responded that their notes helped them feel adequately prepared because they could look back through them. One student’s reason was, “Yes, because I knew what I was going to write. I looked back through my notes that I took which really helped me with the test.” Fourteen interviewed students (45%) referred to the PPA as helping them. This is evident in this student’s response, “Yes, because all the practice assessments helped me be ready for the test.” Again, these responses show that students valued the PPAs as indicated by the fact that the vast majority felt prepared for the SPAs.



*Figure 6.* Post-SPA survey about students’ value of the PPA. (N=61, 73, 81, 75).  
*Note:* These percentages were the students who agreed or strongly agreed with a Likert statements indicating the value of each PPA.

Eighty-four to eighty-nine percent of surveyed students found value in the PPA. Similarly, all thirty-one interviewed students indicated that the PPA was effective in preparing them for the SPA, which further validates the PPA part of the treatment and aligns with the literature surrounding the testing effect. Fifteen of these students (48%) referenced how seeing the format of the PPA prepared them for the SPA, with this quote supporting that, “Yes, because we did a similar format and type of question on the practice assessment as on the final assessment.” This is supported by the literature which argues that the learning procedure (the PPA) should be like that of the testing procedure (Bransford et al., 1979). What’s the best way to study for the SATs? Complete practice SATs and learn from one’s mistakes. What’s the best way to prepare for a driver’s test? Practice driving and complete the things that will be done in the driver’s test. Thus, it is not surprising that students find value in a PPA that is similar in style and content to that of the SPA.



*Figure 7.* Post-SPA survey about students’ value of the Final Review Class. (N=61, 73, 81, 75).

*Note:* These percentages were the averages of the sum of the students who agreed or strongly agreed with Likert statements pertaining to the Final Review Class.

After all four SPAs, the number of surveyed students who expressed value in the final review class (FRC) was high, ranging from 77 to 85% of students. Twenty-nine out of thirty-one interviewed students (94%) commented that the FRC was effective and/or helpful. Thirty-one percent of the interviewed students referenced the exemplar in helping them improve, while in the Google surveys, 82 – 92% of students agreed or strongly agreed with statements in support of the exemplar. Here's a student's reason for the effectiveness of the exemplar, "The exemplar was helpful because I got to compare mine to the exemplar and see all the things that I didn't have on my test that the exemplar had."

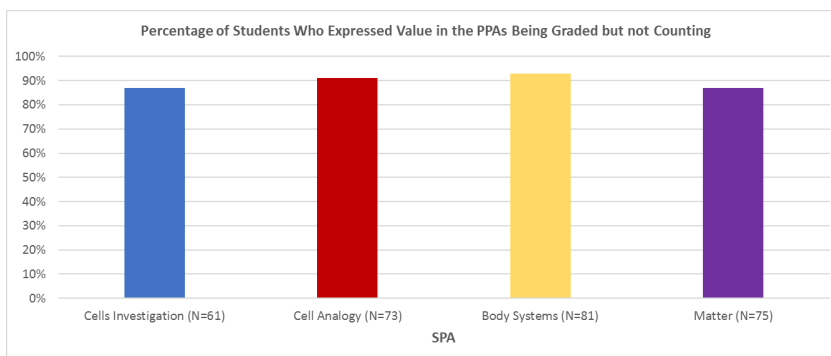
Fifty-two percent of interviewed students and 77 – 84% of Google surveyed students indicated that correcting their mistakes during the FRC helped them learn and improve for the SPA. This student's opinion echoed the sentiment of many of these students, "I learned what to write for the claim because I was confused before. The checking corrected my mistakes so for the real test, I can remember to pay special attention." Clearly, seeing one's PPA, then being able to compare it to a student exemplar and then correct it, enabled students to learn from their mistakes. Seeing an exemplar clarified the expectations that I had for students. It is easier for students to know what they are supposed to do if they can see an example of it. This is why MSU posts exemplar MSSE AR papers on their website. By modeling the structure of my writing after said papers, I improved the quality of this paper to make it the masterpiece that it is.

Forty-eight percent of interviewed students valued the reteaching of the content that took place during the FRC, which I was able to do after I saw the misconceptions and



gaps in learning that were prevalent in the PPAs. Here is a student quote supporting this part of the FRC, “Good, because we explained the function of all three (body) systems and got it in my head and then I went home and researched it to get it more in my head.”

It can be concluded that the final review class component of my treatment was valued and effective because it enabled students to learn from their mistakes by correcting them with the aid of the exemplar and reteaching of content. The aforementioned grade increases from the PPAs to the SPAs is also evidence that the FRC between them was instrumental in improving student learning and achievement.



*Figure 8.* Post-SPA survey about students’ value of the PPAs being graded but not counting towards their overall grades. (N=61, 73, 81, 75).

*Note:* These percentages were the averages of the sum of the students who agreed or strongly agreed with Likert statements pertaining to the grading of the PPAs.

After all four SPAs, the number of students who expressed value in the PPA being graded was very high, ranging from 87 to 93% of students, thus it can be concluded that grading the PPAs was an effective measure in helping students do better on the SPAs. Interviewed students were asked whether getting graded on the PPA was helpful and why. All thirty-one interviewed students indicated that the grading of the PPA was helpful. Seventy-one percent of them indicated that the grading helped them identify and

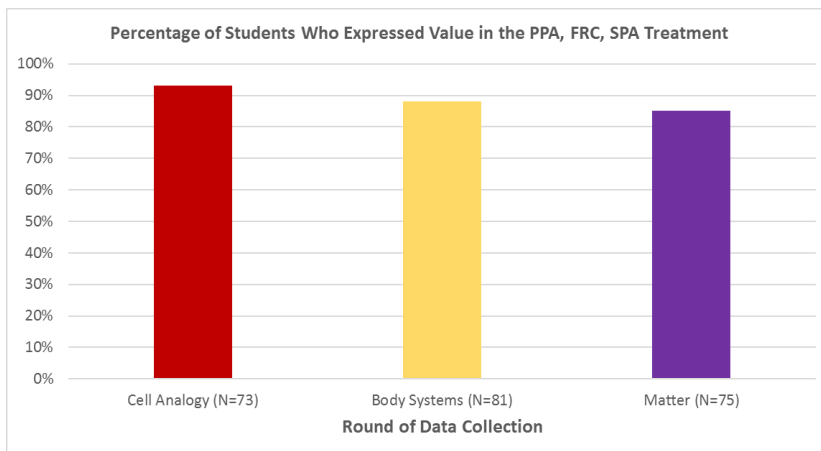
correct their mistakes. Nineteen percent of them valued that the grading told them what they were already good at. Thirty-five percent of students indicated that the low grade on the PPA was motivation for them to make improvements so they would do better on the SPA. This student's response exemplifies these sentiments, "Because I got the grade, I knew I had to prepare more for the actual test. It would be harder if I didn't know my grade because I wouldn't know what I had to fix and what I needed to study more of." One high-achieving student (PPA score = 11, SPA score = 14) referred to her fear of failure, "I wasn't very pleased by my practice assessment grades but then it really helped me for my final assessment because after feeling a bit angry, I practiced more because of the thought of failing the final assessment."

This aligns with the literature in which a meta-analysis of many testing effect research studies found that feedback on practice tests was beneficial to learning (Phelps, 2012). Phelps defined feedback as grades and/or comments.

My grade six students don't know what they don't know. Without receiving evaluative feedback in the form of grades, they often think that their work is similar enough to that of an exemplar. The grades inform students if they do or don't know content and/or skills (how to do/write/communicate something). The feedback provided by the grades on each part of the PPAs told them where they needed to improve. The FRC and exemplar helped them identify how to improve.

By receiving graded feedback on each part/question, students were able to identify their strengths and weaknesses, and then make necessary corrections and/or study deficient components. Prior to the grading of the PPAs, students hadn't been graded on anything for that part of the unit, giving some students a false sense of ability.

For many students, the graded PPA was a wake-up call that let them know where they stood with regards to the skills and concepts necessary for success on the SPA. Furthermore, the overall PPA grade provided students motivation to put in the necessary effort and learning to do better on the SPA.



*Figure 9.* Post-SPA survey about students' value of the treatment. (N=73, 81, 75).

*Note:* These percentages were the averages of the sum of the students who agreed or strongly agreed with a Likert statement pertaining to the overall treatment (PPA, followed by FRC, followed by SPA). Regrettably, this question was not asked for the first round of data collection (Cells Investigation).

Overall, students valued the treatment with 85 – 93% of them agreeing that doing a practice assessment followed by a final review class followed by a final assessment is a good way to prepare for final assessment. Their value of the treatment correlates well with the significant grade increases from the PPAs to the SPAs. The reasons for this effectiveness were:

- By receiving graded feedback, students were able to identify their strengths and weaknesses. Prior to this, students hadn't been graded on anything for this part of the unit, giving some students a false sense of ability. For many students, this was

Commented [KAT12]: Single or double space?

a wake-up call as the grade on the PPAs let them know where they stood with regards to the skills and concepts necessary for success on the SPAs. This motivated them to correct their work, learn from their mistakes and fill in any gaps in content and/or skills.

- By receiving an exemplar during the FRCs, students were able to see more clearly what was expected of them enabling them to do better on the SPAs.
- By grading student work, I was able to identify and tailor each FRC to address most deficiencies in the PPAs, content, and skills, which helped students learn from their mistakes, acquire content and skills and thus do better on the SPAs.

#### INTERPRETATION AND CONCLUSION

My primary research question is:

- How can I use practice performance assessments (PPAs) to prepare students for summative performance assessments (SPAs)?

Due to the significant improvement in student performance between the PPAs and SPAs for almost all students (except for those PPA scores who were perfect or near perfect), I conclude that PPAs are quite effective in preparing students for the SPAs. This aligns with the literature surrounding the testing effect which refers to gains in learning and achievement that occurs when students take a practice test that resembles the format, content, style of questions, and test conditions of the summative test (Adesope, et al., 2017). This opinion was shared by the surveyed students, as 84 - 89% of them agreed that that PPA was good preparation for the SPA. During the interview phase, students' reasons for this were the similar format, content, and style of questions, which aligns with the literature. The most recurring student reason was knowing the expectations and

Commented [KAT13]: Single or double space.

learning from their mistakes. Doing an SPA that is similar in format and content to its PPA will probably result in improvements if students are motivated to learn from their mistakes. Accordingly, grade motivation was given as a reason by 35% of interviewed students which I would argue is the case with the high socioeconomic students that attend our school.

Personally, I will continue to use the graded PPA then FRC (with a student exemplar) then SPA methodology long after I graduate from the MSSE program. During my member feedback session with my students, I shared the results with them so they would be meta-cognitively aware of what is good for their learning and achievement. It is my hope that they will use this information to advocate for graded practice assessments and final review classes in their other classes.

The most common recommendation that I received from students was more practice. For the Body Systems and Matter rounds of data collection, I introduced an optional extra practice assessment, which students could do and correct at home (exemplar posted on Google Classroom), between the FRC and SPA. Before the Body Systems and Matter PPAs, I began doing practice for the practice assessment in order to better prepare students for the PPAs. I hope that this will result in higher PPA and SPA grades. In the future, I will continue to try to give students practice assessments before and after the PPA.

Two of my secondary AR questions are:

- How can I use students results on PPAs to prepare them for SPAs?
- What are students' opinions of PPAs and final review classes?

Commented [KAT14]: an SPA?

Eighty-five to ninety-three percent of students agreed that doing a practice assessment followed by a final review class is a good way to prepare for actual assessments. This opinion coupled with the PPA to SPA score increase, leads me to believe that the PPA, followed by an FRC is a good way to prepare students for an SPA.

The elements of the FRC that students deemed effective were:

- Getting grades on their PPAs that didn't count towards their overall grade (87 – 93% agreement).
- Being able to see an exemplar PPA (79 – 92% agreement).
- Being able to correct their PPAs (77 – 84% agreement).
- Receiving personalized feedback/clarification from the teacher (62 – 73% agreement)

In future, I will continue to do the above during the FRCs. In my Action Research, the SPAs were given two to five days after the PPAs. This aligns with the literature which found that the optimal time between practice and summative tests should be between one and six days (Adesope et al., 2017). I will continue to do this in the future.

My next AR question centered around grading students PPAs:

- How does assigning a grade to students PPAs impact their subsequent performance on SPAs?

Judging from the vast improvement from PPA scores to SPA scores, I would argue that assigning a grade that didn't count towards their overall grade (as in their grades were recorded on my research data sheet and on their PPAs but not in the gradebook) had a significant positive impact on their subsequent SPA performance. This opinion is supported by student opinions in the survey as over 87 – 93% of students favored the grading of their PPAs. This is also supported by the literature in which a meta-analysis

Commented [KAT15]: an?

Commented [KAT16]: I noticed that most other bullet point end with a period. Did you intend one for each of these as well.

of many testing effect research studies found that feedback on practice tests was beneficial to learning (Phelps, 2012). Phelps defined feedback as grades and/or comments.

Some students (1 – 4%) who had perfect PPA scores ended up with lower SPA scores. The reason for this is that once you have achieved perfection, there is nowhere to go but down or stay in the same spot. My reflection on this is, “Should students who have shown exemplary (14 out of 14) or near exemplary (13 out of 14) performance on a PPA be required to take the SPA?” Making these students do the SPA is like passing a driver’s license test and then two days later having to try to pass it again. In future, I am inclined to let these students’ PPA scores count as their SPA scores, and not require them to take the SPA. I piloted this during the COVID-19 imposed distance learning in April and May of 2020, and it was well-received by students although I did not collect any formal data on this. Students informally told me that this reduced their stress and provided extra motivation for them to do well on the PPA. It also had a side benefit of decreasing some of the SPA grading.

My final question to analyze is:

- How did using PPAs impact me as a teacher?

As noted in my journal, using PPAs had an overall positive impact on me as a teacher, because I saw the vast improvement both during the PPA and SPA classes, and after I had graded them. The following journal entries document the positive impact that using PPAs had on me as teacher:

- After all four PPAs, I noted that some students found the PPA quite challenging and were down on themselves afterwards but were much more confident during

Commented [KAT17]: My only idea here is can you determine if they got lucky the first time. Did they make a couple of lucky guesses on answers that panned out?

the SPA. I noted six occasions in my journal where students were reduced to tears during the PPA or when they received their PPA grades as opposed to no occasions for the SPA.

- During all four PPAs, each class asked many more questions than they did during the SPAs. Quotes from my journal during the PPAs which support this include, “Wow that was painful!” and “SO MANY QUESTIONS!”
- During all four FRCs, I noted that upon seeing their PPA grades, students were engaged and motivated to identify and correct their mistakes and/or gaps in learning. In my journal, I noted that 80 – 90% of students were focused and on task for the first two-thirds of the FRC, with that number dropping to 50 – 70% for the last-third of the FRC. This is a high number for a non-investigation or assessment class for my sixth-grade students.
- Grading students PPA performance was a bit deflating during the Body Systems and Cell Analogy PPAs. In my journal, I wrote things like “I can’t believe that students weren’t able to identify the correct body systems,” and “Some students really don’t know the functions of the organelles.” Of course, this helped me plan the FRCs.
- Accordingly, seeing the vast improvement and much higher scores on all four SPAs was uplifting to me, because it validated my treatment and showed increased academic performance in my students. Quotes from my journal include, “That was much better,” and “Most students wrote much better CERs.” It was uplifting to me to see that my treatment was effective. Similarly, the



positive responses that students wrote in the Google survey reaffirmed that my whole PPA, FRC (with exemplar), SPA methodology is a good one.

The quick turnaround required to have the PPAs graded and photocopied, and an exemplar made and photocopied did have a negative impact on me on the two occasions when the PPA was not during the last class of the week. For the body systems PPA, I had to wake up in the middle of the night to finish grading them. In my journal, I noted that “I was exhausted and irritable” during the Body Systems FRCs. In future, I should try to schedule the PPA for the first, third or fourth lesson of the week in order to avoid this scenario.

#### Implications on Teaching Practices

Clearly the treatment was effective, thus in future I will continue to:

- Use the PPA then FRC (with exemplar and graded PPA) methodology.
- Give the SPA to students two to five days after the PPA.
- Make sure that the PPAs resemble the format, content, style of questions, and test conditions of the subsequent SPAs.
- Have an extra practice task that is ungraded both before and after the PPA, for which students can see an exemplar.
- Plan the PPAs so that I have time to grade them.
- Allow students who do well on a PPA to be exempt from the subsequent SPA.
- Survey students about the PPAs, FRCs, and SPAs twice per semester.
- Share with students the results of the treatment, literature, and surveys so that they are meta-cognitively aware of what is good for their learning and achievement.

- Share my findings with colleagues in department meetings, department head meetings, and school-wide professional development workshops.

#### VALUE

First and foremost, my research expands upon and validates the research surrounding the testing effect. My research shows that the testing effect doesn't apply to just content-based tests. It also applies to skills such as focusing a microscope, analyzing data, making an analogy, writing up a lab report (PPA and SPA from the semester after this research was done). Students learn from their mistakes, and it is helpful if those mistakes are identified through grading. My students are motivated by grades, so a graded practice assessment that doesn't count towards their overall grade can induce them to do better.

To further enhance this process, I should find other methods that I can use during the FRCs in order to enhance student learning and achievement. Next semester, I will continue to trial different ideas during the FRCs. I will look in the literature and speak to colleagues in order to improve this element of my lessons.

During the surveys and interviews, the most common request was for more extra practice assessments either before or after the PPA. I began doing this for the Body Systems and Matter PPA/SPA. I continued doing this during the spring semester of 2020, but I did not collect any data on it. In future, I will collect data to see the improvement of students who complete the extra practice assessment (between the FRC and SPA) at home, who complete and correct it, and compare those two groups' improvement of that of the students who did no extra preparation. Then I will

Commented [KAT18]: validates?

communicate those findings to students. I will do the same thing for students who do an extra practice assessment before the PPA.

My initial methodology had a treatment (graded PPA) group and a control group (ungraded PPA) for at least one round of data collection, in order to see how grading the PPAs would impact student performance on the SPAs. In my surveys and interviews, 87 – 93% of students indicated a strong preference for the grading. Thus, I deemed it somewhat unethical to not grade their PPAs because this was something that they valued, and which seemed to be a key component in the PPA to SPA learning and grade improvement. Therefore, I changed my research plan and graded all the PPAs for all four rounds of data collection.

Hence, a next step in the research would be to have a treatment (ungraded PPAs) class and a control class (graded PPAs), which is randomly selected. The groups/classes could switch for each round of data collection so that students could be compared to each other, as well as to themselves. After the fact, photocopied versions of the ungraded PPAs could be graded. They could then be used to see if students who received ungraded PPAs improved as much as on the SPAs as those who received graded PPAs. This would enable a researcher to see the impact that grading has on student performance on SPAs.

As mentioned in the methodology section, all my PPAs and SPAs were open notes. During the SPAs, students had access to the exemplar and their corrected PPAs. The results of my research may not apply to non-open notes performance assessments. Thus, a next step would be to replicate my research without the open-notes component to see if the PPA to SPA score gains are as significant and if student opinions of the treatment are as positive.

Commented [KAT19]: then be used?

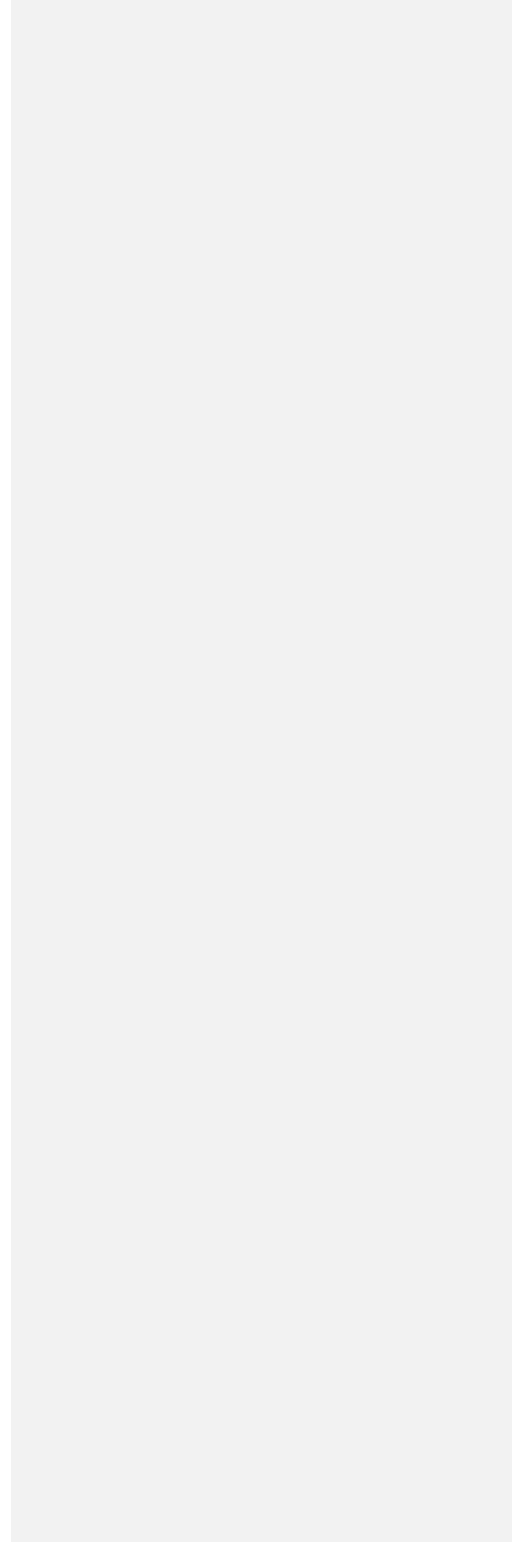
Commented [KAT20]: I would highly encourage this change.

Finally, my findings apply to privileged, motivated, predominantly expatriate Asian sixth-grade science students. Future researchers could study how the testing effect applies to students of different:

- Ages
- Disciplines (English, Social Studies, PE, Art, Modern Languages, etc.)
- Ethnicities
- Nationalities
- Socioeconomic backgrounds

Throughout this study and in the semester afterwards, I observed the positive impact that graded Practice Performance Assessments have on student learning and achievement as measured by students' grades on their subsequent Summative Performance Assessments. The reflections that I collected from students further validated this assessment preparation process and informed me that students valued what I was doing. Furthermore, students gave me valuable suggestions that I incorporated both during my research and in the semester afterwards in order to maximize their learning and achievement. In short, my Action Research into using Practice Performance Assessments to prepare students for Summative Performance Assessments has enabled me to become a better teacher, researcher, and writer.

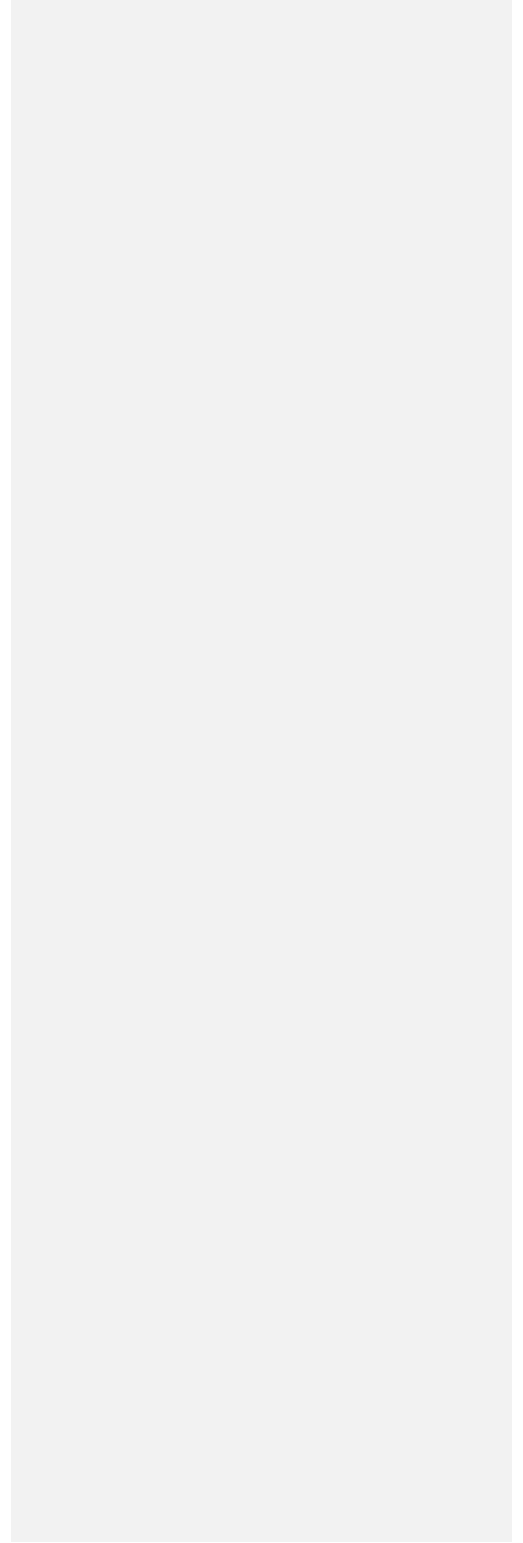
REFERENCES CITED



- Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701. Retrieved from <https://doi-org.proxybz.lib.montana.edu:3443/10.3102/0034654316689306>
- Alt, D. (2018). Teachers' practices in science learning environments and their use of formative and summative assessment tasks. *Learning Environ Research, 21*(3), 387–406. Retrieved from <https://doi-org.proxybz.lib.montana.edu:3443/10.1007/s10984-018-9259-z>
- Allen, G. A., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning & Verbal Behavior, 8*(4), 463–470. Retrieved from [https://doi.org/10.1016/S0022-5371\(69\)80090-3](https://doi.org/10.1016/S0022-5371(69)80090-3)
- Carpenter, S. K., Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review, 14*(3), 474–478. Retrieved from <https://eric.ed.gov/?id=ED505596>
- Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1979). Some general constraints on learning and memory research. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*, 331–354. Hillsdale, NJ: Erlbaum.
- Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*(4-5), 514–527. Retrieved from <http://dx.doi.org/10.1080/09541440701326097>
- Giordano, G. (2005). *How testing came to dominate American schools: The history of educational assessment*. New York, NY: Peter Lang.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74.
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills*, 35–56. New York: Springer.
- Karpicke, J., & Aue, W. (2015). The Testing Effect Is Alive and Well with Complex Materials. *Educational Psychology Review, 27*(2), 317–326. Retrieved from <https://doi-org.proxybz.lib.montana.edu:3443/10.1007/s10648-015-9309-3>
- Lasry, N., Charles, E., & Whittaker, C. (2014). When teacher-centered instructors are assigned to student-centered classrooms. *Physical Review Special Topics—Physics Education Research, 10*(1), 1–9.

- NGSS Lead States. (2013). Next Generation Science Standards: For States, By States. MS-LS1 From Molecules to Organisms, MS-PS1 Matter and Its Interactions. Retrieved from <http://www.nextgenscience.org/>
- Panadero, E., Jonsson, A., & Strijbos, J. W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation*, 311–326. Boston, MA: Springer.
- Phelps, R. P. (2012). The effect of testing on student achievement, 1910–2010. *International Journal of Testing*, 12(1), 21–43.
- Roediger, H., & Karpicke, J. (2006). Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, 17(3), 249-255.
- Shavelson, RJ, Ruiz-Primo, Maria Araceli, & Shavelson, Richard J. (1996). Rhetoric and reality in science performance assessments: An update. *Journal of Research in Science Teaching*, 33(10), 1045-1063.
- Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: the testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review*, 27(2), 247–264. Retrieved from <https://doi-org.proxybz.lib.montana.edu:3443/10.1007/s10648-015-9310-x>
- U.S. Department of Education. (2008). Great expectations holding ourselves and our schools accountable for results. Retrieved from <https://files.eric.ed.gov/fulltext/ED504194.pdf>

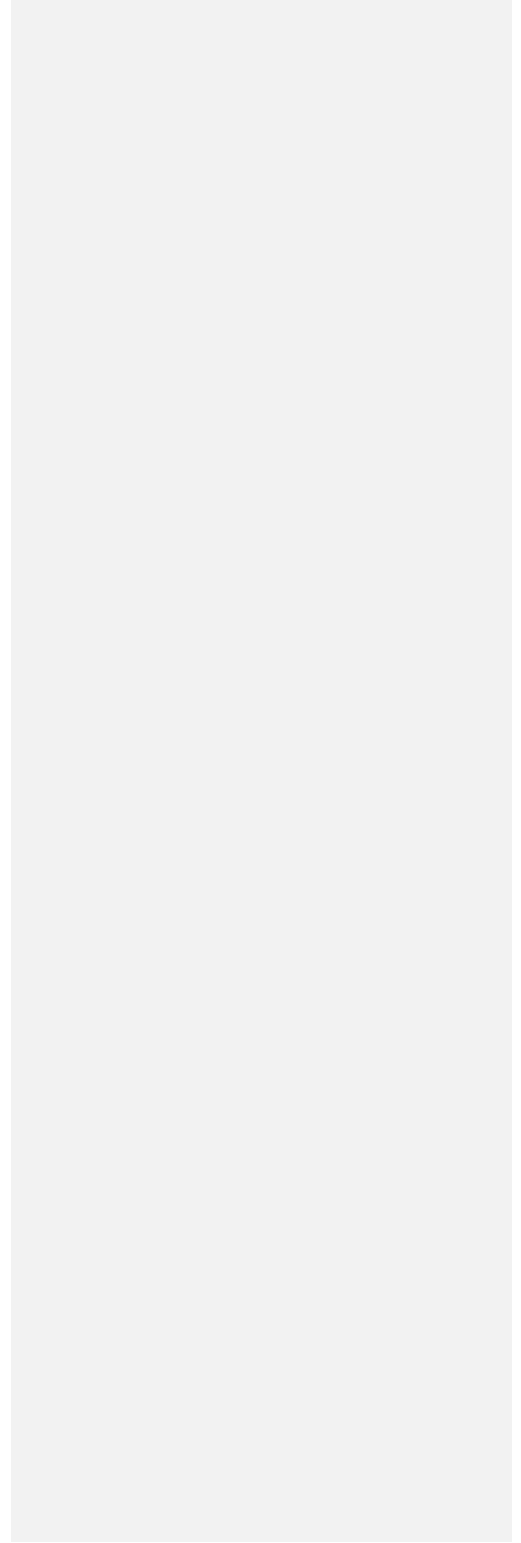
APPENDICES





APPENDIX A

INSTITUTIONAL REVIEW BOARD EXEMPTION





2155 Analysis Drive  
c/o Microbiology & Immunology  
Montana State University  
Bozeman, MT 59718  
Telephone: 406-994-4706  
FAX: 406-994-4303  
E-mail: cherylj@montana.edu

Chair: Mark Quinn  
406-994-4707  
mqquinn@montana.edu  
Administrator:  
Cheryl Johnson  
406-994-4706  
cherylj@montana.edu

**INSTITUTIONAL REVIEW BOARD**  
**For the Protection of Human Subjects**  
**FWA 00000165**

**MEMORANDUM**

**TO:** Neil Freeman and Walter Woolbaugh

**FROM:** Mark Quinn *Mark Quinn CJ*  
Chair, Institutional Review Board for the Protection of Human Subjects

**DATE:** September 26, 2019

**RE:** "The Effects of Using Practice Performance Assessments to Prepare Students for Summative Performances in Grade 6 Science" [NF092619-EX]

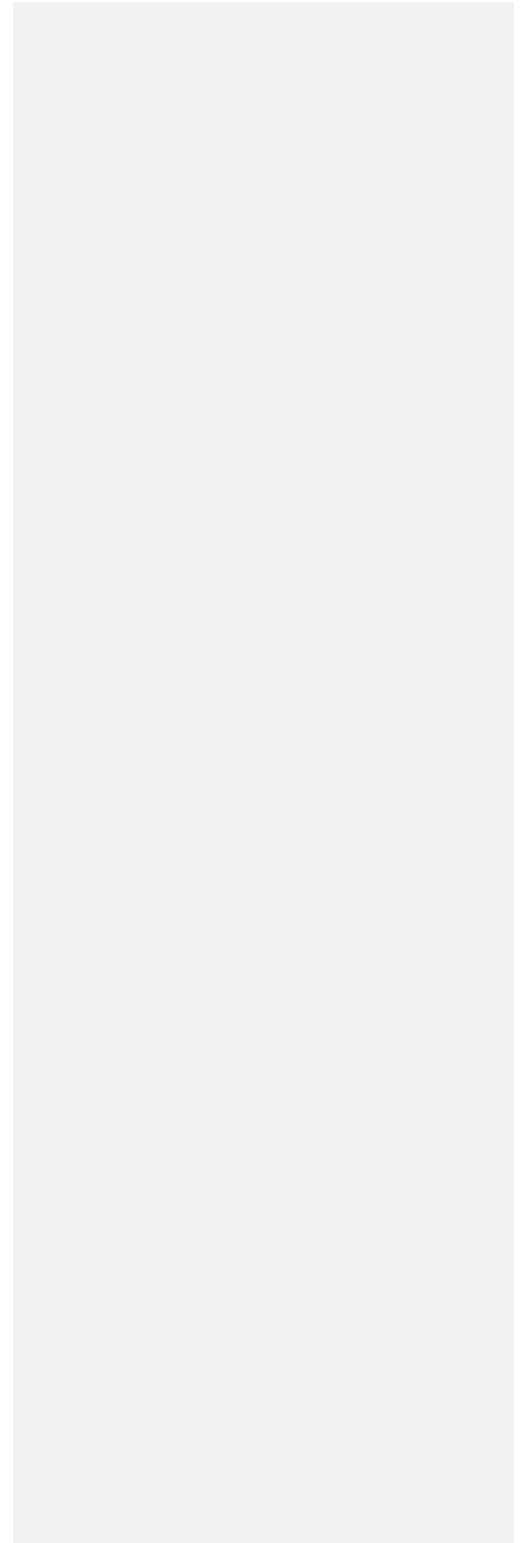
The above research, described in your submission of September 25, 2019, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

- (b) (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
- (b) (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation; and (iii) the information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by section 16.111(a)(7).
- (b) (3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if: (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.
- (b) (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available, or if the information is recorded by the investigator in such a manner that the subjects cannot be identified, directly or through identifiers linked to the subjects.
- (b) (5) Research and demonstration projects, which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.
- (b) (6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed, or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the FDA, or approved by the EPA, or the Food Safety and Inspection Service of the USDA.

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

APPENDIX B

NGSS PERFORMANCE EXPECTATIONS FOR THE PPAs AND SPAs



*NGSS Performance Expectations and Description of Performance Assessment (PA)*  
(NGSS Lead States, 2013)

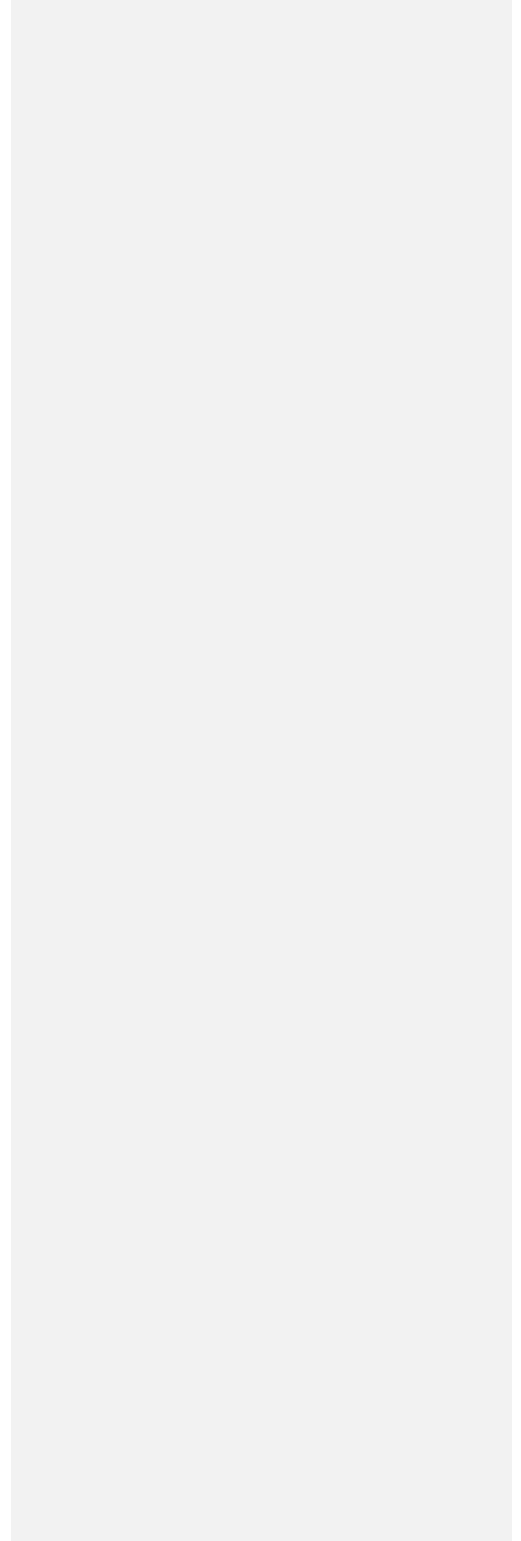
NGSS PE	NGSS PE	PA Description
MS-LS1-1 From Molecules to Organisms: Structures and Processes	Conduct an investigation to provide evidence that living things are made of cells; either one cell or many different numbers and types of cells.	<p>Students prepared their own wet mount and focus it on the microscope, showing the teacher once its focused. The teacher assessed whether students have done this or not.</p> <p>Students wrote a CER (Claim, Evidence, Reasoning) to support a claim that their unknown sample is made of cells. Their evidence was their drawing of a small part of what they observed. It should have clearly showed what one cell of their sample looks like. Their reasoning should have been evidence-based and should link their evidence back to their claim.</p>
MS-LS1-2 From Molecules to Organisms: Structures and Processes	Develop and use a model to describe the function of a cell as a whole and ways parts of cells contribute to the function.	Students constructed an analogy for something (airplane, restaurant, classroom) that is like a plant cell, clearly detailing how the part of their thing (airplane, restaurant, classroom) is like the similar part of a plant cell. Their focus should have been on the similarity of the function of the parts, and how that function contributed to the overall function of the plant cell and the thing (airplane, restaurant, classroom).
MS-LS1-3 From Molecules to Organisms: Structures and Processes	Use argument supported by evidence for how the body is a system of interacting subsystems composed of groups of cells.	Students were given evidence and had to detail how that evidence showed that two body systems are interacting. They used the CER framework to first write a claim stating which two body systems were interacting in the given scenario. Their reasoning was evidence-based which showed how the interaction of the two body systems caused the evidence to occur, and thus how that evidence demonstrates that the two given body systems were interacting. The circulatory, respiratory, and digestive systems were studied during the unit.
MS-PS1-4 Matter and its Interactions	Develop a model that predicts and describes changes in particle motion, temperature, and	Students used the CER framework to make a claim about how changes in thermal energy affected a given scenario. The evidence was given to them. Their reasoning was evidence-based and included a model that showed the

	state of a pure substance when thermal energy is added or removed.	changes that occurred to the molecular motion, temperature, and/or state of a pure substance when thermal energy was added or removed. Their model and description should have shown the pure substance before and after the thermal energy change.
--	--	---

The above contains the NGSS Performance Expectation (PE) and a brief description of each of the four performance assessments that went along with said PE.

APPENDIX C

CELLS INVESTIGATION PRACTICE PERFORMANCE ASSESSMENT



**Inner Cheek Skin CER** Name \_\_\_\_\_ Homebase \_\_\_\_\_  
 During this task, you will: “Conduct an investigation to see if your inner cheek skin is living.”

**Here’s what you need to do:**

1. Prepare a slide with your inner cheek skin, focus it on high power and show your teacher \_\_\_\_\_
2. Complete the following **Claim, Evidence, Reasoning** framework, which satisfies the **bolded** part above, which is the point of this investigation:

**Claim** (Is your inner cheek skin living? Complete Sentence!)

**Evidence** (Sketch your inner cheek skin on high power)

Sketch components:

1. Title
2. Colour
3. Accurate sketch
4. Magnification
5. Labelled organelles
6. One Cell Circled and Identified

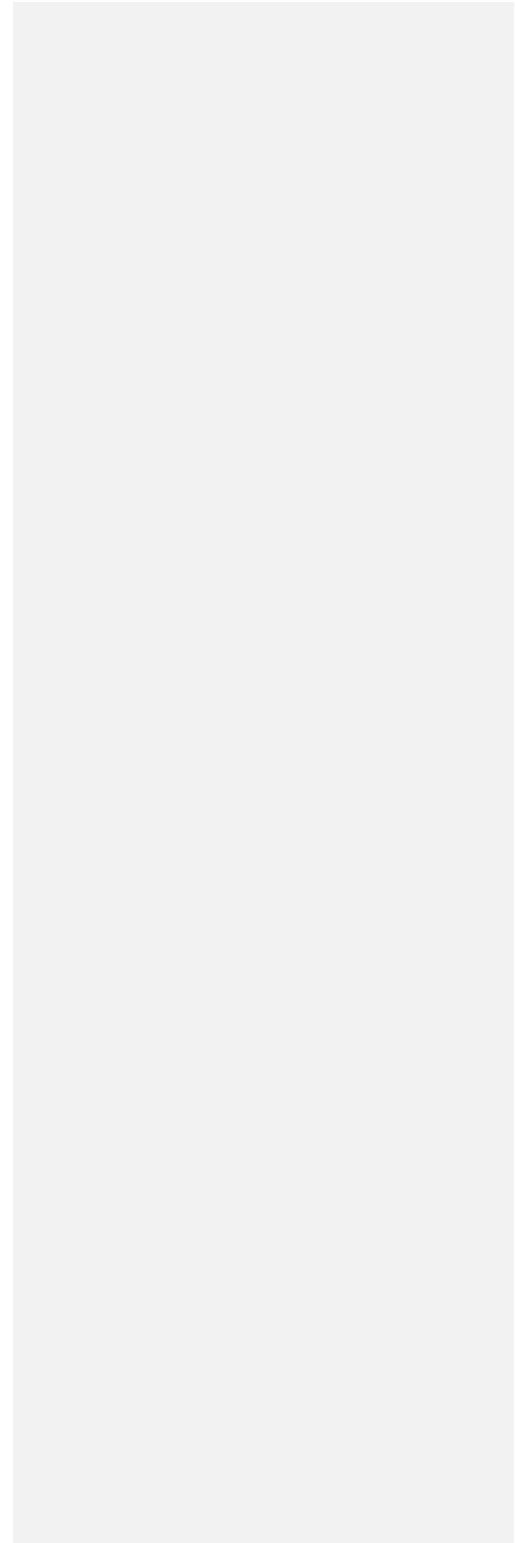
<b>Reasoning</b> - How does your evidence support your claim?
<b>Step 1:</b> Show what you know
<b>Step 2:</b> Connect the evidence to the claim. How does your sketch show that your inner cheek skin is living or not?
<b>Step 3:</b> Closing statement. Therefore.....

SELF-ASSESS using checkmarks in pencil on the rubric on the back.

<b>Knowledge and Understanding</b>	Always (7) <b>Exemplary</b>	Usually (5) <b>Proficient</b>	Somewhat (3) <b>Developing</b>	Rarely (1) <b>Emerging</b>
Claim is accurate.				
Sketch (evidence) is accurate.				
Sketch (evidence) has the correct title and magnification.				
Organelles (parts) are correctly labelled.				
Overall				

±

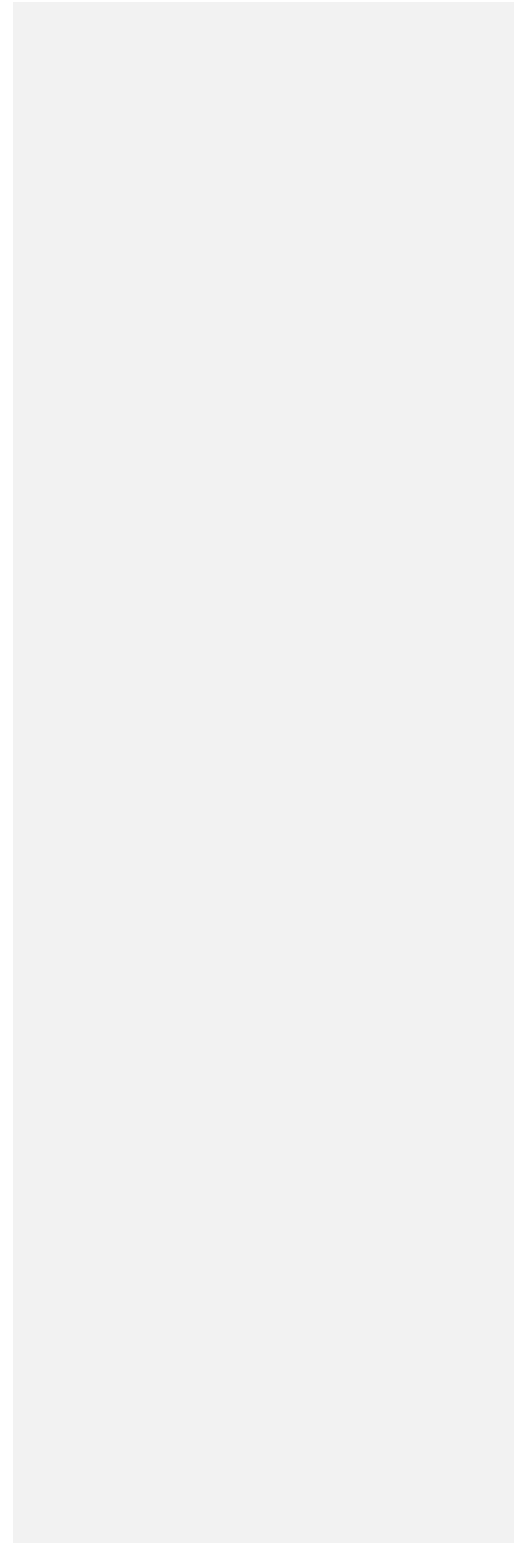
<b>Transfer of Learning</b>	Always (7) <b>Exemplary</b>	Usually (5) <b>Proficient</b>	Somewhat (3) <b>Developing</b>	Rarely (1) <b>Emerging</b>
Student was able to focus their slides on the cells at high power without teacher or peer assistance.				
Student's reasoning connected the evidence to the claim.				
Student's evidence clearly identified what one cell of their inner cheek skin looked like.				
Overall				





APPENDIX D

CELLS INVESTIGATION SUMMATIVE PERFORMANCE ASSESSMENT



**Bongza CER** Name \_\_\_\_\_ Homebase \_\_\_\_\_

During this task, you will: “**Conduct an investigation to see if the Bongza specimen is living.**”

**Here’s what you need to do:**

1. Prepare a slide with Bongza, focus it on high power and show your teacher \_\_\_\_\_
2. Complete the following **Claim, Evidence, Reasoning** framework, which satisfies the **bolded** part above, which is the point of this investigation:

**Claim** (Is Bongza living? Complete Sentence!)

**Evidence** (Sketch your inner cheek skin on high power)

Sketch components:

1. Title
2. Colour
3. Accurate sketch
4. Magnification
5. Labelled  
organelles
6. One Cell Circled  
and Identified

<b>Reasoning</b> - How does your evidence support your claim?
<b>Step 1:</b> Show what you know
<b>Step 2:</b> Connect the evidence to the claim. How does your sketch show that <u>Bongza</u> is living or not?
<b>Step 3:</b> Closing statement. Therefore.....

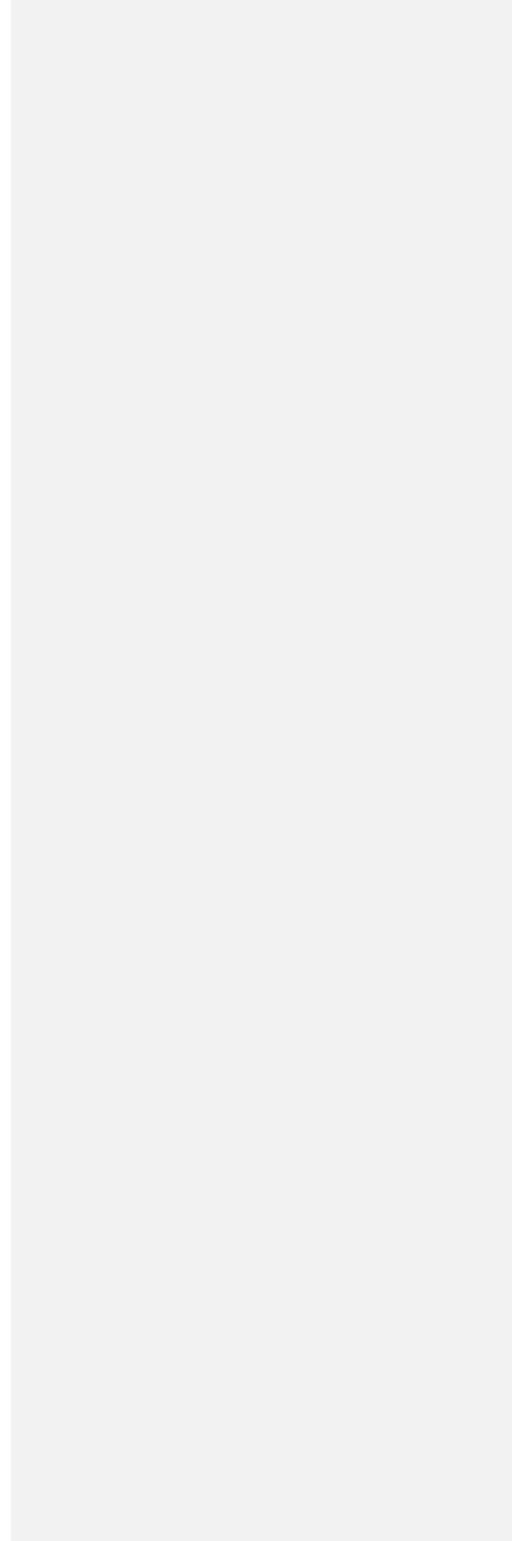
SELF-ASSESS using checkmarks in pencil on the rubric.

<b>Knowledge and Understanding</b>	Always (7) <b>Exemplary</b>	Usually (5) <b>Proficient</b>	Somewhat (3) <b>Developing</b>	Rarely (1) <b>Emerging</b>
Claim is accurate.				
Sketch (evidence) is accurate.				
Sketch (evidence) has the correct title and magnification.				
Organelles (parts) are correctly labelled.				
Overall				

<b>Transfer of Learning</b>	Always (7) <b>Exemplary</b>	Usually (5) <b>Proficient</b>	Somewhat (3) <b>Developing</b>	Rarely (1) <b>Emerging</b>
Student was able to focus their slides on the cells at high power without teacher or peer assistance.				
Student's reasoning connected the evidence to the claim.				
Student's evidence clearly identified what one cell of <u>Bongza</u> looked like.				
Overall				

APPENDIX E

CELL ANALOGY PRACTICE PERFORMANCE ASSESSMENT



Name \_\_\_\_\_

Homebase \_\_\_\_\_

### Cell Analogy PRACTICE Assessment

Your chosen analogy for a cell is: An Airplane

**Sentence Frame:**

The XYZ is/are like a cell membrane.

The function of the XYZ is...

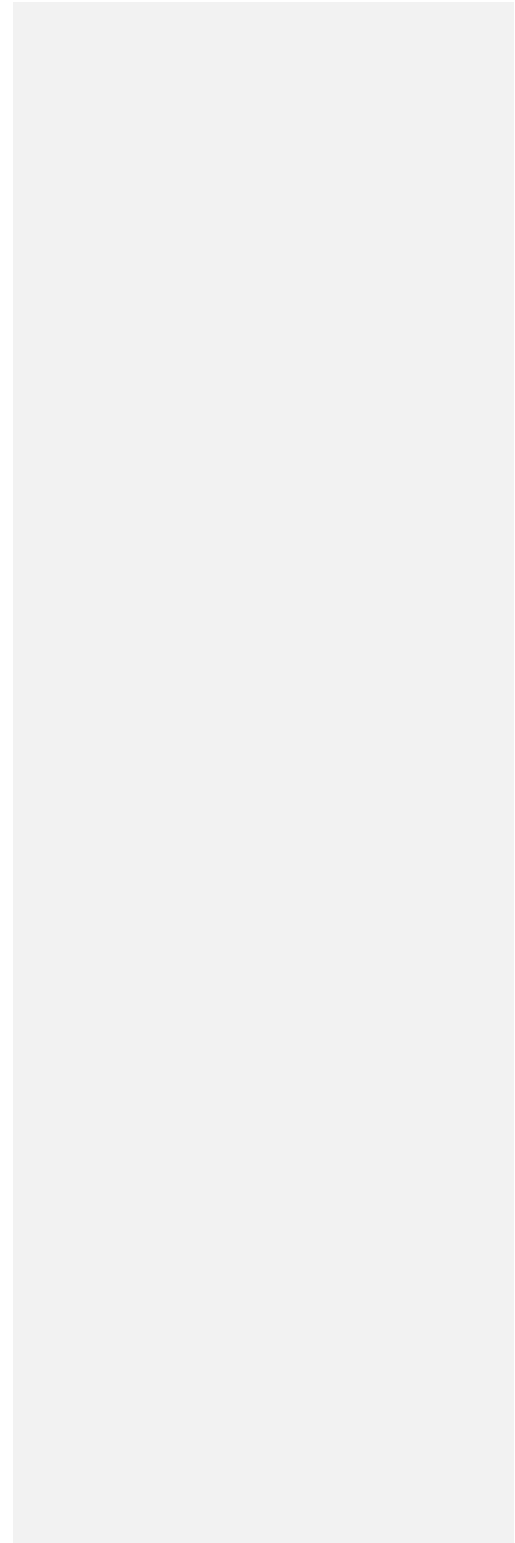
This is just like the cell membrane because...

The limitations with this analogy are ...

Use the sentence frame above to explain your choice of analogy for each of the listed organelles.

Cell Membrane

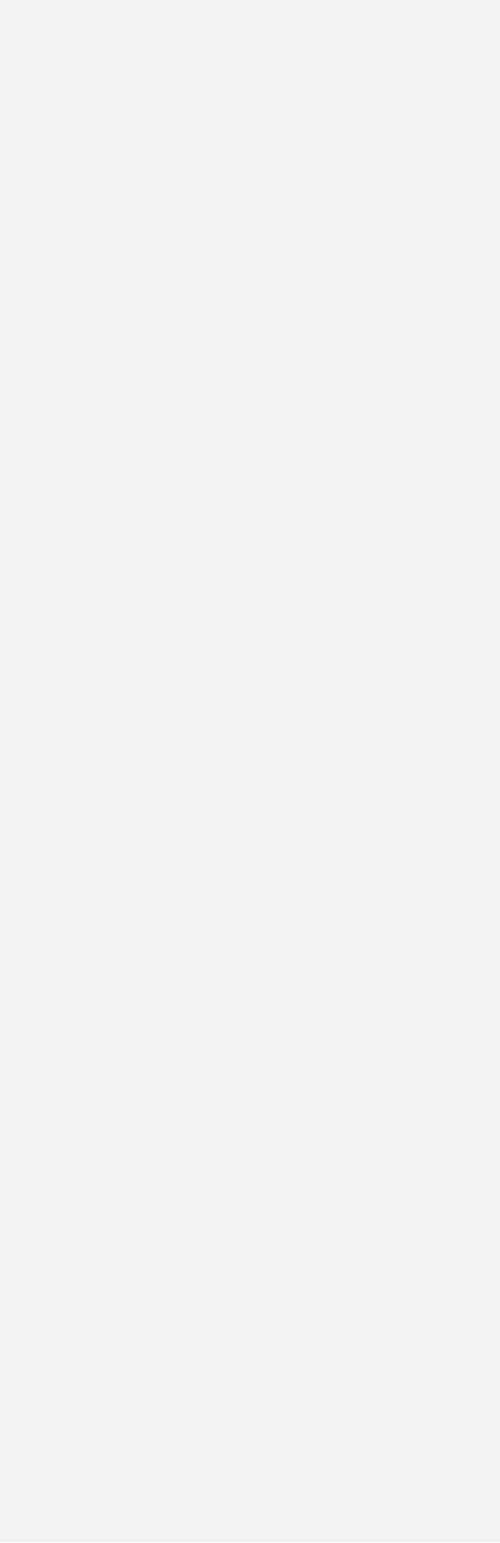

Cell Wall

The XYZ is/are like a cell membrane.  
The function of the XYZ is...  
This is just like the cell membrane because...  
The limitations with this analogy are ...

Lysosome


Mitochondria

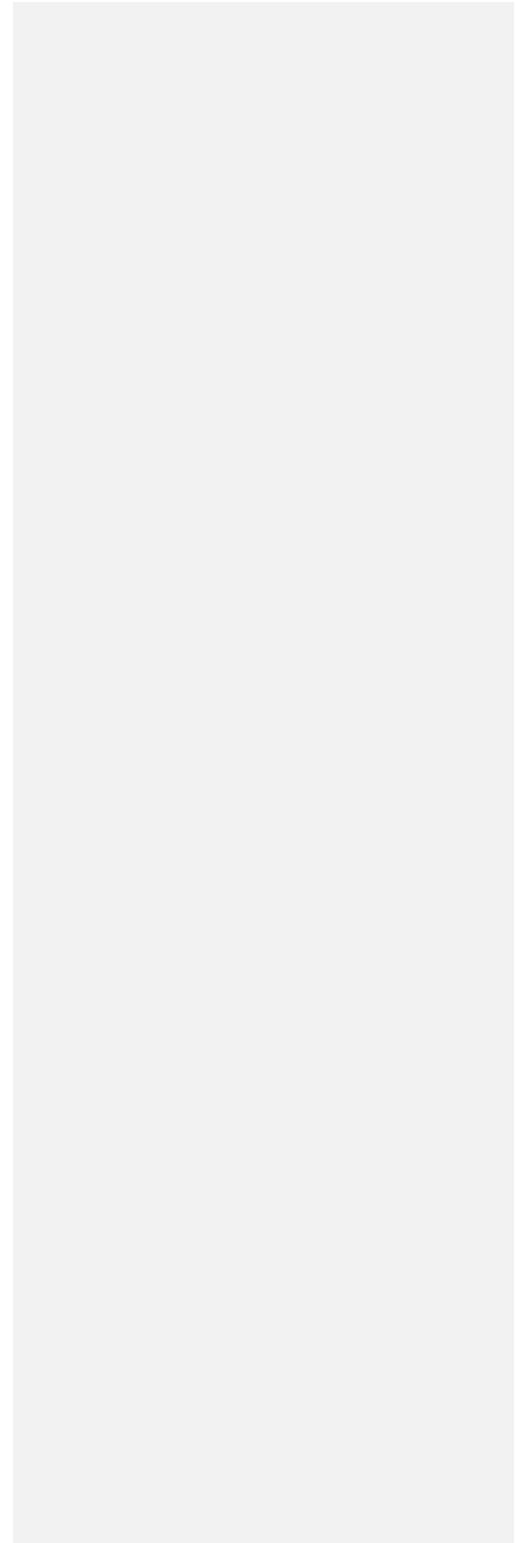



The XYZ is/are like a cell membrane.  
The function of the XYZ is...  
This is just like the cell membrane because...  
The limitations with this analogy are ...

Vacuole


Nucleus


Cytoplasm

The XYZ is/are like a cell membrane.  
 The function of the XYZ is...  
 This is just like the cell membrane because...  
 The limitations with this analogy are ...

Chloroplast

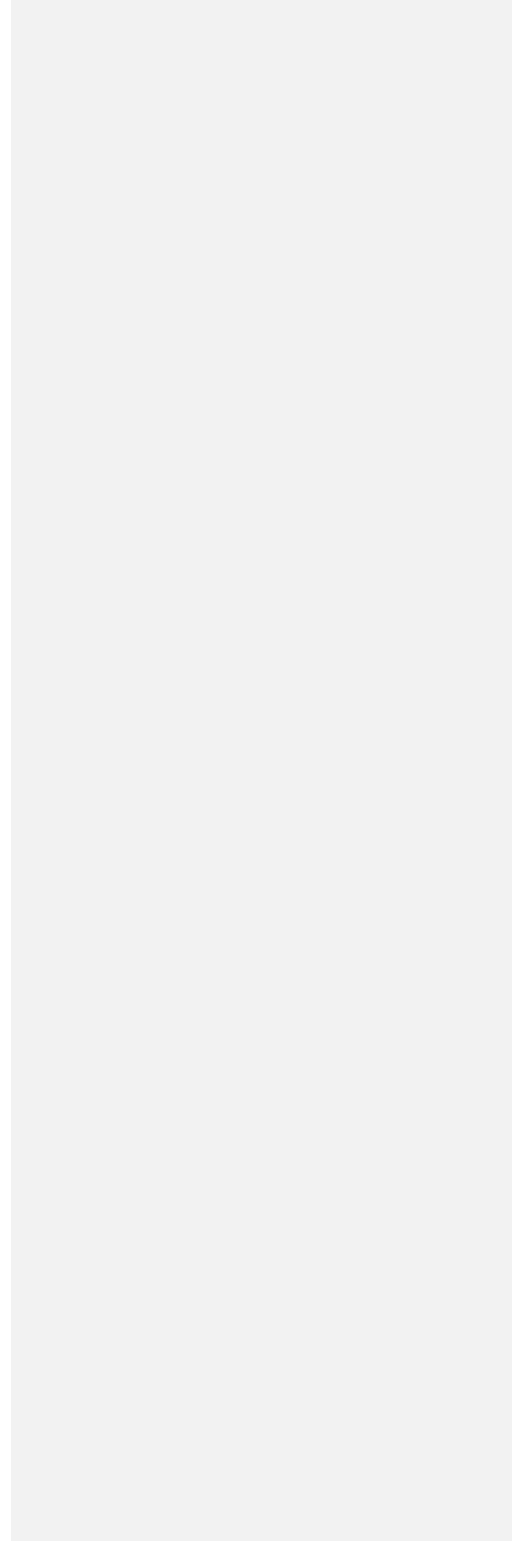

Knowledge and Understanding	Always (7) Exemplary	Usually (5) Proficient	Somewhat (3) Developing	Rarely (1) Emerging
Student's knowledge of the functions of each cell organelle is complete and accurate.				
Student's explanation demonstrates an understanding of the role the organelle plays in the functioning of the cell.				
OVERALL				

Transfer of Learning	Always (7) Exemplary	Usually (5) Proficient	Somewhat (3) Developing	Rarely (1) Emerging
Student chooses plausible corresponding parts in their analogy which have a similar function to the organelle.	In at least 1 instances, more than one part is chosen.			
Student explains how the function of their analogous part(s) is/ore similar to that of the organelle.	In at least 1 instance a limitation to their analogies is explained.			
OVERALL				



APPENDIX F

CELL ANALOGY SUMMATIVE PERFORMANCE ASSESSMENT





---

**Sentence Frame:**

The XYZ is/are like a cell membrane.

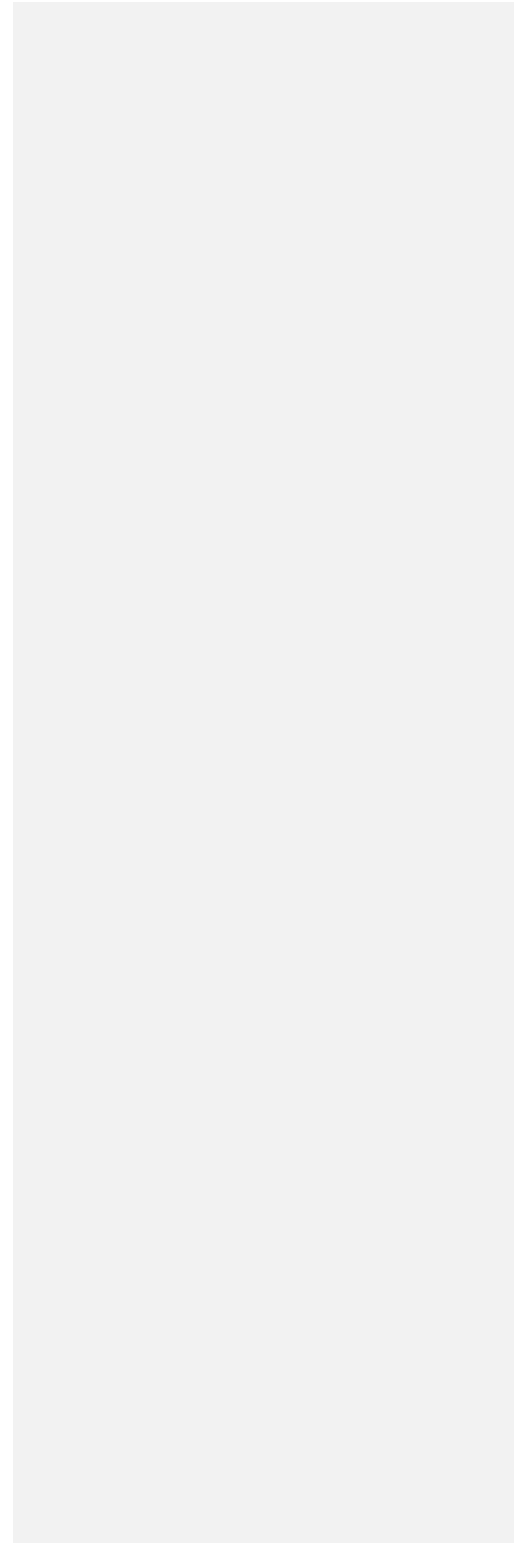
The function of the XYZ is \_\_\_\_\_

This is just like the cell membrane because \_\_\_\_\_

The limitations with this analogy are \_\_\_\_\_

**Lysosome**


**Mitochondria**

**Sentence Frame:**

The XYZ is/are like a cell membrane.

The function of the XYZ is \_\_\_\_\_

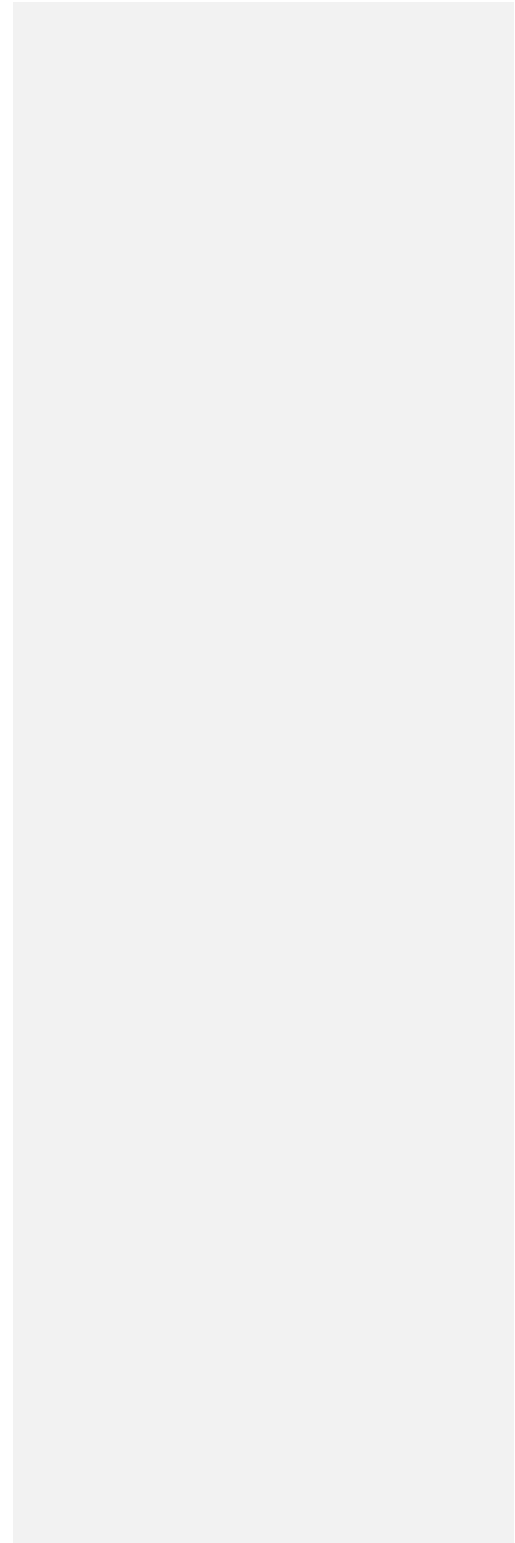
This is just like the cell membrane because \_\_\_\_\_

The limitations with this analogy are \_\_\_\_\_

**Vacuole**


**Nucleus**


**Cytoplasm**

**Sentence Frame:**

The XYZ is/are like a *cell membrane*.

The function of the XYZ is \_\_\_\_\_.

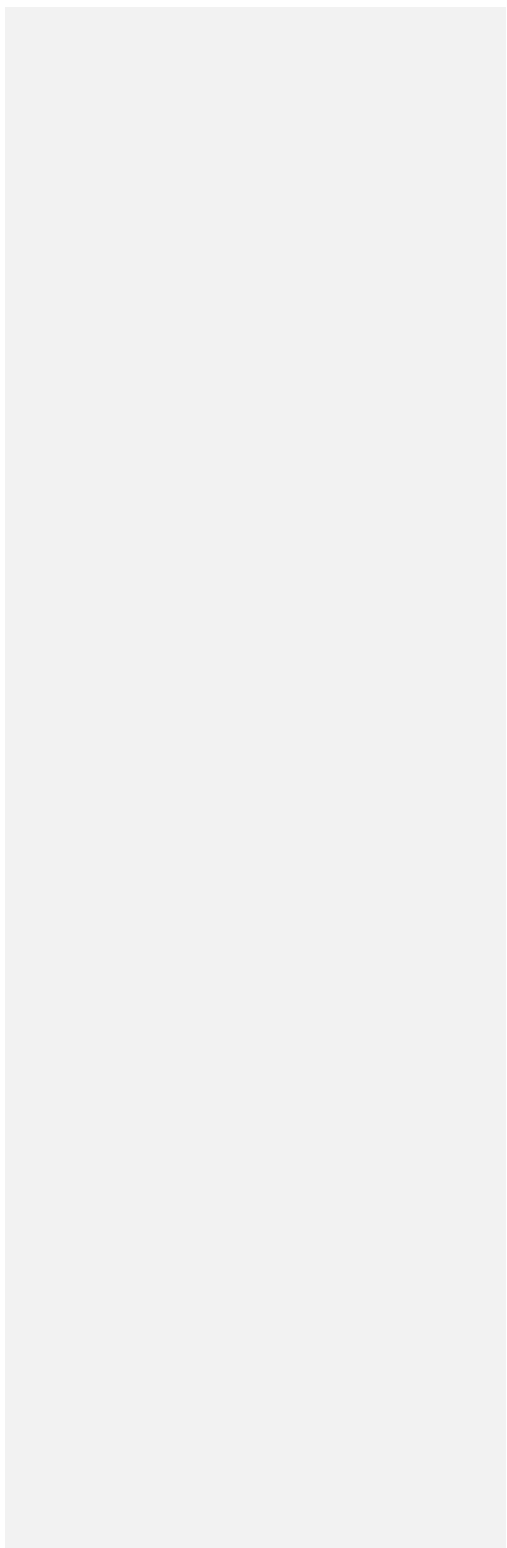
This is just like the *cell membrane* because \_\_\_\_\_.

The limitations with this analogy are \_\_\_\_\_.

Chloroplast

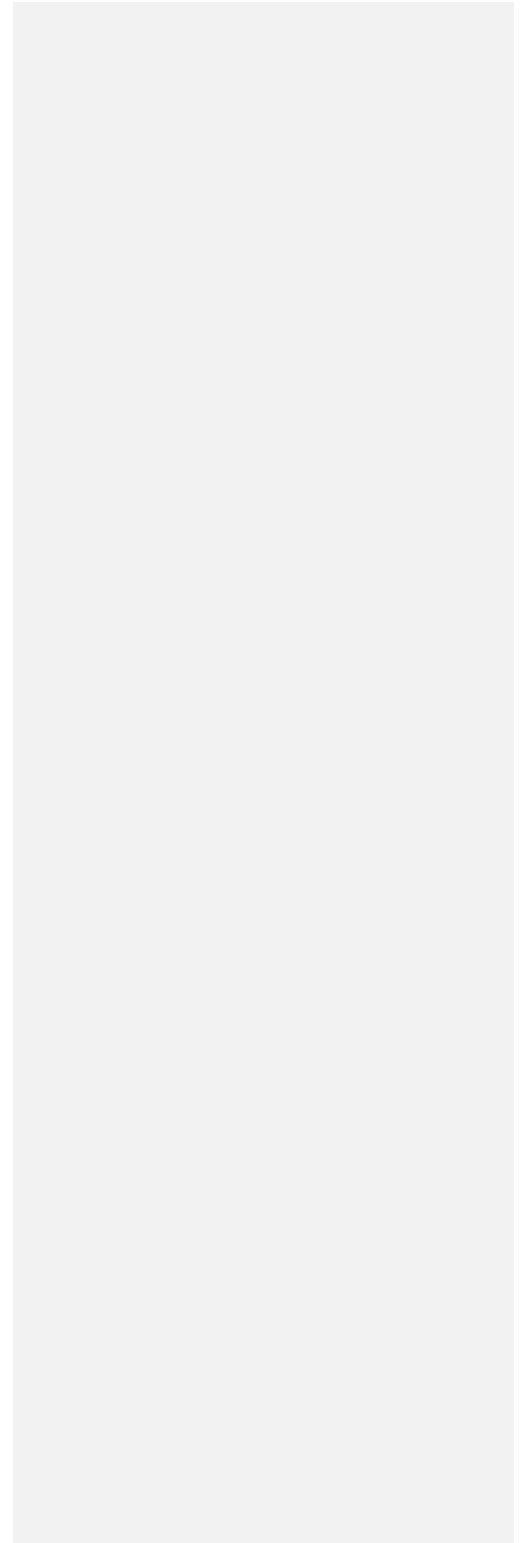

Knowledge and Understanding	Always (7) Exemplary	Usually (5) Proficient	Somewhat (3) Developing	Rarely (1) Emerging
Student's knowledge of the functions of each cell organelle is complete and accurate.				
Student's explanation demonstrates an understanding of the role the organelle plays in the functioning of the cell.				
<b>OVERALL</b>				

Transfer of Learning	Always (7) Exemplary	Usually (5) Proficient	Somewhat (3) Developing	Rarely (1) Emerging
Student chooses plausible corresponding parts in their analogy which have a similar function to the organelle.	In at least 1 instances, more than one part is chosen.			
Student explains how the function of their analogous part(s) is/are similar to that of the organelle.	In at least 1 instance a limitation to their analogies is explained.			
<b>OVERALL</b>				



APPENDIX G

BODY SYSTEMS PRACTICE PERFORMANCE ASSESSMENT



**Body Systems Practice Assessment** Name \_\_\_\_\_ Homebase \_\_\_\_\_

**Claim:** What body systems are working together to make the evidence occur? (Complete sentence)

---

---

**Evidence:** In the morning before eating, the amount of sugar in Ms. vB's blood is measured at 100 mg/dL.  
Two hours after eating breakfast, the amount of sugar in her blood is measured at 140 mg/dL.

<b>Reasoning</b> - How does your evidence support your claim?
<b>Step 1:</b> Show what you know.
<b>Step 2:</b> How does your evidence show that these two body systems are working together?
<b>Step 3:</b> Describe how and why these two body systems are working together to make the evidence occur. Draw a model to support your explanation.
<b>Step 4:</b> Closing statement. Therefore.....

<b>Transfer of Learning</b>	<b>Always (7) Exemplary</b>	<b>Usually (5) Proficient</b>	<b>Somewhat (3) Developing</b>	<b>Rarely (1) Emerging</b>
Student correctly <u>identifies</u> the interacting body systems.				
Student correctly <u>explains</u> how the body systems interact to cause the evidence to occur.				
Student's model is correct and supports their explanation.				
Student connects evidence back to their claim				

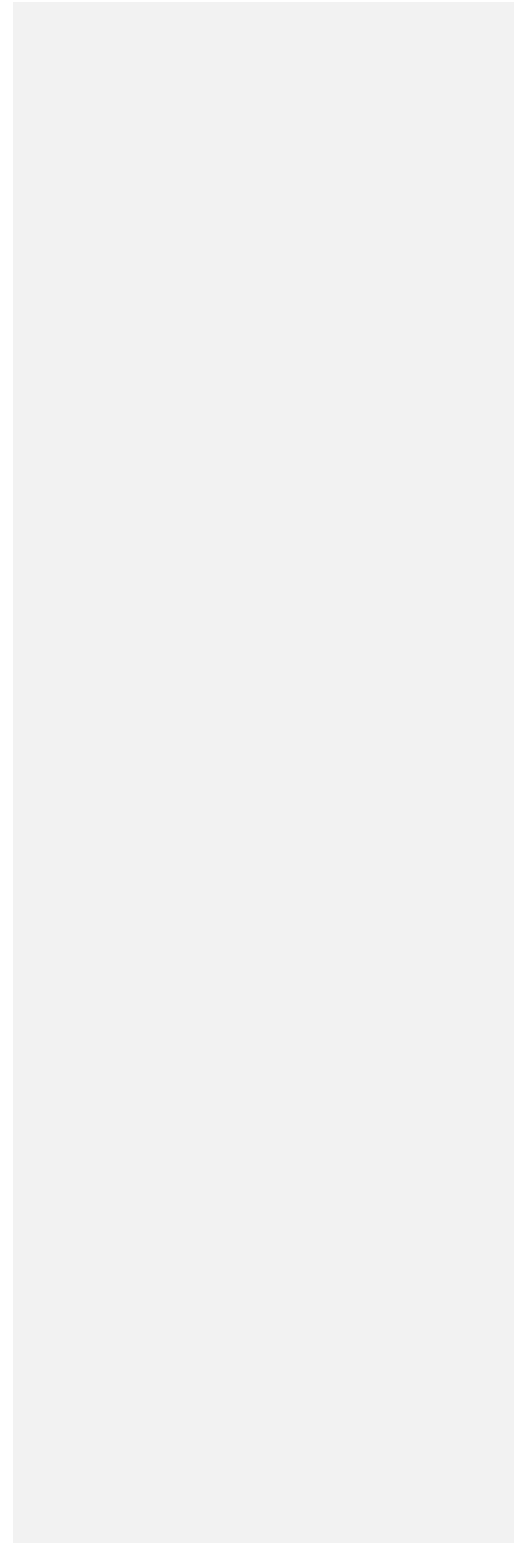
2a )	Why does our body have veins and arteries? _____ _____ _____
b)	Why do our cells have a nucleus? _____ _____ _____
c)	So, why do organisms have smaller parts? (Use other examples to support your answer) _____ _____ _____ _____ _____

3a)	What is the relationship between the heart's muscle and the function of the heart? _____ _____ _____
b)	What is the relationship between the pores of the cell membrane and its function? _____ _____ _____
c)	So, what is the relationship between the structure of our body's parts and their function? (Use other examples to support your answer) _____ _____ _____ _____ _____



APPENDIX H

BODY SYSTEMS SUMMATIVE PERFORMANCE ASSESSMENT



**Body Systems Summative Assessment** Name \_\_\_\_\_ Homebase \_\_\_\_\_

**Claim:** What body systems are working together to make the evidence occur? (Complete sentence)

\_\_\_\_\_

\_\_\_\_\_

**Evidence:** Before smoking, the amount of nicotine in a Smoker Sally's blood was measured at 10 ng/mL.  
After smoking two cigarettes, the amount of nicotine in her blood was measured at 500 ng/mL.

<b>Reasoning</b> - How does your evidence support your claim?
<b>Step 1:</b> Show what you know.
<b>Step 2:</b> How does your evidence show that these two body systems are working together?
<b>Step 3:</b> Describe how and why these two body systems are working together to make the evidence occur. Draw a model to support your explanation.
<b>Step 4:</b> Closing statement. Therefore.....

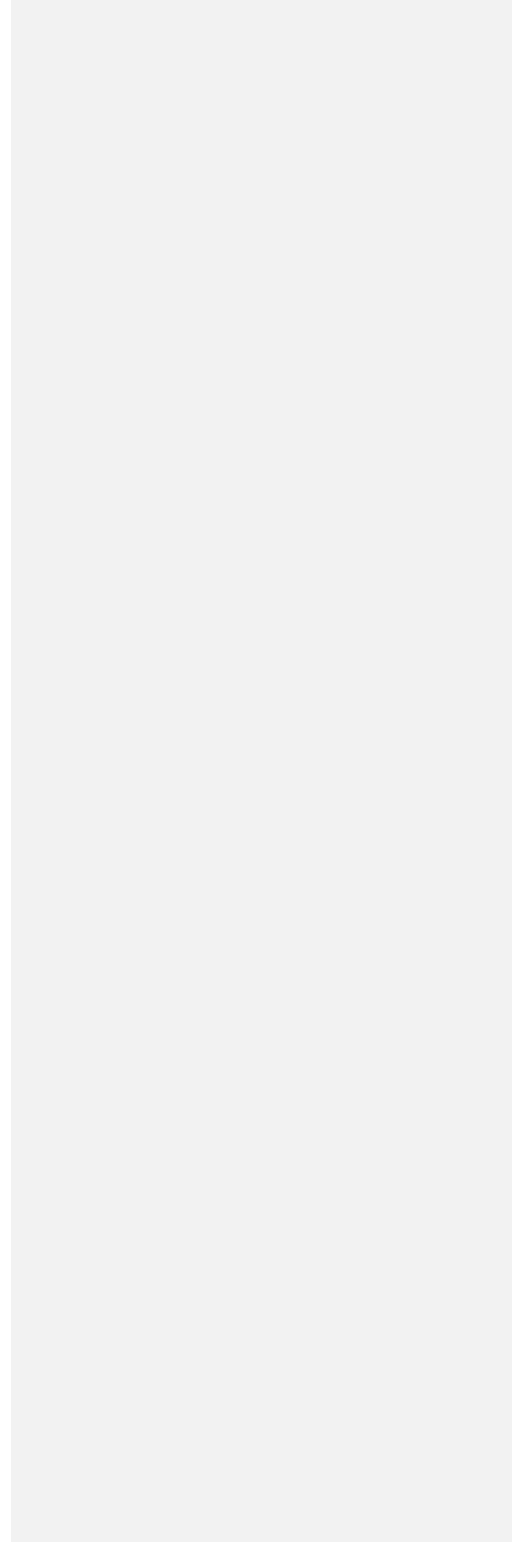
<b>Transfer of Learning</b>	<b>Always (7) Exemplary</b>	<b>Usually (5) Proficient</b>	<b>Somewhat (3) Developing</b>	<b>Rarely (1) Emerging</b>
Student correctly identifies the interacting body systems.				
Student correctly explains how the body systems interact to cause the evidence to occur.				
Student's model is correct and supports their explanation.				
Student connects evidence back to their claim				

2a)	Why does our body have lungs? _____ _____ _____
b)	Why do our cells have vacuoles? _____ _____ _____
c)	So, why do organisms have smaller parts? (Use other examples not from this or from your practice assessment to support your answer) _____ _____ _____ _____ _____


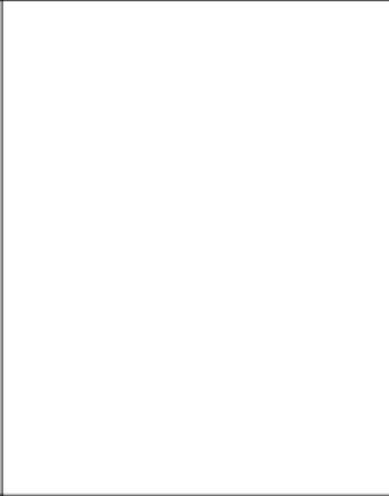
3a)	What is the relationship between the size of the capillaries and their function? _____ _____ _____
b)	What is the relationship between enzymes/bile/acids and the function of the digestive system? _____ _____ _____
c)	So, what is the relationship between the structure of/in our body's parts and their function? (Use other examples not from this or from your practice assessment to support your answer) _____ _____ _____ _____ _____

APPENDIX I

MATTER PRACTICE PERFORMANCE ASSESSMENT



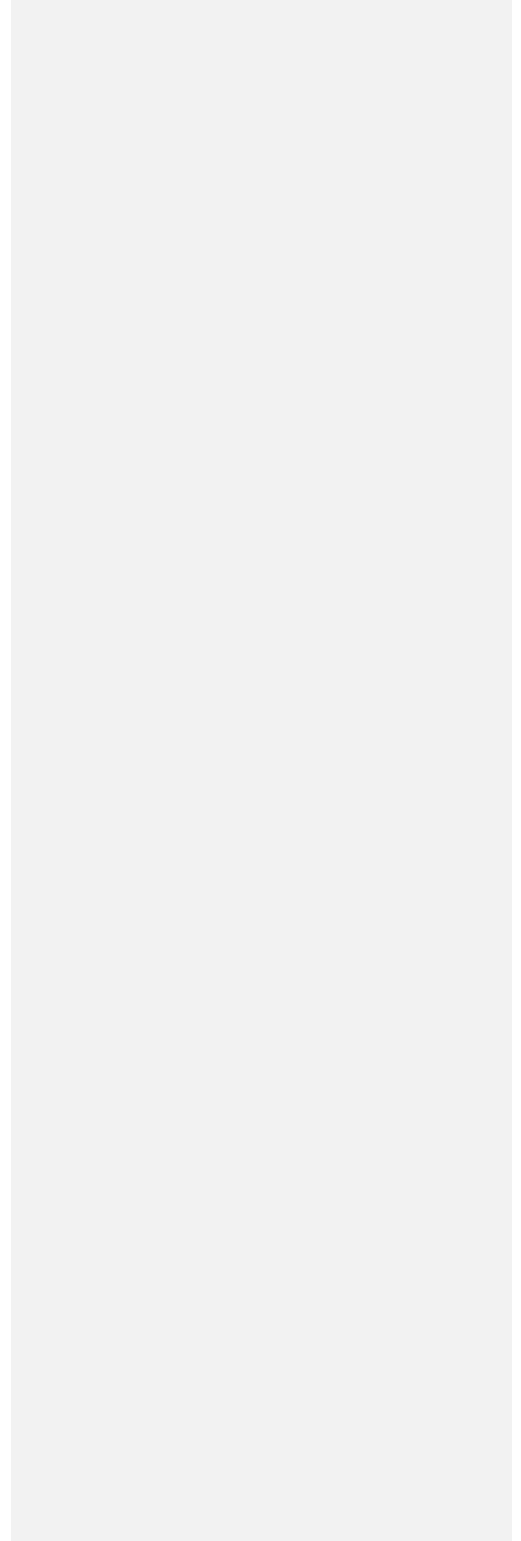


<b>Reasoning</b> - How does the evidence and theory support your claim?	
Step 2: How does your evidence support your claim?	
Step 3: Describe what was happening to the solid molecules in the ball before and after the thermal energy was changed. <input type="checkbox"/> Draw a before and after model to support your explanation.	
Before:	After:
	
Step 4: Closing statement. Therefore.....	

Knowledge and Understanding	Always (7) Exemplary	Usually (5) Proficient	Somewhat (3) Developing	Rarely (1) Emerging
Student correctly identifies the effect that thermal energy changes have on solids.				
Student's evidence shows what happened to the ball. Evidence is labelled.				
Student's models are correct and supports their explanation.				
Student connects the evidence back to their claim				

APPENDIX J

MATTER SUMMATIVE PERFORMANCE ASSESSMENT







**Matter and Thermal Energy Changes Assessment** Name \_\_\_\_\_**Claim:** What effect does thermal energy change have on solids and gases? (Complete sentence based on evidence)

As the thermal energy of the solids and gases in the popcorn bag \_\_\_\_\_

\_\_\_\_\_

**Evidence:**

Before microwaving	After microwaving
<p data-bbox="233 672 373 703">Outside the Bag</p>  <p data-bbox="233 882 373 913">Inside the Bag:</p> 	<p data-bbox="625 672 763 703">Outside the Bag</p>  <p data-bbox="625 882 763 913">Inside the Bag:</p> 
<p data-bbox="233 1113 649 1144">Describe the changes that you observe in the evidence:</p> <p data-bbox="233 1165 1045 1186">_____</p> <p data-bbox="233 1207 1045 1228">_____</p> <p data-bbox="233 1249 1045 1270">_____</p>	

**Reasoning** - How does the evidence and theory support your claim?**Step 1:** Show what you know about solids and gases and how they are affected by thermal energy increases.

\_\_\_\_\_

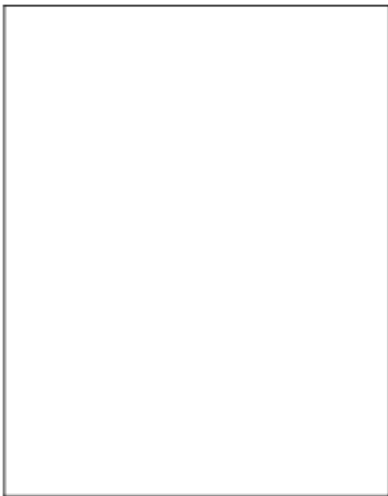
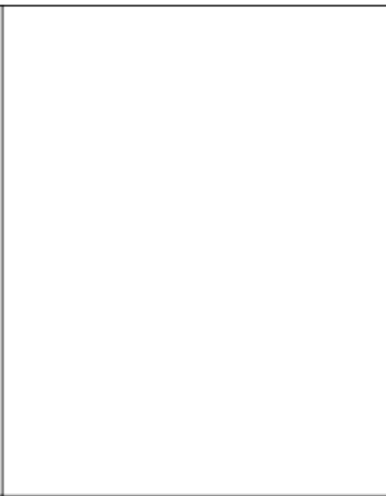
\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

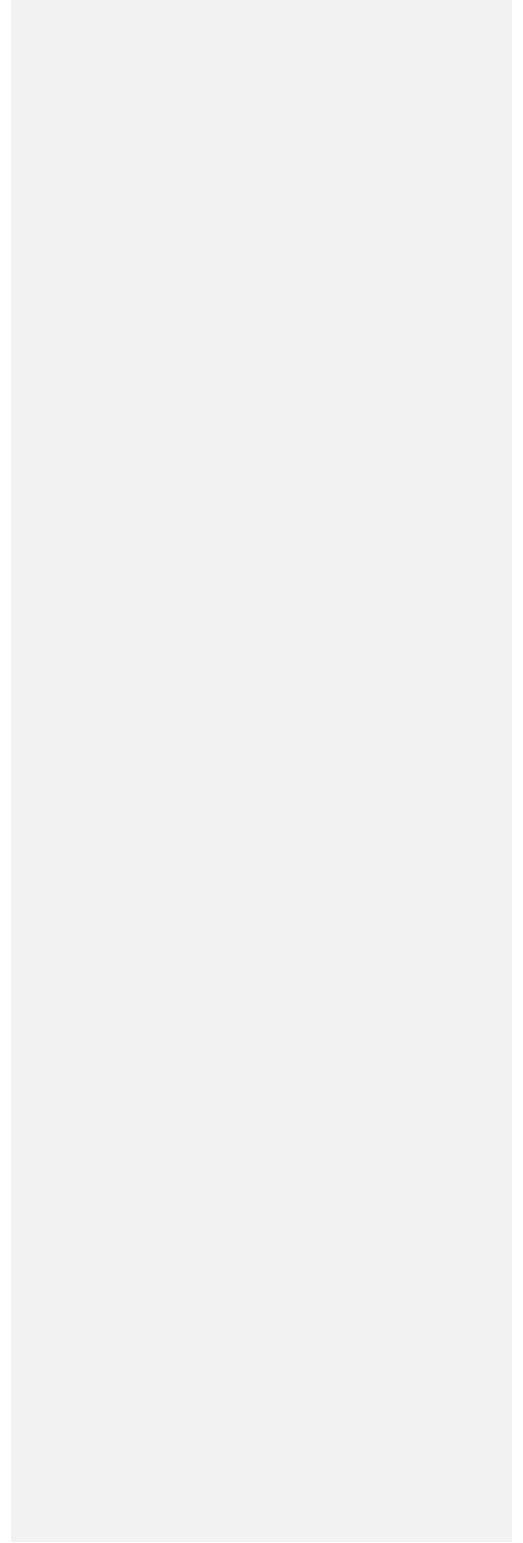


<b>Reasoning -</b> How does the evidence and theory support your claim?	
Step 2: How does your evidence support your claim?	
Step 3: Describe what was happening to the gas and solids in the bag of microwave popcorn before and after the thermal energy was changed. Draw a before and after model to support your explanation.	
Before:	After:
	
Step 4: Closing statement. Therefore.....	

<b>Knowledge and Understanding</b>	<b>Always (7) Exemplary</b>	<b>Usually (5) Proficient</b>	<b>Somewhat (3) Developing</b>	<b>Rarely (1) Emerging</b>
Student correctly identifies the effect that thermal energy changes have on solids and gases.				
Student's evidence describes the changes seen.				
Student's models are correct and supports their explanation.				
Student connects the evidence back to their claim				
<b>OVERALL</b>				

APPENDIX K

CELLS INVESTIGATION GOOGLE SURVEY QUESTIONS



## "Living or Non-Living" Assessment Survey

Participation in this research is voluntary and participation or non-participation will not affect a student's grades or class standing in any way.

\* Required

I am comfortable with the content and practices (skills) of related to this assessment. \*

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
I can prepare my own microscope slide and focus it accurately on low or medium power.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I can prepare my own microscope slide and focus it accurately on high power.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I can identify whether a microscopic specimen is non-living, plant or animal.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I can identify the parts of cells (nucleus, cell wall, cell membrane, chloroplasts).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I can identify one plant cell.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I can identify one animal cell.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

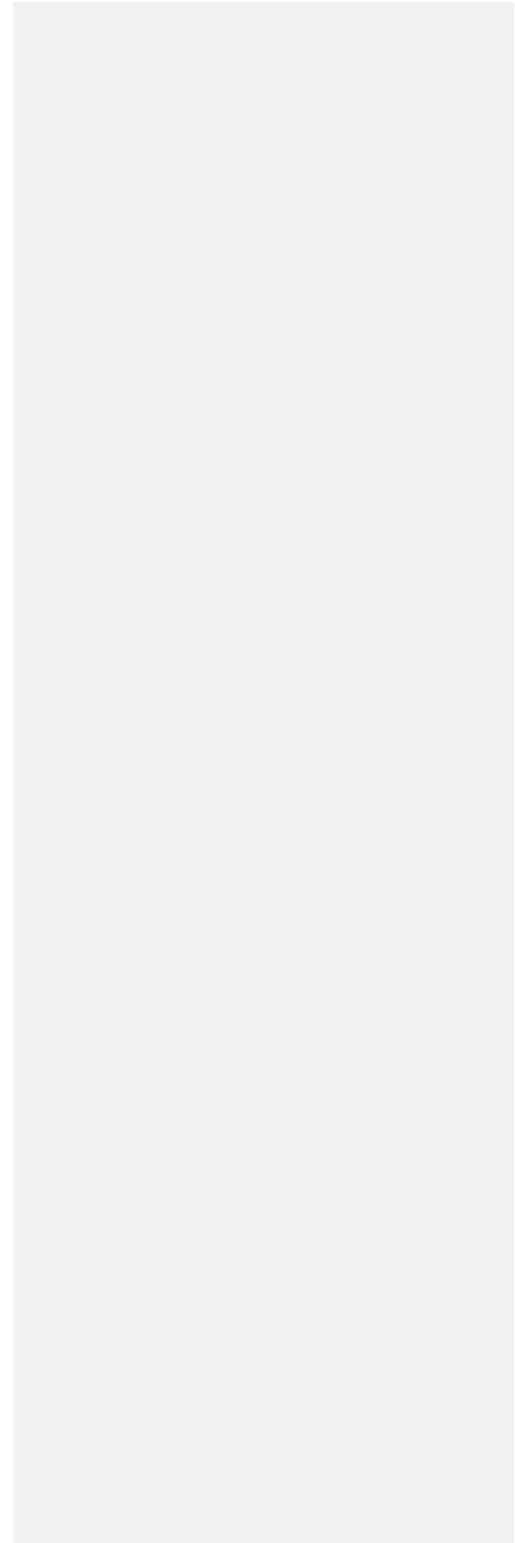
	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
I am confident in my ability to write a scientific claim.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am confident in my ability to write reasoning that supports a claim about whether a specimen is living or non-living.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am confident in my ability to write a CER for science class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am confident in my ability to write a CER for another class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am confident in how I did on today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I felt well prepared for today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The "Inner Cheek Cells" CER practice assessment was good preparation for today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The final review class was good preparation for today's Bongza	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
The answer key for the "Inner Cheek Cells" practice assessment helped me learn from my mistakes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Getting graded for each part of the "Inner Cheek Cells" practice assessment helped me improve for the today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would prefer if Mr. Freeman grades our practice assessments, as long as they don't count towards our PowerSchool overall grade.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Correcting my "Inner Cheek Cells" practice CER assessment helped me do better on today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Doing the extra "Onion Skin" CER investigation helped me prepare for today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
I put in an adequate amount of preparation for today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was focused and put forth my best effort during the 'Inner Cheek Cell' practice investigation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was focused and put forth my best effort during the class before today's Bongza assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was focused and put forth my best effort during the microscope and cells part of this unit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I think that it was appropriate that today's Bongza assessment was open notes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How could Mr. Freeman have prepared you better for today's Bongza assessment?

Your answer \_\_\_\_\_



What could you have done (during the unit, lessons tutorials or at home) in order to have done better on today's Bongza assessment?

Your answer \_\_\_\_\_

What would you change about the unit so far?

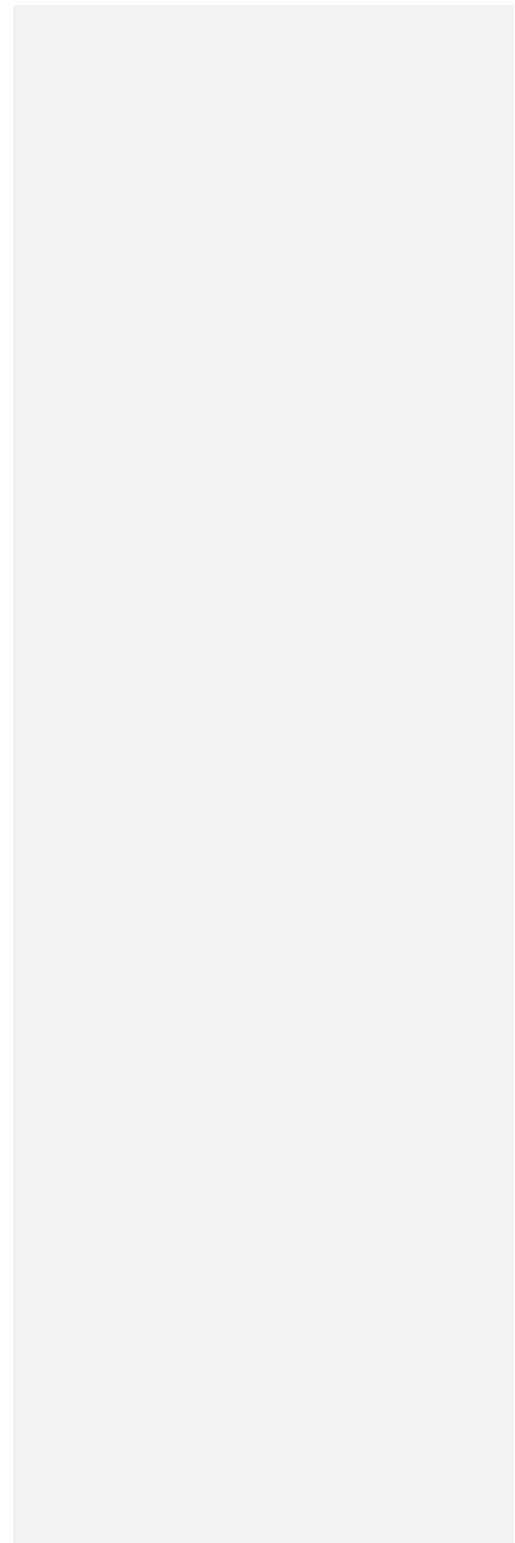
Your answer \_\_\_\_\_

**Submit**

Never submit passwords through Google Forms.

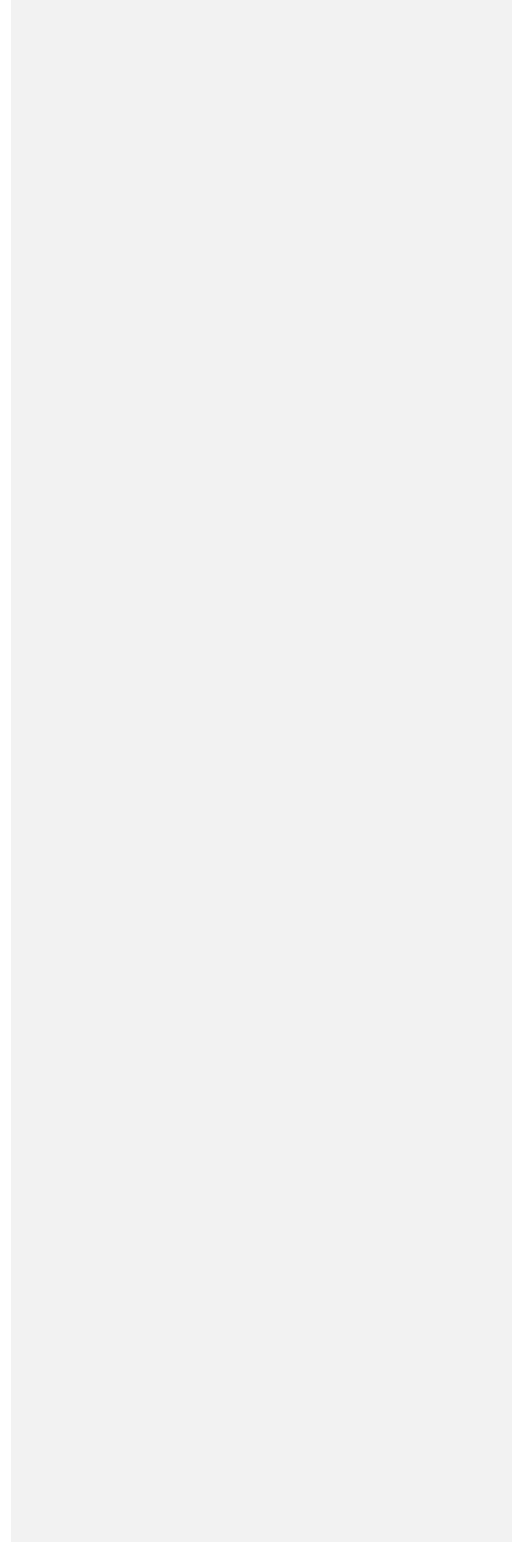
This form was created inside of International School Manila. [Report Abuse](#)

Google Forms



APPENDIX L

CELL ANALOGY GOOGLE SURVEY QUESTIONS





## "Cell Analogy" Assessment Survey

Participation in this research is voluntary and participation or non-participation will not affect a student's grades or class standing in any way.

Your email address ([freemann@ismanila.org](mailto:freemann@ismanila.org)) will be recorded when you submit this form. Not **freemann**?

[Sign out](#)

\* Required

### 1. Please indicate your level of agreement with each of the following statements. \*

*Check all that apply.*

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
I am confident in my knowledge of the structure of the 8 organelles that we learned in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am confident in my knowledge of the function of the 8 organelles that we learned in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Comparing my Airplane analogy to the exemplar prepared me for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The compare and contrast brainframe (my Airplane assessment versus the Exemplar) helped prepare me for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Asking Mr. Freeman for personalized clarification or feedback during the final review class helped prepare me for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Doing analogies (ISM and Airplane) helped me learn about the structure and function of the organelles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The homework helped me learn about the structure and function of the organelles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher directed notes that we took in class helped me learn about the structure and function of the organelles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The Kahoots helped me learn about the structure and function of the organelles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The fact that today's assessment was open notes helped me do well on today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am confident in how I did on today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I felt well prepared for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The "Airplane" practice assessment was good preparation for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The final review class was good preparation for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The exemplar for the "Airplane" practice assessment helped me learn from my mistakes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
Getting graded for each part of the "Airplane" practice assessment helped me improve for the today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would prefer if Mr. Freeman grades our practice assessments, as long as they don't count towards our PowerSchool overall grade.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Correcting my "Airplane" practice assessment helped me do better on today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Doing an extra (not Airplane) analogy was good preparation for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I put in an adequate amount of preparation for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was focused and put forth my best effort during the "Airplane" practice assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was focused and put forth my best effort during the class before today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was focused and put forth my best effort during the organelles structure and function part of this unit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I think that it was appropriate that today's Cell Analogy assessment was open notes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The "ISM as a Cell" Goosechase was good preparation for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Revising the "ISM as a Cell" Goosechase was good preparation for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Science assessments should be open notes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Doing a practice assessment followed by a final review class is a good way to prepare for actual assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would have done well on this assessment if it weren't open notes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The analogy choices on today's practice assessment were fair.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The exemplar prepared me well for today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
It was fair that we were not allowed to use a copy of the exemplar during today's Cell Analogy assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
I would have done better if we'd been allowed to use the exemplar during today's assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. How could Mr. Freeman have prepared you better for today's Cell analogy assessment?

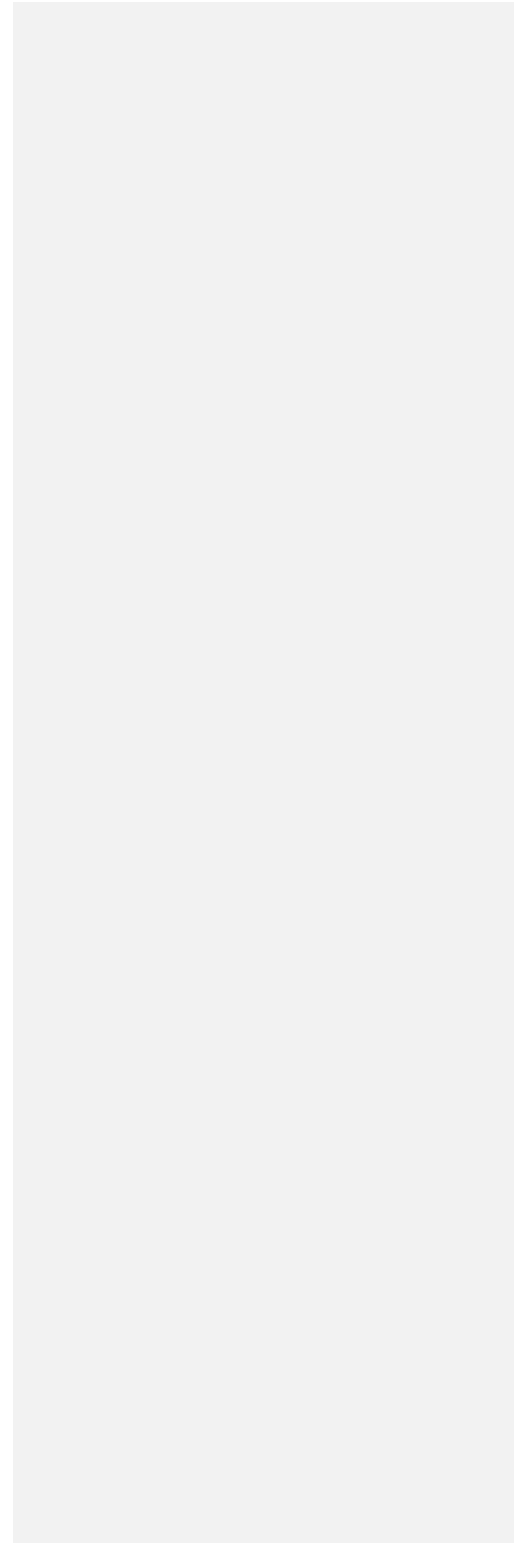
---

3. What could you have done (during the unit, lessons tutorials or at home) in order to have done better on today's Cell analogy assessment?

---

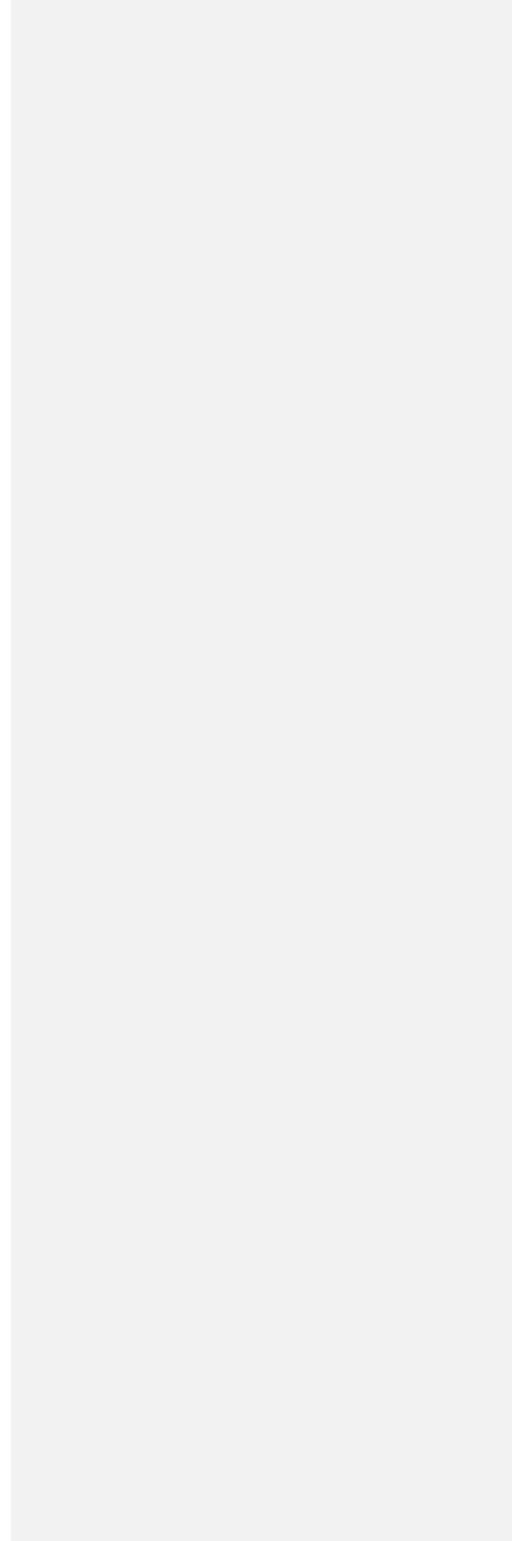
4. What would you change about the unit so far?

---



APPENDIX M

BODY SYSTEMS GOOGLE SURVEY QUESTIONS



### "Body Systems" Assessment Survey

Participation in this research is voluntary and participation or non-participation will not affect a student's grades or class standing in any way.

Your email address ([freemann@lsmanila.org](mailto:freemann@lsmanila.org)) will be recorded when you submit this form. Not you? [Switch account](#)

\* Required

I am confident in my knowledge of the circulatory, respiratory, and digestive systems. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I am confident in my ability to write an evidence-supported Body Systems CER. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I know why organisms are made of smaller parts. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I know the general relationship between the structure of part(s) of an organism and the function of that part(s). \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I am confident in how I did on the Body Systems final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Doing the Body Systems practice assessment helped me learn about the interaction of the body systems. \*

- Strongly Disagree
- Disagree
- Neutral (Didn't do the extra practice assessment)
- Agree
- Strongly Agree

Comparing my Body Systems practice assessment to the exemplar helped me prepare for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The exemplar for the Body Systems practice assessment helped me learn from my mistakes. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The Body Systems practice assessment was good preparation for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Getting graded for each part of the Body Systems practice assessment helped me improve for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Asking Mr. Freeman for personalized clarification or feedback during the final review class helped prepare me for the final Body Systems assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I would prefer if Mr. Freeman grades our practice assessments, as long as they don't count towards our PowerSchool overall grade. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I felt well prepared for the Body Systems final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The exemplar prepared me well for the Body Systems final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Correcting my Body Systems practice assessment helped me do better on the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The final review class was good preparation for the Body Systems final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Doing a practice assessment followed by a final review class is a good way to prepare for final assessments. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The Body Systems extra practice assessment and exemplar were good preparation for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral (Didn't do the extra practice assessment)
- Agree
- Strongly Agree

How could Mr. Freeman have prepared you better for the Body Systems final assessment?

Your answer \_\_\_\_\_

Submit

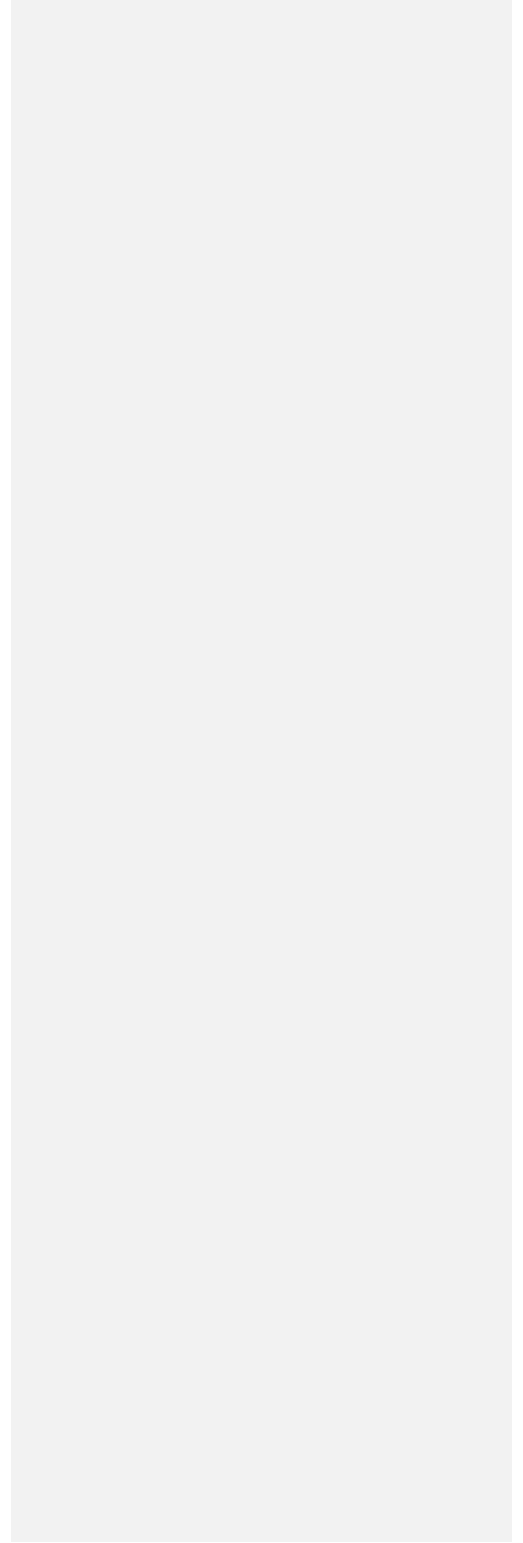
Never submit passwords through Google Forms.

This form was created inside of International School Manila. [Report Abuse](#)

Google Forms



APPENDIX N  
MATTER GOOGLE SURVEY QUESTIONS



## "Effect of Thermal Energy Changes on Solids, Liquids and Gases" Assessment Survey !

Participation in this research is voluntary and participation or non-participation will not affect a student's grades or class standing in any way.

This form is automatically collecting email addresses for International School Manila users. [Change settings](#)

I am confident in my knowledge of solids, liquids and gases. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I am confident in my knowledge of how changes in thermal energy affect solids, liquids and gases. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I am confident in my ability to write an evidence-supported CER about the effect that changes in thermal energy have on Solids, Liquids and/or Gases. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I am confident in how I did on the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Doing the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" practice assessment \* helped me learn about the effect that changes in thermal energy have on solids.

- Strongly Disagree
- Disagree
- Neutral (Didn't do the extra practice assessment)
- Agree
- Strongly Agree

Comparing my "Effect of Thermal Energy Changes on Solids, Liquids and Gases" practice assessment to the exemplar helped me prepare for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The exemplar for the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" practice assessment helped me learn from my mistakes. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

111

The "Effect of Thermal Energy Changes on Solids, Liquids and Gases" practice assessment was good preparation for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Getting graded for each part of the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" practice assessment helped me improve for the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Asking Mr. Freeman for personalized clarification or feedback during the final review class helped prepare me for the final "Effect of Thermal Energy Changes on Solids, Liquids and Gases" assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I would prefer if Mr. Freeman grades our practice assessments, as long as they don't count towards our PowerSchool overall grade. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I felt well prepared for the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The exemplar prepared me well for the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Correcting my "Effect of Thermal Energy Changes on Solids, Liquids and Gases" practice assessment helped me do better on the final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

The final review class was good preparation for the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" final assessment. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Doing a practice assessment followed by a final review class is a good way to prepare for final assessments. \*

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

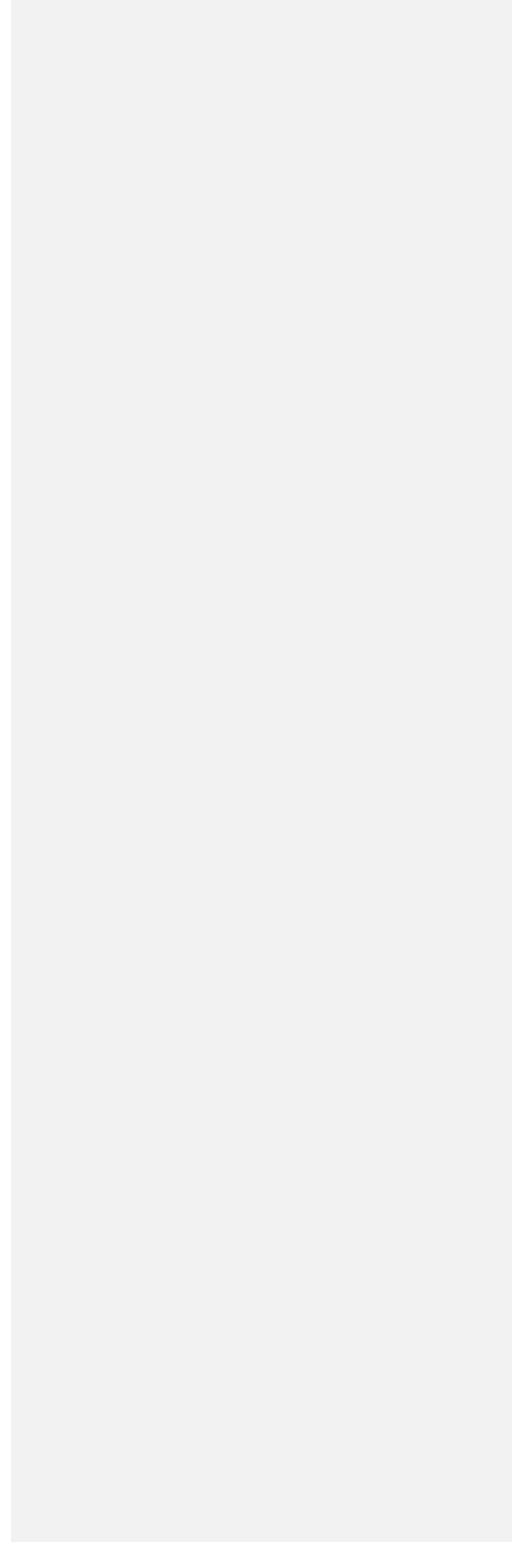
The "Effect of Thermal Energy Changes on Solids, Liquids and Gases" extra practice assessment \* and exemplar were good preparation for the final assessment.

- Strongly Disagree
- Disagree
- Neutral (Didn't do the extra practice assessment)
- Agree
- Strongly Agree

How could Mr. Freeman have prepared you better for the "Effect of Thermal Energy Changes on Solids, Liquids and Gases" final assessment?

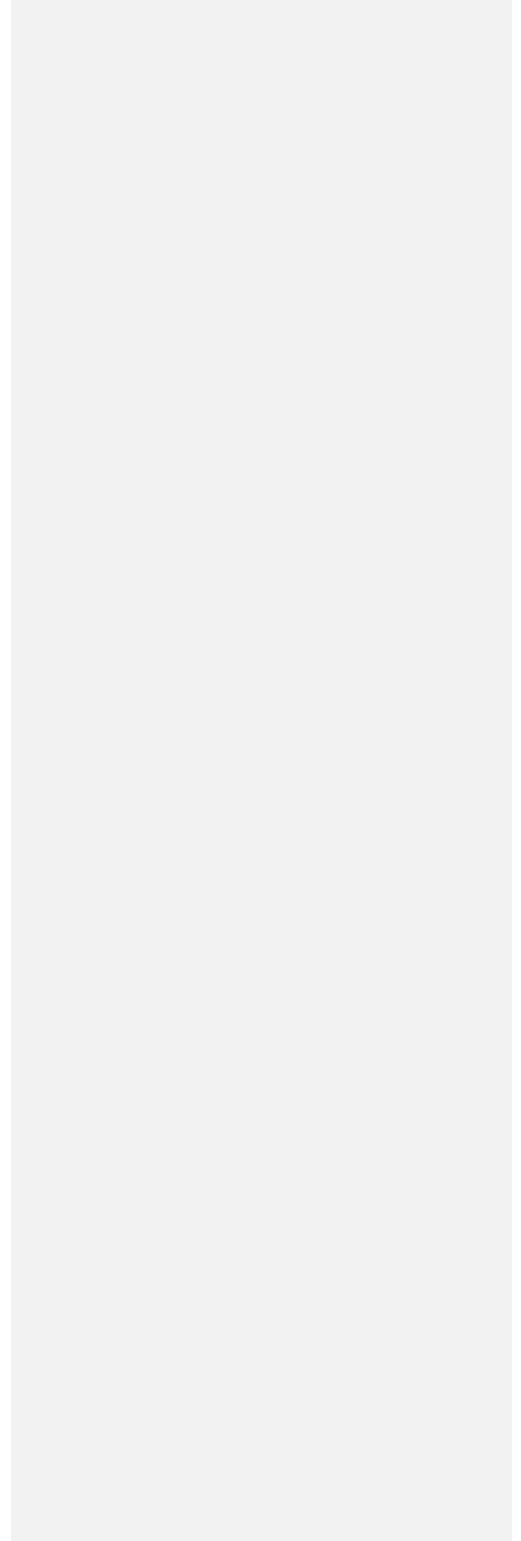
Short answer text  
.....

APPENDIX 0  
INTERVIEW QUESTIONS



How comfortable are you with the _____ part of this unit? Why?
Did you feel adequately prepared for the _____ final assessment? Why or why not?
Was the _____ practice assessment good preparation for the final assessment? Why or why not?
How did seeing your grades on the practice assessment affect your learning and achievement on the _____ final assessment? Why/
How effective was the class between the practice and summative assessment? Why?
Was the summative assessment fair, hard, too hard, correct difficulty? Why?
How could I prepare you better for future final assessments?

APPENDIX P  
STRATEGIES FOR THE FINAL REVIEW CLASS





1. Provide students with a student exemplar of the PPA and ask students to do a compare and contrast brainframe comparing their PPA to the exemplar.
2. Provide students with a student exemplar of the PPA and a non-proficient student PPA. Ask students to do a compare and contrast brainframe comparing those two PPAs.
3. Have students correct their PPA using the exemplar.
4. Reteach any concepts that are deficient in the PPAs to the whole class.
5. Have students prepare flashcards and use them to quiz themselves and/or each other.
6. Have students discuss the strategies that they used or should use in order to be successful on the PPA and/or SPA.
7. Kahoot! or Quizlet which addresses misconceptions as identified by the PPA.
8. Extra practice work which addresses deficiencies as identified by the PPA.
9. Strategy to be determined possibly with the input of my new sixth grade science teaching partner, such as the next two:
10. Code the deficiencies students had on the PPAs.
11. Teach mini-lessons to students who displayed a coded deficiency.