

THE EFFECTIVENESS OF CONCEPTUAL AND QUANTITATIVE FORMATIVE
ASSESSMENTS IN THE HIGH SCHOOL PHYSICS CURRICULUM

by

Jeffrey J. Noblejas

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

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2014

STATEMENT OF PERMISSION TO USE

In presenting this professional paper in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the MSSE program shall make it available to borrowers under rules of the program.

Jeffrey J. Noblejas

July 2014

TABLE OF CONTENTS

INTRODUCTION AND BACKGROUND	1
CONCEPTUAL FRAMEWORK.....	4
METHODOLOGY	9
DATA AND ANALYSIS.....	17
INTERPRETATION AND CONCLUSION	39
VALUE.....	47
REFERENCES CITED.....	51
APPENDICES	52
APPENDIX A: Mid-Semester Survey Questions.....	53
APPENDIX B: Pre-Treatment Survey Questions.....	57
APPENDIX C: Correspondence between Survey and Research Questions.....	60
APPENDIX D: Pre-Treatment Interview Questions	62
APPENDIX E: Correspondence between Interview and Research Questions	65
APPENDIX F: IRB Exemption Approval	68

LIST OF TABLES

Data Triangulation Matrix 13

LIST OF FIGURES

1. Conceptual-Based Question Example.....	11
2. Problem-Based Question Example	11
3. Box-whisker of Overall Exam Scores by Unit and Class Period.....	18
4. Box-whisker of Conceptual Questions by Unit and Class Period	21
5. Box-whisker of Quantitative Questions by Unit and Class Period.....	22
6. Box-whisker of FA's by Unit	24
7. Box-whisker of FA's by Class Period	25
8. Pre- and Post-survey Comparisons on the Essential Nature of FA's.....	28
9. Pre- and Post-survey Comparisons on How FA's Help Learning	30
10. Pre- and Post-survey Comparisons Conceptual and Problem-solving Type of Learning	31
11. Post-survey Questions on the Direct Effect of FA's on Learning and Tests	33

ABSTRACT

This study attempted to determine the effects of formative assessments in a flipped curriculum physics classroom. It looked at the effects on student achievement, student attitude towards learning, and the way curriculum is taught. It also attempted to determine whether the type of formative assessment, conceptual or quantitative, makes a difference in the outcome. The data showed that overall performance on the formative assessments doesn't correlate well with summative assessment performance. However, there is a possible correlation between the type of formative assessment given and performance on corresponding questions on exams. Overall, students had a positive perception of formative assessments as a tool for learning. Formative assessment were also shown to be an effective means to adjust curriculum during and after instruction.

INTRODUCTION AND BACKGROUND

Background

In my classroom I have noticed that there were many times when I felt that students understood the concepts that I taught only to be surprised in their performance in unit exams. This brought about the question of how I can better gauge student understanding in class to help them master the material. This is an important issue for teachers to address because in order to teach well we have to know what our students understand and don't. This is an important question in my school in particular because I work at a private high school which attracts high performing students. The students are well behaved and highly motivated. At the same time, however, the students and parents have high expectations when it comes to performance, grades, and learning. Having students perform well on the summative assessments is a main concern.

The motivation behind the study is the realization of the significance of teaching guided by feedback and data. The realization came about through an Education, Curriculum, and Instruction (EDCI) course I took last year on evaluations and assessments. Also, the formal observations I have done of other teachers has shown that most, if not all, teachers in my school don't have a formal system of formative assessment in place. Most of the formative assessments are informal in nature. An example would be verbally asking students questions during the course of a lesson. This speaks to the broader significance of my study. If I do succeed in creating a formative assessment system that is shown to be effective, it can serve as an example for other teachers to adopt and adapt in their own classes. Also, I would hope that my study will

be able to contribute, even if only slightly, to the larger educational knowledge base especially in the area of conceptual understanding and computational problem-solving. On a more immediate level I am hoping that my study will improve student performance in my classes. My hope will be to get to the point where I feel that I accurately know my students' level of understanding coming into a summative assessment. I also hope to understand how much I need to focus on conceptual understanding versus problem-solving procedures in my classes. Lastly, I hope that the process will help students to be more cognizant of their learning and help in the process of them becoming life-long learners.

Support Team

My support team consisted of four people. Two worked at my school. The third was a professor at a university. The last person was a personal friend.

Emily- Science Teacher, St. Ignatius College Preparatory

Emily is the other teacher at my school who taught honors physics. She double majored in physics and chemistry at the University of San Diego and received her PhD in Bioengineering from UC Berkeley/UCSF. She was a good resource because we taught the same class but with slightly different methodologies. She gave feedback through our daily conversations of what was done in class and during planning sessions for our curriculum.

Paul- Assistant Principal of Professional Development, St. Ignatius College Preparatory

Paul's role in our school is to plan professional development and support teachers in their professional growth. Paul provided valuable feedback from an administrator's perspective. He asked clarifying questions and was supportive throughout the process. A few of the conversations with him led to possible ways my treatment can be useful in the broader school setting.

Evelyn- Sociology Professor, University of San Francisco

Evelyn is a tenured professor at University of San Francisco and a longtime friend. She was selected to give an outside educator perspective. My conversations with her centered on implementation of the research. She is well versed in Action Research because it is taught in her University and her department. She always provided feedback on the final report.

Jess- good friend, Oakland, CA

Jess is a close friend of mine for many, many years. He works in the non-profit sector and specializes in cross-cultural leadership training. Jess was chosen to be my non-traditional choice. He mainly helped with proof reading of the final draft and provided impressions about the work.

Purpose and Research Questions

The purpose of this capstone project is to examine the impact of a formative assessment system on student learning. It will also attempt to determine if there is a difference in effectiveness between two types of formative assessments in physics, conceptual and quantitative.

The research question to be addressed is:

1. What are the effects of formative assessments in facilitating students understanding in physics?

The sub-questions are:

2. What are the effects of conceptual formative assessments on students' conceptual understanding? What are the effects of quantitative formative assessments on computational problems-solving?
3. How do formative assessment affect student's attitude on learning?
4. How do formative assessment affect the way curriculum is taught?

CONCEPTUAL FRAMEWORK

In the policy brief, *Formative Assessment: Improving Learning in Secondary Classrooms* (2005), the listed benefits of formative assessments include raising achievement gains, building students' 'learning to learn' skills, and improving student's ability to learn. These benefits ultimately arise from the goal of "identifying learning needs and responding to them" (p. 1). This implies that formative assessments are not only beneficial in the mastery of content but in the overall learning capacity of the student.

The brief goes on to say that there are several key elements to good formative assessments. They include frequent assessments, interactive assessments (student involvement in the process), the provision of tools that student use to judge their work, communication of learning goals, variation in instructional methods, a mixed approach in assessment (e.g. diagnostic assessments, open questions, questions regarding causality), and feedback that is timely, specific, and tied to the standards. Most importantly, teaching must change in response to the assessment.

Emphasizing this point, the brief argues that in order to be effective formative assessment must be seen as one part of a broader change in classroom culture. Not only must teachers assess student understanding and adjust their teaching in response but

connections to standards, teaching strategies, and metacognition must be made. The implication is that formative assessments should not just be another teaching strategy in a teacher's repertoire. Philosophy and goals for the classroom must realign with the assessments.

A major part of the formative assessment process is to use the outcomes to alter teaching. In order to facilitate this Walvoord, Bardes, and Denton (1998) developed a four-step process. It includes creating the assessment, creating the rubric, aggregating student scores, and developing a plan on how to change teaching based on the results. Adoption of the process ensures that all components of the assessments are covered especially the last step which involves developing a plan to change teaching which may easily be neglected.

In science, and in physics in particular, pre-conceptions are major obstacles in learning. Students don't come to the classroom as blank slates but with their own physical conceptions based on their everyday experience. Teaching in physics doesn't just entail giving students new information and problem-solving procedures, but also uncovering students' misconceptions and addressing them. Conlin, 2009, developed a way to deal with misconceptions revealed in assessments. He suggests looking for a perspective in which the misconception is reasonable. Most of the time misconceptions are not wrong ideas (they drawn from experience after all) but they are just incorrectly applied. In this sense teachers should consider these ideas to be resources (Conlin 2009). Conlin found this approach more effective than simply correcting students by offering the right explanation. This idea can be used to frame the post discussion in a formative assessment process. The discussion can center on the differing conceptions, where they

come from, what situations they are reasonable, and what conditions are important to consider for producing the right explanation.

Paul Hewitt, 2000, a pioneer in the conceptual movement, states that the physics is a challenging and frustrating experience for students “because physics courses usually emphasize the final stage of the learning cycle- application, skipping over the first two stages- exploration and concept development” (p. 41). His premise is that conceptual development is essential to computational type problems. He defines conceptual physics as an “emphasis on concepts rather than derivation and on critical thinking rather than computation” (p. 41). This sets up a distinction between two types of problems in physics. The first is conceptual which is primarily descriptive and focuses on “showing the relationships between ideas and concepts.” It can be described as qualitative. The other type of question is computational which focuses on “algebraic manipulations and computations.” It can be described as quantitative.

Other categories can be used to classify the qualitative and quantitative understandings in physics. Docktor, Strand, Mestre, and Ross (2010), performed a study on five types of questions: categorization, conceptual questions, finding errors, equation instantiation, and problem-solving. The first two can be considered qualitative in nature while the last three are more quantitative. Categorization problems are 3-problem categorization tasks which involve identifying the physics principle to apply, justifying why the principle applies, and planning out the steps to solve the problem. Conceptual questions are conceptual multiple choice questions. Finding errors involve being “shown a worked-out solution that includes a physics error and asked to identify and describe the error in writing. In Equation instantiation students are “shown a problem and worked out

solution in symbolic form; asked to assign/match the appropriate values for each quantity in the final expression” (Doktor, et al, 2010, p. 139). Problem-solving involves traditional free-response problems. Even though results of the study did not turn out to be reliable due to unanticipated circumstances, the categorization of problems are useful and relevant.

In formulating a model for formative assessments in the physics classroom, Robert Dufresne and William Gerace (2004) provides examples of different types of formative assessments. The assessments are categorized based on tasks students perform. *Exploring Naïve Concepts* are problems that target student pre conceptions. In *Honing and Clustering* students comparing and contrasting different situations that require several concepts. These situations are typically one that can be carried out experimentally. *Analyze and Reason Using Concepts* involves two processes in solving complex problems. They define analyzing as “break(ing) down a situation into basic parts to understand the whole” (p. 430). Reasoning, on the other hand, involves “putting together the parts to draw conclusions or make judgments” (Dufresne & Gerace, 2004, p. 430). *Concept-Based Problem Solving* involves identifying the principle to solve a computational problem. One method that was given to do this is to have students decide which of two procedures presented for a problem is the valid one. Lastly, in *Organize and Interrelate Groups of Concepts* students are given two problems. They have to decide if the same approach can be used to solve the problems and then identify the principle to be used for each problem.

Hung and Jonassen (2006) looked at two different types of conceptual reasoning processes in a study that tried to determine the effectiveness of each in solving physics

problems. The first one is called covariational which looks at the correlational relationship between quantities. It looks at repeated associations to determine direct or inverse relationship between quantities. The second method is called mechanism-based. This type of reasoning looks at cause and effect relationships between quantities. In addition to establishing the relationship between the quantities as in the covariational process it also identifies the quantity that serves as the cause and the quantity that is affected.

The two reasoning processes were taught and used to solve different physics problems. The covariational method employed a simulation where participants changed values to determine the relationship between quantities. The mechanism-based method made use of Influence Diagrams which are a type of diagram that shows the chain of cause and effect relationships in a problem. The authors expected the mechanism-based approach to produce better results. The results of the study showed that students using the mechanism-based process scored better on conceptual questions while there wasn't a significant difference between the two groups when it came to computational problems. The authors hypothesized that there wasn't a difference in outcomes in the computational problems because the students didn't work with the different processes long enough to internalize and integrate the reasoning. They reverted to matching an equation to the problem as a problem-solving strategy

In summary, formative assessments are beneficial in the mastery of content and in the overall capacity for students to learn. Important characteristics include, among others, frequent assessments, interactive assessments, a mixed approach in assessment (e.g. diagnostic assessments, open questions, questions regarding causality), and feedback

that is timely, specific, and tied to the standards. Also, changes in teaching made in response to the assessment are important as the assessments themselves.

A division between qualitative and quantitative types of questions can serve to be useful in teaching physics. It can ensure that the former will not be neglected for the benefit of the latter which can typically be the case. Different categories within the qualitative and quantitative types can be made. Examples include identifying principles, finding errors, matching values to equations, comparing and contrasting problems, analyzing and reasoning in a problem, and solving problems using correlation and causation type reasoning.

Major stumbling blocks in learning physics are misconceptions. Incorporating them in the assessment can therefore be beneficial in addressing them.

METHODOLOGY

Treatment

My treatment consisted of implementing a formative assessment system within a flipped curriculum lesson cycle. Two different types of formative assessments were administered- conceptual and quantitative.

The beginning of a lesson cycle consisted of students learning concepts and problem-solving through reading PowerPoint slides for homework. Staggered throughout the slides were questions that checked for student understanding of basic definitions and concepts. Also included were sample computational problem-solving questions that were worked out. For the more challenging sample problem-solving questions video tutorials were created and made available to the students. The end of the homework lesson

typically consisted of two computational problems which students worked out on their own. Answers were given for those problems but no work was shown.

The treatment consisted of multiple choice formative assessments (FA's) taken in the class period following the homework lesson. They were administered in the beginning of class and consisted of three to four questions. Occasionally, a free response question was used in place of multiple choice. The FA's were created through Google forms and students completed them on their iPads within a time period of about five minutes. The results were discussed immediately afterward. In the discussion students were randomly called upon to explain their thinking. It was encouraged for other students to participate in the discussion as well. At the end of the discussion the percentage of correct answers for each question were revealed and discussed.

The discussion then turned to the homework lesson and any misunderstandings that resulted from it. If additional practice was needed, students worked in pairs on additional practice problems that incorporated the concepts from the homework and assessment.

The formative assessment consisted of either conceptual or quantitative questions. The conceptual questions focused on definition, concepts, principles and qualitative predictions. An example of a conceptual-based question is given in Figure 1.

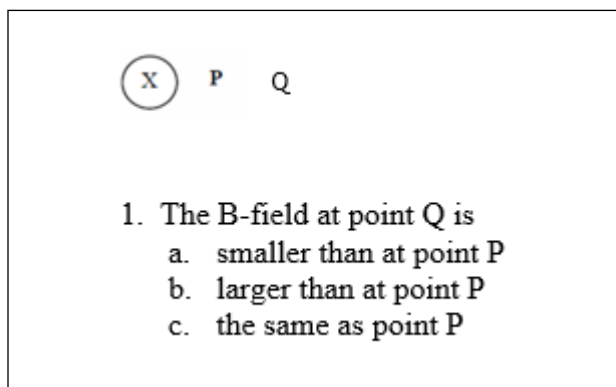


Figure 1. Example of conceptual question.

Quantitative questions involved numbers in the description and answer choices. Calculations, however, were not required. Sometimes equations were provided as well. An example of a quantitative question is given in Figure 2.

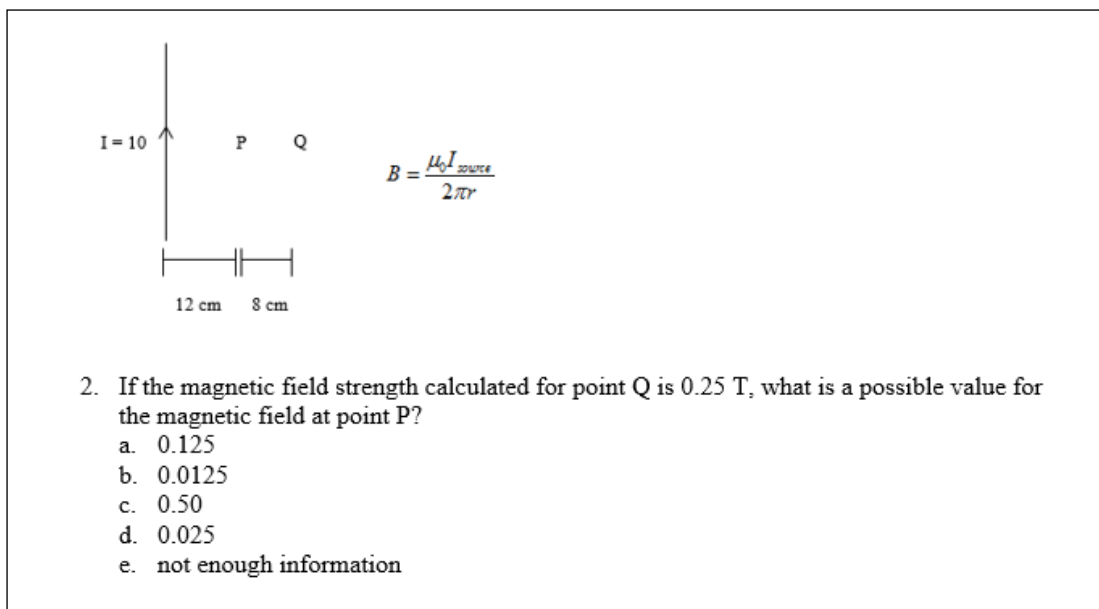


Figure 2. Example of a quantitative question.

Some of the formative assessments were designed to address misconceptions. These questions were intentionally included to assess, discuss, and remedy the misconceptions.

The treatment was applied to three periods of honors physics in three units of study during the second semester. The units were electric fields, magnetic fields, and electric potential and electromagnetic waves. Each unit was two to three weeks in length. The three class periods were cycled through treatment and non-treatment in a staggered way. For example, during the first unit of study one class section was given conceptual assessment questions, the second section was given quantitative questions, while the third was not given any formative assessments. For the second unit of study each class got a different type of formative assessment or the non-treatment. This was also done for the third unit so that each class period had two treatment units (conceptual and quantitative formative assessments) and one non-treatment unit (no assessments).

Instrumentation

The treatment was conducted on three sections of honors physics at a private Catholic high school in San Francisco, California. The population comprised of 74 students, 34 of which are female and 40 of which are male. Most of them come from middle to upper class socio-economic backgrounds. The ethnic breakdown of the participants mirrored the population of the school: 58% Caucasian, 22% Asian American, 13% Latino, and 4% African American. Five of the students have been diagnosed with learning disabilities and are provided with accommodations. Because of the prerequisites of the class all but one of the students are in honors or accelerated pre-calculus. The one

exception is in AP Calculus BC. Also because of the prerequisites the students have already taken biology and chemistry. Many students have taken honors chemistry.

In an attempt to answer the research question four data collections methods were employed. They were the formative assessments themselves, summative assessments, pre-and post-surveys, and pre-and post-interviews. The formative and summative assessments were the primary sources of quantitative data while the surveys and interviews provided the more qualitative viewpoints on the treatment. Table one summarizes the data sources and the research questions they address.

Data collection started in the second semester. However, the formative assessment system that was used for the treatment was implemented at the beginning of the school year. At the beginning of the second quarter an anonymous mid-semester survey was administered to all participants in the study. The survey related to the tools used in the class. The goal of the survey was to get feedback on the effectiveness of the different tools which included the formative assessment system. The results of the survey were not part of the instrumentation but they were used as a tool to improve the treatment and to develop some of the questions in the instrumentation surveys and interviews. The Mid-Semester Survey questions are included in Appendix A.

Table 1. *Data Triangulation Matrix*

Research Question	Data Source			
	1	2	3	4
1. Formative assessments in general	formative assessments	unit summative assessments	pre- and post-surveys	pre- and post-interviews
2. Conceptual formative assessment	formative assessments	unit summative assessment	pre- and post-surveys	pre- and post-interviews

3. Computational formative assessments	formative assessments	unit summative assessment	pre- and post-surveys	pre- and post-interviews
4. Attitudes on learning	pre- and post-surveys	pre- and post-interviews	teaching log	
5. How Curriculum is taught	teaching log			

At the beginning of the data collection period a non-anonymous pre-treatment survey was administered to all participants through a Google Form. At the time of the survey the students had a semester's worth of experience with the formative assessments. The pre-treatment survey was intended to probe their perceptions of the assessments and their attitude in learning. The Pre-treatment Survey is listed in Appendix B. Appendix C indicates how the survey questions correspond to each research question.

In the week following the Pre-treatment Survey a Pre-treatment Interview was conducted. The pre-treatment interview questions are listed in Appendix D. Appendix E indicates how the interview questions correspond to each research question.

Six students were interviewed. Two were lower performing students, two were middle performing, and two were high performing students. The performance level of the students was determined by overall class grades and performance on the formative assessments from the previous semester. Students were chosen randomly within each performance level. Four of the six students were interviewed individually while remaining two were interviewed together. The last two were interviewed together because of scheduling conflicts of one student due to absences and extra-curricular activities.

The formative assessments consisted of three to four multiple choice questions given at the beginning of class. Occasionally there were free-response questions. They were administered electronically through a Google Form. As stated in the treatment, each class section had either conceptual or quantitative questions for the two treatment units. In the non-treatment unit the classes were not given any formative assessments.

The summative assessments consisted of unit exams. The unit exams were administered at the end of each unit. They only consisted of free response questions which were conceptual or computational in nature.

At the end of the three units a post-treatment survey was administered. Most of the questions on the post-survey were identical to the pre-treatment survey in order to analyze any changes in attitudes and perceptions.

Similarly, post-treatment interviews were conducted with the same participants in the pre-treatment interviews. Some questions from the pre-treatment interview were used to determine any changes in perceptions. However, additional question relating to the difference in conceptual and problem-based formative assessments and the difference in the presence and absence of formative assessments were included.

A teacher log was kept throughout the data collection process. After every class, observations were recorded as well as notes on changes made to or planned for instruction that resulted from the results of the formative assessments.

Several measures were taken to ensure validity and reliability of the results. A major component to ensure validity was the triangulation of data. For the major research questions at least three data sources were collected. They were a mixture of quantitative and qualitative sources. The data analysis below shows that the qualitative interview and

survey data provided invaluable insights that complimented the interpretations of the quantitative results to provide a more complete perspective on the treatment.

Discussions with support team members were also valuable in ensuring validity. Discussions with Emily in particular provided useful insights during the data collection process. She provided key feedback on some of the FA questions before they were administered. She also gave useful ideas during the data analysis process.

One measure that was taken to ensure the reliability of the FA data was the inclusion of a survey question at the end of each FA quiz. The question asked: “Were you able to complete the HW Lesson before class? If not, why?” It was emphasized to the students that the question was only going to be used in interpreting the FA results. Its inclusion was intended to determine to what extent incorrect results were due to failure to do the absence of learning rather than incorrect learning.

As mentioned, the FA’s were implemented starting the beginning of the school year and the data was taken in the second semester. The reasons were to work out any unforeseen problems, to determine how to write the FA questions to measure student learning effectively, and to acclimate the students to the assessment process. Having a vetted formative assessment system before data collection helped ensure the reliability of the data.

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 F).

DATA AND ANALYSIS

The data revealed many trends and patterns on the role formative assessments have in student learning. The major findings indicate that:

- The performance on formative assessment does not correlate well with performance on the summative assessments. Also, the presence of formative assessments did not seem to affect them as well.
 - There is a possible correlation between the implementation of conceptual formative assessments and performance on conceptual questions on summative assessments.
 - A similar possible correlation between quantitative FA's and computational questions on summative assessment was shown.
 - Students answer conceptual FA's more successfully than quantitative FA's but prefer quantitative questions when asked for their preference between the two.
 - The perception students have of the role of the FA's correspond well with their intended function of determining what is known and how well it is known.
 - The FA's are not seen as essential but as a helpful supplement to learning.
 - The students have an overall positive attitude toward the formative assessments
- The data on the summative assessments, formative assessments, surveys, and interviews will be discussed in turn.

Summative Assessments

The summative assessments include the scores for the unit exams in each of the three units in the study. Each question on the exam was classified as conceptual or quantitative. The number of possible points for each item was determined by the number

of steps or parts involved to obtain the correct answer. The number of points for each item was tallied for each student for each exam, and the students were separated according to their class section. The overall score for each student was calculated into a percentage using the number of points earned for the total number of points on the exam.

Figure 3 is a box-and-whisker plot of the overall scores separated by unit and class period. The red shading indicates the non-treatment unit for the class section and the blue shading indicates the treatment units.

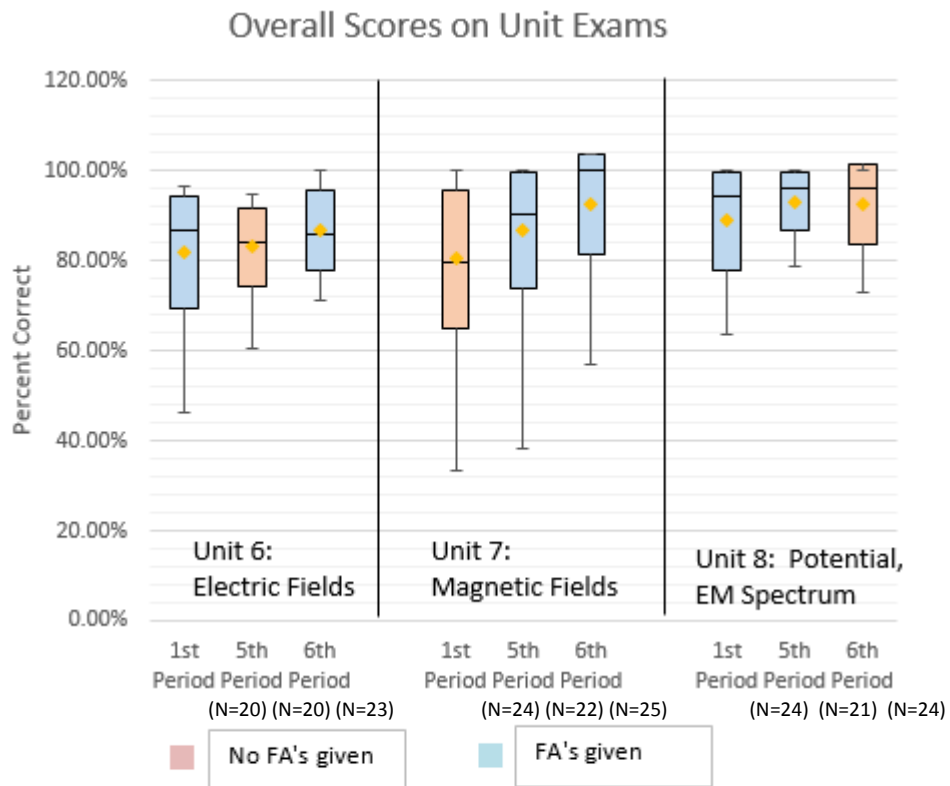


Figure 3. Box-whisker of overall exam scores by unit and class period.

The data shows that non-treatment units only resulted in lower average and median scores for one the three units. This implies that there isn't a clear correlation between the presence of formative assessments and exam performance. The data,

however, does show a correlation between class performance and class period. 1st period performed the lowest in terms of mean, standard deviation, and range in all three units while 6th period had the best average in two of the three units (it was only second by about 1% for the third unit). A possible reason for these results may be the distribution of students into specific classes by the student assignment process of the school in which this study took place. In terms of mean and median scores, 1st period was scored consistently lowest while 6th period score the highest in two of the three units with a close second on the third. The defacto tracking of students in high school naturally occurs when student select their courses, especially between honors and regular courses. This results in different classes having different levels of academic strengths. The data may reflect this trend.

A surprising outcome from the graph is how clearly it shows the difficulty level of each unit. With one exception, the average for all class periods progressively rose with each unit. The exception is 1st period. The averages for that period were 81.96%, 80.46%, and 88.86% for Units 6, 7, and 8 respectively. It can be seen, however, that the difference between Unit 6 and 7 is just over one percentage point. This trend makes sense when examining the content of each unit. Traditionally Unit 6 on electric fields is one of the hardest units in the course. The idea of the field is introduced and is conceptually new to most students. Also, the amount and level of mathematical calculations is fairly high when calculating the total electric field. The subsequent units get more conceptually abstract but become mathematically simpler. The trend in the data suggest that mathematical complexity may be the key factor in determining physics exam scores.

The data point that stands out in Figure 3 is unit 7 for 1st period. It has the lowest mean and one of the larger standard deviations. A closer look at the scores for that exam revealed that there was one test score in particular that was very different from the rest. One student scored a 33.33%. The class average was 80.46%. Taking out that one score raised the average to 82.51% and lowered the standard deviation from 15.33 to 11.85%. Although the one score did not account for the entire point difference, its omission brought the statistics of the class closer to the other data points. Extreme outliers (exams below 50%) were identified for only two other cases. One was for the unit 6 exam, 1st period (46.43%) and other for unit 7 exam, period 5 (38.10%). The outliers contributed a good amount to the statistic of those data points, particularly the standard deviation. However, their omission would not change the relative relationships between the class sections and therefore the overall trends would remain.

The scores for the type of question were also analyzed. The number of points for each type of question, conceptual or quantitative, was tallied for each student. The score for conceptual type questions was then calculated into a percentage using the number of points earned for the conceptual questions and the total possible points on those questions. The same was done for quantitative questions. The box-and-whisker plot for conceptual questions are shown in Figure 4.

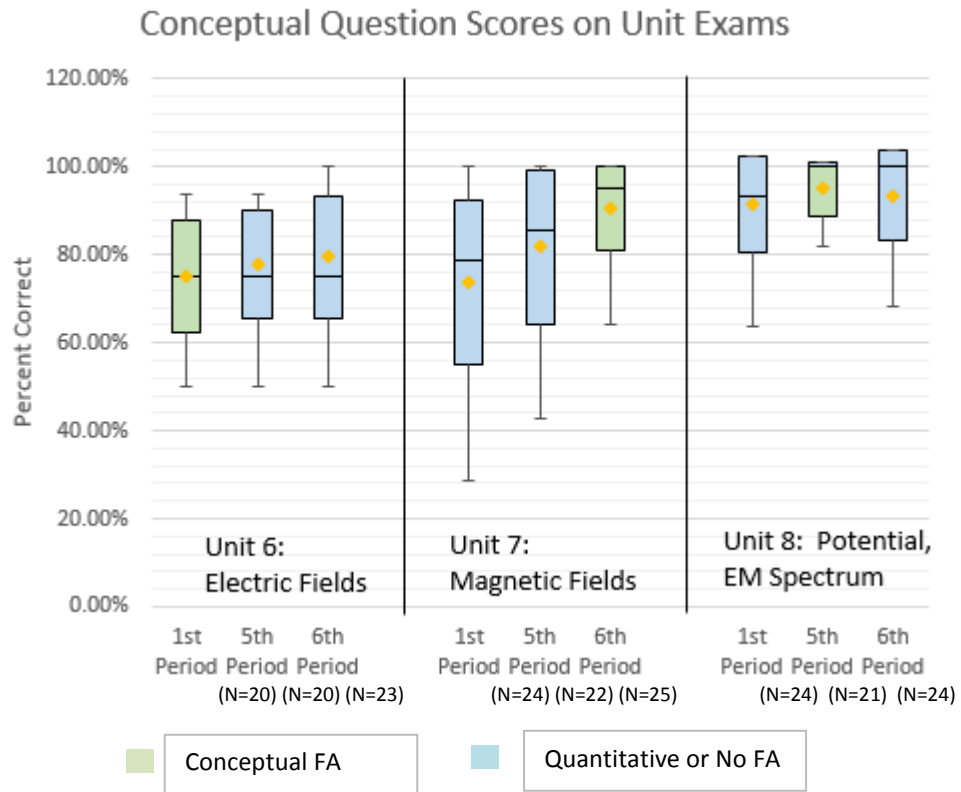


Figure 4. Box-whisker of conceptual questions by unit and class period.

The data suggests that there may be a correlation between the implementation of conceptual formative assessments and performance on conceptual questions on unit exams. In two of the three units the class that was given conceptual formative assessments had the highest mean, smallest standard deviation, and smallest range. The median was the highest or tied for the highest in all three units. In one of the units, however, the class that had conceptual formative assessments had the lowest average.

The box-and-whisker plot for quantitative questions are shown in Figure 5.

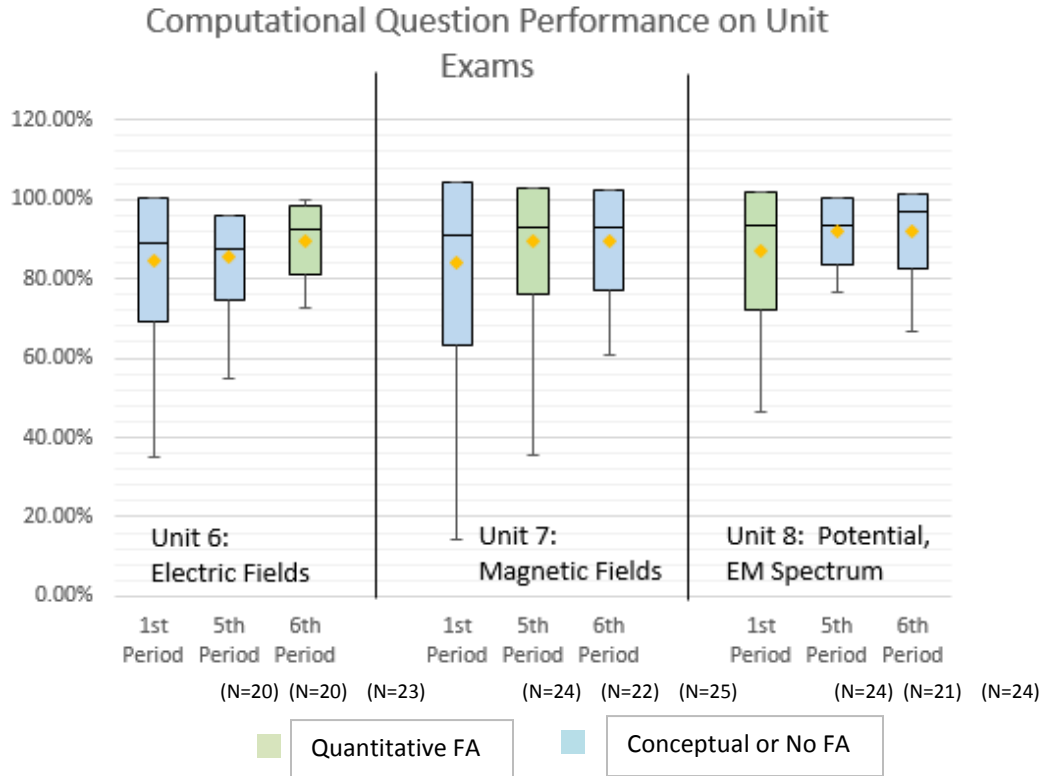


Figure 5. Box-whisker of Computational Questions by Unit and Class Period.

The data suggest a similar pattern to that of conceptual assessments but the trend in statistics is not as strong. In two of the three units the class that was given quantitative formative assessments had the highest mean and median. In one of the units, however, the class that had quantitative formative assessments had the lowest average. There wasn't an apparent trend in standard deviation or range.

Students with scores over one standard deviation below the mean were examined. To simplify the analysis the focus was on students who fell into this category in more than one unit in the study. This was also done to mitigate the situations where extraneous circumstance may have contributed to an uncommonly low performance. An example

may be several large assignments being due around the test date. From all three class sections six students were below one standard deviation in more than one unit. Four of them were girls and two were boys. All six students had inconsistent performance throughout the school year on exams. In terms of grades they fall towards the middle to lower end of the class.

All six students had low performance in two of the three units in the study. For five of the six students, one of the low scoring exams was in the non-treatment unit. However, there wasn't a pattern on the second low scoring exam. Between the students it was about an even mixture between conceptual and quantitative formative assessment treatments. Overall, in nine of the twelve exams taken between the six students the students scored lower on the quantitative questions than on the conceptual questions. Taken together these pieces of data suggests that there may be a connection between the implementation of formative assessments and performance on exams for these mid to lower performing students. It is possible that the students would not have scored as low if formative assessments were given in the low performing no-treatment units. However, at this point the connection is tenuous. The data isn't clear on the low performance in the treatment units. There isn't a pattern to the type of FA's given and the type of incorrectly answered questions on the exams. Other confounding factors such as time and effort spent on the class assignments may play a bigger role.

Formative Assessments

Each formative assessment consisted of three or four multiple choice questions. The formative assessments for each treatment unit were either all conceptual or all quantitative. Because of the small number of questions per assessment, the scores were

collected for the class as a whole. For each formative assessment the overall score for the class was calculated into a percentage using the total number of correct responses in the class. Then the statistics were calculated for all the formative assessments in the unit.

Figure 6 organizes the statistics by unit.

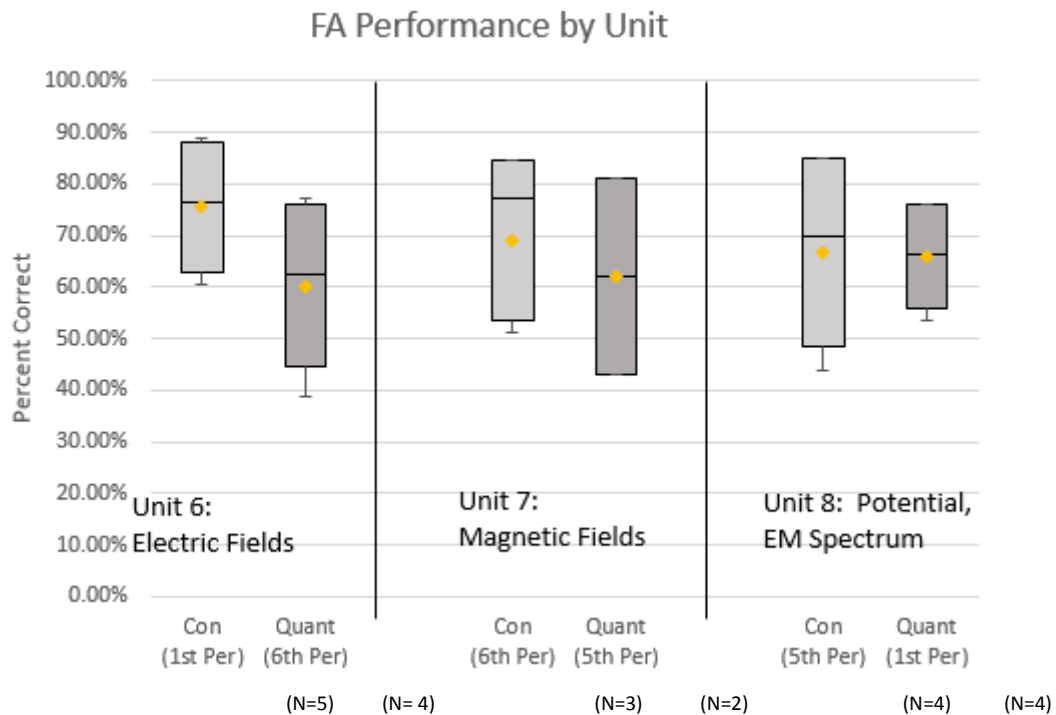


Figure 6. Box-whisker of FA's by Unit.

The data shows that the conceptual FA's were answered more successfully than the quantitative ones. The mean and median was higher for the conceptual in all three units. The standard deviation and range was lower for two of the three units. An interesting trend in the data is that the means of the conceptual FA's progressively decreased with each unit while the quantitative FA's exhibited the opposite trend. The spread also increased progressively with conceptual FA's as shown by the standard deviations.

These trends make sense in terms of the content of the different units. Unit 6 on electric fields is the most computationally involved out of the three because it involves resolving vectors into components and finding resultants. The calculations for the successive chapters get progressively simpler. In unit 7 on magnetic fields the vectors are limited to be along the same line and unit 8 on potential uses scalar and not vector addition. Conceptually, the opposite trend is true for the difficulty. Unit 7 is conceptually more challenging than unit 6 because it involves using hand rules to determine the 3D direction of the vectors. Moreover, Unit 8 is probably the most conceptually difficult because the concept of potential is very abstract and can be easily confused with potential energy.

Figure 7 shows the same data but separates the FA data by class period.

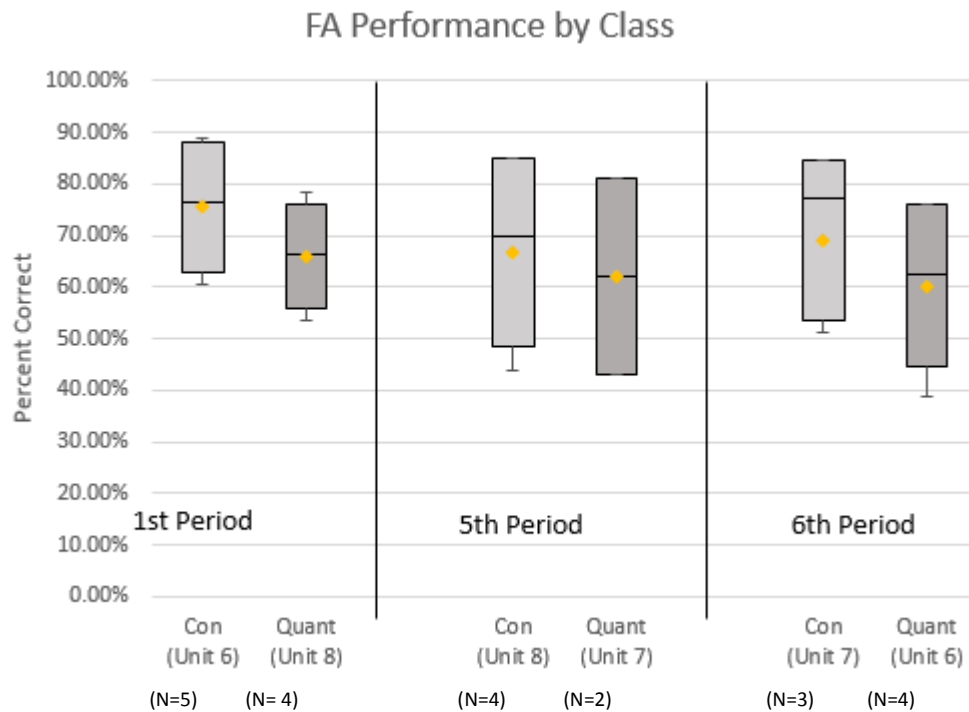


Figure 7. Box-whisker of FA's by class period.

The graph shows that 1st period had the highest mean and lowest standard deviation for both conceptual and quantitative FA's. 6th period had the lowest mean for quantitative formative assessments. This pattern doesn't match well with the pattern seen in the summative assessments where 1st period showed the lowest performance and 6th period showed the highest. This indicates that performance on the formative assessments is not a good indicator for performance on the summative assessments. This lack of correlation makes sense in light of the purpose of the formative assessment which is to determine how much the student has learned. A major determining factor for performance on summative assessments would be how much is relearned and corrected in between the formative and summative assessments.

The individual formative assessment results of outliers identified in the summative assessment analysis above were analyzed to determine if there are any connections between the FA performance and SA performance for these students. Five of the six students indicated that they did not complete the homework lesson for at least one formative assessment (this question was included at the end of each formative assessment) in the two low performing units. This may be an indication of a smaller amount of time spent on the material compared to other students. For all the formative assessments given in this study very few students indicate that they do not complete the homework lesson (two or three for each FA at most, if any). However, since this is self-reported, the accuracy of the number of students who actually complete the homework lesson may not be accurate.

The only other observed pattern for these students is that they consistently missed at least one question in the FA's for the low performing treatment unit. There was only

one exception to this. One student did not get any questions incorrect in one formative assessment (this represents one case in 21). It is not uncommon for a student to miss a question on the formative assessment. Therefore, this would not be a sufficient predictor for low performance on tests. However, this may be a useful indicator of students in need of extra assistance. In the future, flagging students who both consistently miss FA questions and report not being able to complete homework lessons may be an effective method of identifying students who may need extra attention.

Surveys

Responses to the key questions in the pre- and post- surveys were analyzed and compared to gain qualitative insight on the role the formative assessments have in learning. The intent was to determine how the FA's aid in learning and the student's perception of them.

One question that was sought to be answered was whether the students perceived the FA's as an essential part of learning or merely a useful supplement. Two questions were included to determine this. The first question asked whether the FA reinforces learning while the second asked if they are a key part learning. The result for the pre- and post-surveys are shown in Figure 8.

FA Survey Results Questions #1 and #4

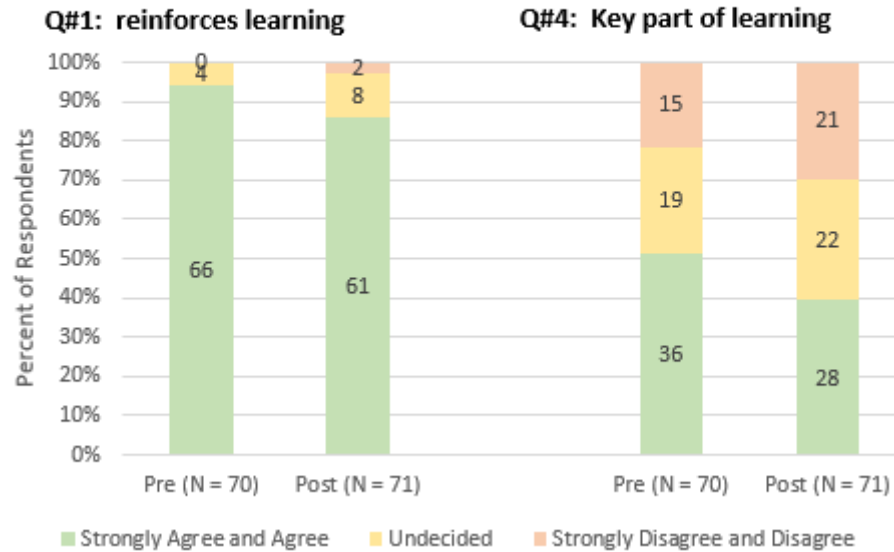


Figure 8. Pre- and post-survey comparisons on the essential nature of FA's.

The graphs show that over 85% of the students in the post-survey agreed or strongly agreed that the FA's reinforce learning while only about 40% felt that they were a key part of learning. In the open-ended part of the latter question, question #4, sample student comments included:

“They help explain some questions that I may have been confused on from the homework.”

“Doing the homeworks (sic) are a key part of learning physics. The assessments are just making sure we learned correctly.”

“It's not necessarily a key part of my learning, but I do find them helpful in inspiring me to do homework.”

“The formative assessments are a type of review for me, especially on Mondays after a long weekend.”

These responses show that students find the FA's to be helpful but not essential to learning. This is also corroborated by the interviews as discussed below.

The graphs also show a lowering trend in the number that agreed or strongly agreed in between the pre- and post-surveys. On the question about reinforcing learning the number that agreed went from 94% to 86%. On the question on FA's being a key part of learning the percentage went from 51% to 40%. This trend was also seen in some of other questions discussed below. The non-treatment was the only unit all year that the students did not have formative assessments. It is possible that the FA's were not seen as effective as before because of the students being able to successfully learn the material without them.

The surveys also sought to determine how the FA's aided in the student learning. In writing the formative assessment questions there was an attempt to ask questions in ways different from the ones in the homework lesson and also to provide new situations to which the concepts were applied. One question on the survey asked if the FA's provided new questions and situations. This was to determine if the students perceived a difference. Another question asked if the FA's showed what they don't know or need to work on. The results for these questions are shown in Figure 9.

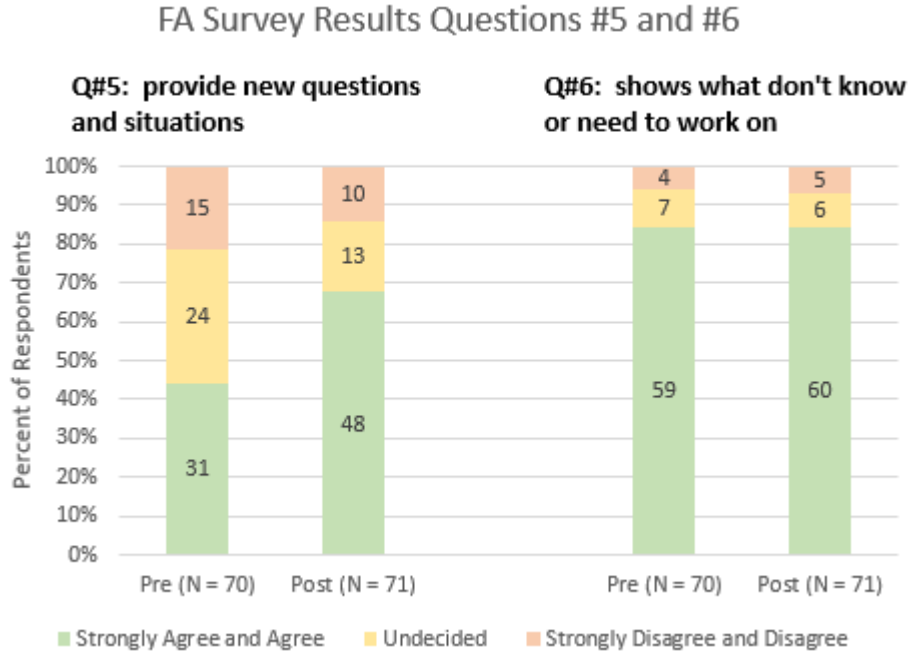


Figure 9. Pre- and Post-survey comparisons on how FA's help learning.

The graphs show that by the end of the treatment period 68% agreed that the FA's provided new questions and situations. This is a significant increase from the 44% in the pre-survey. It is also the only time that showed an increase between the pre- and post-surveys. This may be attributed to the fact that the questions for the treatment were intentionally written into conceptual or quantitative types. Having that constraint required the creation of questions in ways that normally would not have been written in order to meet the conditions. Before the treatment, the formative assessments did not have to fall into certain categories and usually it was easier to write questions that were very basic in nature. This result shows that the way questions are constructed is important in their effectiveness.

The question related to the FA's showing what the students don't know or need to work on was consistent in the pre- and post- at 85%. This data shows that the perception of the FA's match what they are intended to do. This was also corroborated in the interviews where students made the point that they are helpful in determining what they need to work on.

Two questions were asked to determine the effectiveness of the FA's in the conceptual and quantitative categories. One question asked if they helped in learning concepts and definitions. The other asked if they helped in learning how to solve numerical problems. Figure 10 shows the results.

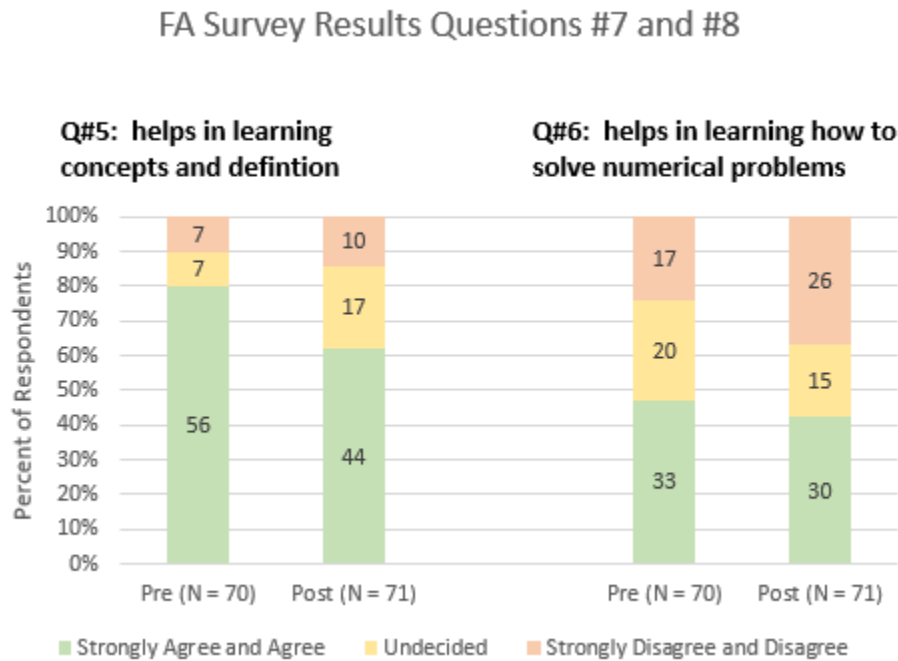


Figure 10. Pre- and post-survey comparisons conceptual and problem-solving type of learning.

About 62% agreed that the FA's helped in the learning of concepts and definitions in the post-survey while only 42% agreed that they help in learning how to solve numerical problems. These questions also showed the decline in perceived effectiveness between pre- and post-surveys discussed above. The FA's were designed to present questions and situations to fully develop the concepts beyond their basic definitions. For example, after students learned that the electric force on a particle is caused by the electric field the formative assessment questions explored that concept further by looking at situations where the field is weaker or where the particle carries less charge. Therefore, the connection to the definitions and concepts are more subtle and indirect. The data, therefore, can be seen as the students perceiving the value of seeing concepts in different contexts.

The connection between conceptual understanding and problem-solving is a subtle one as well. Success in solving a computational problem requires conceptual understanding of the situation and of the principles involved. The data shows that less than half of the students make the connection between concepts and problem-solving. Some of the data from the interviews, as discussed below, suggest that students see conceptual and computational problems as distinctively different and don't see that the later relies on the former.

Included on the post-survey but not the pre-survey were direct questions on the impact of the FA's on learning and exams. They were only present in post-survey because the idea to include them resulted from the responses in the pre-survey. One question asked if the student learned better as a result of the FA's. The second question

asked if the student did better on the tests because of the FA's. The data for the two questions are shown in Figure 11.

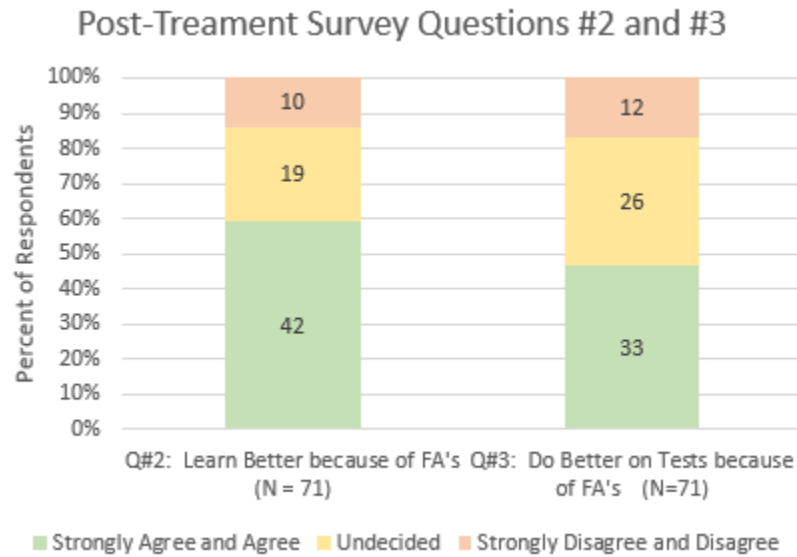


Figure 11. Post-survey questions on the direct effect of FA's on learning and tests.

The graphs shows that 60% agreed that they learn physics better with the FA's while only 46% felt that they do better on tests because of them. The results are fairly significant. Close to half of the students feel that they directly contribute to learning and performance.

Interviews

The interview data reinforced many of the findings of the other instruments in the research. The students interviewed felt that the FA's made a difference in their learning, but were more ambivalent about their impact on test scores. Just like in the surveys they felt that the FA's were a good way to tell how well they know the material. In terms of attitude in learning the FA's seem to have a positive effect.

The surveys revealed that the formative assessments were seen as helpful but not essential to learning. The interviews probed deeper into this notion and tried to determine to what extent. All of the students interviewed said that the FA's helped their learning. Two of those students were more cautious in their response by only saying that they sometimes help. The reasons they gave corresponded to the same reasons cited in the surveys. They saw the FA's as review and reinforcement opportunities. The FA's determined how well they knew the material and what they needed to work on. They cleared up concepts. One student stated, "They don't show what you know but what you need to work on." This shows that the student sees the FA's as a means to an end rather than a way to determine an end. He goes further in saying, "They are a good review or a good way to start review, but they shouldn't be a substitute for review." He seems to be trying to get across the point that FA's are one helpful way to learn concepts but that they shouldn't be the main or most important way to do so. This connects to the survey finding that the FA's are not seen as key to the learning process by most students.

A reason for their effectiveness given by another student was that they reminded her of the importance of doing the homework lessons. She said that this occurs when she fails to do the homework and get many of questions wrong. This shows the value of the FA's as a feedback mechanism for students even when it is assessing what has yet to be learned.

When asked if the FA's made a difference in exam performance most of the students were at a loss. Several said "I don't know." A couple said yes but were unsure in their responses. Two students noted that the exam questions were qualitatively different. One said, "Not so much on test and quizzes because those are more like

homework problems.” The other said, “(It) reflects conceptual understanding, but questions are not like the formula stuff, not really solving for anything- so not like that.” Here both students see the exam questions as primarily being computational problems even though there is a mixture of conceptual and computational problems. This corresponds with the observation made overall that student focus most of their learning on the calculation-based problems. They are the hardest type of problems because they incorporate math as well as physics concepts and they are usually multi-step. An important goal in the future will be to determine ways to get student to realize that success in solving computational problems builds on conceptual mastery.

Another goal of the interviews was to determine what corrective steps students take after getting feedback from the formative assessments. When asked if they change how or what they study as a result of their performance on the FA’s many said that they don’t. Several indicated that they do minor corrective actions to understand the answer like looking back at the homework lesson or asking the people around them. One student said, “Well, we’ll review it in class anyways.” This points out a possible area of improvement. In the future it might be helpful to teach the students direct ways to strengthen the weaknesses revealed by the FA’s. A possible reflection time after the assessment where students can write down their mistakes and reasons why they are wrong might be helpful in engendering a habit of taking action on feedback.

The interviews also sought to determine whether the FA’s have positive effect on student’s attitude in learning. As stated above, all students indicated that they were helpful in some way by mainly determining what they needed to work on. All students said that the FA’s are at the right level of difficulty. Several students stated there is a

good balance of easy and difficult questions. Five of the six interviewees said that they felt more relaxed in doing them as compared to doing tests. One student, though, pointed out that she felt more prepared for tests. Another said that even though they feel less pressure people still try to get it right. He said, “I feel like people definitely do try to get them right ‘cause they don’t want to see it up on the board like 25 people got this right and one person got it wrong and you don’t want to think, Oh, I’m that one person that got it wrong.”

He was referring to the post discussion where the answers and the percentage of the class that got them correct was shown. This shows that even though the FA’s aren’t graded there are some incentives for students to try their best.

When asked for ways they can be improved most of the students couldn’t think of anything. The two suggestions were to give more free response questions and the other was to possibly add computational questions. Those suggestions will be taken into account in the future. In the end of the interviews the students were asked if they wanted to give any other comments on the FA’s. Four of the students said they felt they are good.

Taken together these responses indicated that the students approach the formative assessments in a positive way and make use of it as a tool for their learning. One of the six students wasn’t as positive about the FA’s as the rest. She was one of the students that struggled in the class. She stated that she doesn’t like doing them all the time because she would rather use the time to go over the homework. When asked if she noticed the times when the FA’s were not given she said no but she noticed when they were given every day. Again she was stating that she would rather go over the

homework. Her perspective was understandable. She struggled with the computational problems which involve a series of steps. She was sometimes overwhelmed with the class and her other classes as well. Her situation highlights a possible use of the formative assessment system developed for this study. Because of a lack of time individual students were not tracked and monitored through the formative assessments. The data collected from each FA could provide a way to recognize students who struggle early on in the unit. It is a process that will be considered strongly for next year.

One interesting outcome from the interviews occurred when students were shown examples of a conceptual FA question and a quantitative one. They were then asked which they felt was more effective. Four of the six students chose the quantitative FA, one said both were important, while one chose the conceptual FA. The student that picked the conceptual FA said she chose it because no calculator was needed and it probably was easier. She misunderstood the question because none of FA's required actual calculations. When clarified she felt that the conceptual one was still probably easier. The reason given by the other students as to why they chose the quantitative FA was because they felt they were more concrete. One student said,

For me personally, I prefer the one with the numbers because that just gives me something easier to grasp instead of like something like the first one where it would be like it's sort of vague, I don't really like it when it's that vague. 'Cause I like it like shown at me in my opinion, for me at least.

Another student mentioned something similar.

Cause I could see it. This is kind of like I see three things and I kind of have to interpret it, it's like almost like being lazy. This one is right in front of you. It's there for you. And I kind of feel like you get a much quicker start to it...

The preference for quantitative formative assessments is counter to the FA data that shows that student answers conceptual FA's more successfully (Figure 7). However, the quote from the second student above may indicate an explanation for this trend. Quantitative FA's may appear quicker to apprehend and therefore students may be quicker to instinctively select an answer. If the conceptual FA's require more thought then they students may be more apt to think slower and more carefully.

The interviews revealed that the FA's had a positive effect on student's attitude in learning. Three students mentioned that physics was challenging but five said that they liked it more than biology and chemistry. The students said that they don't do similar assessments in other classes. Two students said they do warm-ups but the teacher didn't use them to assess the understanding in the class. They did say that they thought they would be helpful in math or science where the material are more factual and less open to interpretation (four students).

There weren't significant differences in the responses of the students in the low, medium, and high group. Most of the questions focused on the role of the formative assessments in learning and the students had similar views. This results seems to reinforce the notion that there isn't a direct connection between performance on the FA's and SA's. If there was a perceived correlation on the part of the students then it would be expected that lower performing students would view them as less effective.

The one perceived difference between the groups, however, was mentioned above. Five out of the six students had a noticeably positive attitude towards the formative assessments and physics. The one outlier was one of the lower performing students. It would make sense that performance does have an effect on attitude. It is

natural to not enjoy a subject as much if it proves to be difficult. This is especially true when the stress of maintaining high grades is factored in.

INTERPRETATION AND CONCLUSION

The four questions this research sought to answer are:

1. What are the effects of formative assessments in facilitating students understanding in physics?
2. What are the impacts of conceptual formative assessments on students' conceptual understanding? What are the impacts of problem-based formative assessments on students' conceptual understanding?
3. How do formative assessment affect students' attitude on learning?
4. How do formative assessment affect the way I teach the curriculum?

Each question will be addressed in light of the data collected and the trends in the data discovered.

1: Effects of Formative Assessments on Student Understanding of Physics

The summative assessment data indicate that there is no clear correlation between the implementation of formative assessments and mean exam scores. Therefore, there is no evidence of a direct effect on the accumulated knowledge and skills at the end of a learning unit (Figure 3). This does not mean there is no correlation at all. The data does show a correlation between class period and mean and median scores (Figure 3). As discussed in the data analysis, this may be a result of defacto tracking that results from the type of classes students select and their eventual assignment into them. The SA results would then be expressing the relative academic strengths inherent in the class periods. If the students were grouped by random assignment the result could possibly be

different and possibly show a positive trend between formative and summative assessments.

The exam data also shows a correlation between unit of study and mean scores (Figure 3). With the exception of one data point, the mean increased with each successive unit indicating that the overall difficulty decreased from unit 6 to unit 8. This pattern is significant in interpreting the results of the formative assessments.

The formative assessment data also supports the notion that there isn't a direction connection between formative and summative assessment performance. The data, paradoxically, actually shows an opposite trend between performance on formative assessments and performance of summative assessments. The lowest performing class on the unit exams, 1st period, had the highest mean and lowest standard deviation in both the conceptual and quantitative formative assessments (Figure 7).

The trend becomes clear, however, when looking at the difficulty of each unit overall and in their respective types of questions. As discussed above, the summative assessment data shows that overall difficulty increases with each successive unit. As discussed in the data analysis, the formative assessment data shows that conceptual difficulty increases with each successive unit and the computational difficulty decreases with each successive unit (Figure 6). The mismatch between formative and summative performance exhibited in 1st Period became clear when it was realized that 1st Period fortuitously had formative assessments in the easiest conceptual unit (Unit 8) and the easiest quantitative unit (Unit 6).

Taken together, the results of the summative and formative assessments show that other factors, such as conceptual, computational, and overall difficulty and students

sorting into class periods, contribute more to the overall performance of the class on unit exams than the presence of or performance in formative assessments. As discussed below for research questions two and three, however, there seems to be a connection between the type of formative assessment given and performance on those corresponding types of questions on unit exams.

The surveys and interviews gave better insight on how the FA's impact student learning. 60% of the students agree that they learn better as a result of the FA's but only 46% feel that they do better on the exams because of them (Figure 11). The latter statistics reinforces the finding from the SA and FA data that shows there may not be a direct connection between them. In terms of how they impact learning, 86% of the students feel that they reinforce learning while only 40% feel that they are a key part of learning (Figure 8). Getting more specific, 85% feel that the FA's show what they don't know or what they need to work on and 68% feel that present new situations or questions (Figure 9). Taken together, the survey shows that students feel FA's are useful tools that reinforce learning by pointing out gaps in their knowledge. This perception exactly reflects the purpose of the formative assessments.

The interviews showed similar results. All six students interviewed indicated that the FA's helped their learning even though two said they only do so sometimes. The reasons given by the students parallel those revealed in the survey. The students said that they determine how well they learned the materials, they show what they need to work on, that they are a type of review. One student said that the FA's reminded her the importance of doing the homework lessons. Another student said that sometimes they directly help with questions on the exams.

The interviews showed that students were unsure about the impact of the FA's on exam performance. This reinforced the finding of the other instruments. One student stated that the exam questions were less like FA's more like homework problems while another implied the exam were different because they were primarily computational problems.

Taking the findings of all the data sources, the formative assessments fulfill their primary purpose of being a useful tool for students to gain feedback on their performance. It is through this feedback that the FA's impact student learning. Students can correct misconceptions and determine what areas of learning need work. The study also shows that limitations of the FA's. They are not good predictors of summative performance. They are one part of the learning process that ends with the summative assessments. In between the FA's and SA's the student's corrective action should have the most impact on performance.

2: Impact of Conceptual and Quantitative FA's on Corresponding Questions on SA's

In order to answer this question the results in the formative assessments were analyze with those in the summative assessments. The data actually showed that there may be a possible correlation between the implementation of conceptual formative assessments and performance on conceptual questions on exams (Figure 4). In two of the three units, the class that was given conceptual formative assessments had a higher mean and lower standard deviation on the conceptual questions on the exam. The exact same trend is seen for quantitative formative assessments (Figure 5). The connection makes sense if the FA are seen as a mode of practice. Implementing conceptual formative assessments constitutes more practice on those types of questions. The students would be

exposed to a greater number and variety of those questions and should be better at answering them for a particular unit. This was suggested when a student mentioned in an interview that some questions on the tests reminded her of formative assessment already done. The same would be true for quantitative questions. The increased familiarity should result in better results.

The surveys indicate that the students felt that the FA's helped more in learning concepts and definitions than solving numerical problems. About 62% agreed on the former while only 42% on the latter (Figure 10). The results of formative assessments indicate that they also perform better on the conceptual FA's than on the quantitative ones (Figure 6). In the interviews, however, when asked which type of question was more effective after seeing an example of each, more students chose the quantitative one. A few of the students indicated that having numbers made the problem more concrete and easier to grasp. As stated in the data analysis, a probable reason is that the conceptual questions require more thought to apprehend which may cause students to approach them more slowly and thoughtfully. This brings up a possible question to explore in the future. What is the link between how a question is posed and how successfully it is answered?

To answer the research question, conceptual FA's produced better results on conceptual questions on summative assessments and quantitative FA's produce better results on quantitative questions. The type of question influence how the students approach the problem which can influence the subsequent performance on the formative assessment.

3: Effects of Formative Assessments on Student Attitude towards Learning

Data from the surveys and interviews were used to answer this research question. When the students were asked directly if they enjoyed doing the formative assessments only 32% said they did. The highest percentage of students, 42%, were undecided. This indicates some ambivalence on the part of the students regarding the FA's.

The interview produced more positive feedback. Five of the six students said that they enjoyed physics and they like it better than biology or chemistry. In my teaching log I recorded five instances between two class periods when students asked if there was going to a formative assessment during the non-treatment. I also did not observe any signs of resistance when doing them during the treatment period. I did not notice any sighs nor did anyone ask if we could not do them. One of the students even said it was her favorite part of class. In the interviews when asked if they had any additional comments four of the students said that they liked them or that they were good. Two also indicated that they are helpful

Taking all these together it seems like the formative assessment have a positive impact on the students. To the students, the FA's were probably just part of the regular class routine. This wouldn't normally engender strong feelings either way. When their opinion is elicited, however, their opinion tends to be positive.

The interviews also sought to determine if the FA's change how students approach learning. The responses indicate that students only do minor corrective measures as result of the FA's. They tried to determine why they got the questions wrong through asking people around them, looking back at parts of the homework lesson, or through paying attention to the post-FA discussion. There wasn't any indication that they had a formal way to reflect on why they got questions wrong and how to avoid errors in the

future. None of the interviewees said that they do something similar to the FA's in other classes. When asked if they thought it would be helpful if they did, nobody said yes resoundingly. Most of them said probably either in a math or science course.

4: Effects of Formative Assessments on How Curriculum is Taught

To answer this research question entries from the teaching journal were analyzed. Part of the formative assessments process is to “close the loop” or adjust teaching/take action based on the results of the assessments. Therefore, there is a built-in mechanism to change how curriculum is taught. Most of adjustments were made during the discussion phase of the FA. During this phase the correct answer is explained usually by a student who was called upon. If a significant number of students chose a wrong answer extra time would be taken to explain why it is wrong. In the case where a majority of students got the answers incorrect, the concepts were retaught and extra practice would be given. This only occurred a few times during the study because there were only a few times when the results indicated significant reteaching and practice was needed.

Adjustments to the curriculum were also made in between classes. For example, one formative assessments revealed that students were still confusing the electric field for the electric force. It was noted in the teaching journal that “Students try to use the term “force field” and “They talk about the force at point P.” That information was kept in mind for the next class period which did not have formative assessments because it was during its non-treatment unit. In discussing the homework I made sure to make the distinction between the two concepts through questioning. There were several occasions where results from the FA's highlighted misconceptions and mistakes on the part of the students and were used to inform teaching in subsequent classes periods.

Information gathered from the formative assessments was also used with feedback gathered throughout the class period to make changes in the curriculum. In one class the formative assessments indicated some difficulty with finding the direction of the magnetic force on a current-carrying wire and moving charged particle. Reviewing the homework and additional in-class problems showed that more practice was needed. In the teaching journal it was stated that “Students still get the B \times Ir Claw and FBI rules confused. They need more mastery of each situation before doing a combined problem.” Given all those indicators, the homework was changed and the schedule for the unit was adjusted. In this situation the formative assessments served as an additional source of information that was used to address the needs of the class.

The data from the formative assessments will also serve to shape the curriculum in the future. Next year, part of the process in planning each unit will be to review this year’s formative assessment results to determine which concepts were particularly difficult for the students and which misconceptions were common. Lesson plans will be revised with that data in mind. For the FA questions that will be reused next year the data on those questions can serve as a baseline to determine if changes to the curriculum produced better results on the FA’s.

The formative assessments proved to be a valuable tool in determining how curriculum is taught. They provided feedback that was used immediately to correct misunderstandings. It was used within the lesson along with other forms of feedback to determine if more time was needed to learn a topic. It was used to anticipate misunderstanding in subsequent classes. Lastly, they will be used to shape the curriculum in the future.

VALUE

The research did not reveal new or original insight on the role FA's have in learning. It reaffirmed their function and purpose of being a tool for both the teacher and student to determine what was learned well and what needs remediation. It provided useful insight on the particular FA system developed in the particular context in which they were implemented.

First, the FA's are an invaluable tool in the flipped curriculum developed in the class. Since basics concepts are learned in the homework lessons, a feedback mechanism is needed to gauge student understanding before moving onto activities, problems-solving, and more complex ideas in the classroom. The multiple choice FA's administered through Google Forms fulfilled that role in the most efficient way possible. Instant feedback in form of percentages and graphs were a concrete way to see metrics of understanding of the students.

The finding of the study provided valuable insight on the limitations and possibilities of the FA's developed. First, the data showed that the FA's are not a good indicator of how the class will do in the summative assessment exam at the end of the unit. This puts the focus on the relearning and re-teaching after the FA's rather than on the learning done before it. The data also showed that the type of questions posed, conceptual or quantitative, can make a difference in student performance in that area. This can be valuable in the future. If a unit is known to be conceptually hard, for example, conceptual FA questions would be primarily given. The data from this study has also shown that the FA's can be helpful in identifying which units are difficult conceptually and quantitatively relative to one another.

In the context of the classroom environment the FA system developed in this study is a good first step in establishing a culture of feedback, reflection, and remediation. The formative assessments were shown to be positively received by the students. After several units the FA's became a part of the class routine. They were just something done in the beginning of the class. As noted, some students noticed their absence in the non-treatment unit. Assessment, feedback, and correction became part of the learning process within the class.

Inspired by the results of this study, a planned next step in the upcoming year is to develop a process of reflection. The goal is to establish a habit of self-reflection in learning. It will involve provide a time, a space, and a process where students ask themselves, "given the feedback, what did I do wrong, why, and how do I improve in the future?"

Another area that will be explored in the upcoming year is the optimum use of the data collected in the FA's. Technology has afforded the opportunity to collect information easily and quickly. The experience from the study has shown that a considerable amount of information can be collected on student learning. The process of doing the data analysis for this study and the fruitful insight it has provided has shown the potential of "big data" in the classroom. Extended reflection in the future on this study will involve determining ways that the data can be leveraged to affect student learning in new ways. An example would be using the formative assessment results to track individual students more closely throughout the unit. From the analysis of the outliers in the summative assessment analysis a possible process is to flag students who consistently answer the FA's incorrectly and who indicate they were not able to complete the

homework lesson. Those students can then be offered additional assistance during or outside of class.

In the larger context of the science department and school this study can hopefully can be model for other teachers to see how a formal formative assessment system can be implemented in a classroom and use some of its findings to judge whether it can be effective in their particular context. Even though a direct impact on summative assessments was not shown other benefits were demonstrated. A particular advantage of having such a system would be to have quantitative measures of student progress which can be used to adjust teaching. It can be a successful a mode of practice to focus on particular types of questions. Lastly, it can serve as a valuable feedback mechanism for students to assess their own learning.

This research project provided experiences, data, and insight that will be very valuable in the future. It can have an impact on a curriculum, class, and school level. But one area in which it definitely made an impact is on a personal level. Designing, implementing, and completing this action research has had a profound influence on how I see data and its role in program evaluation. Having gone through the rigors of collecting and analyzing data I understand the daunting task of determining the effectiveness of curricular programs and approaches. Good sounding ideas, anecdotal evidence, and gushing narratives of positive experiences now fall suspect in the absence of critical analysis. That applies to my impressions and experiences as well. Like a good scientist, I will now look for the data and scrutinize how it was collect and analyzed before holding strong opinions. But more importantly, I feel that I will be entering the next stage of my

teaching career, that of a teacher researcher. This will not mark the end of my action research but hopefully the beginning of many more.

REFERENCES CITED

- Conlin, L. D. (2009). Finding Method in Madness. *Physics Teacher*, 47(2), 125-126.
- Docktor, J. L., Strand, N. E., Mestre, J. P., & Ross, B. H. (2010). A Conceptual Approach to Physics Problem Solving. *AIP Conference Proceedings*, 1289(1), 137-140.
- Dufresne, R. J., & Gerace, W. J. (2004). Assessing-To-Learn: Formative Assessment in Physics Instruction. *Physics Teacher*, 42(7), 428-433. doi:10.1119/1.1804662
- Ernst, J. R. (2011). Implementing Frequent Assessment to Increase Student Performance Levels in a High School Chemistry Class. Retrieved 4/1/13 from <http://etd.lib.montana.edu/etd/2011/ernst/ErnstJ0811.pdf>
- Formative Assessment: Improving Learning in Secondary Classrooms*. (2005) Paris: OECD. Retrieved 4/1/13 from <http://www.oecd.org/dataoecd/19/31/35661078.pdf>
- Hewitt, P. G. (2000). Conceptually speaking. *Science Teacher*, 67(1), 41.
- Hung, W., & Jonassen, D. H. (2006). Conceptual Understanding of Causal Reasoning in Physics. *International Journal of Science Education*, 28(13), 1601-1621. Doi: 10.1080/09500690600560902
- Walvoord, B. E., Bardes, B., & Denton, J. (1998). Closing the Feedback Loop in Classroom-Based Assessment. *Assessment Update*, 10(5), 1.

APPENDICES

APPENDIX A

MID-SEMESTER SURVEY QUESTIONS

Mid-Semester Survey

Please answer the following questions as honestly as possible so that I can make this class the best it can be for you.

* Required

1. I do the homework lessons on time *

- all the time
- most of the time
- half of the time
- less than half of the time
- rarely

2. Explanation or elaboration to response for #1 (optional)

3. The homework lessons help me understand physics *

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

4. The Conceptual Questions (red slides) help me understand physics concepts *

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

5. The Problem Solving Questions (grey slides) help me understand how to the physics problems. *

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

6. How many tutorial videos have you watched? *

- 0
- 1-2
- 2-4
- most
- all

7. The tutorial videos help me understand how to solve physics problems.

Answer if you have watched at least 1 tutorial video

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

8. Please list things, if any, that were particularly helpful about the homework lessons and/or tutorial videos

9. Please list things, if any, that needs improvement related to the homework lessons and/or tutorial videos

10. The formative assessments at the beginning of the class help me understand physics *

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

11. Explanation or elaboration to response for #10 (optional)

12. The labs and activities help me understand physics *

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

APPENDIX B

PRE-TREATMENT SURVEY QUESTIONS

Pre-Treatment Survey

For each item below, indicate whether you Strongly Agree, Agree, are Undecided, Disagree, or Strongly Disagree. Please circle your response.

Participation in this research is voluntary and participation or non-participation will not affect your grade or standing in any way.

1. The formative assessments reinforce what I learn from the homework lessons.

Strongly Agree Agree Undecided Disagree Strongly
Disagree

2. I am able to do the formative assessments successfully.

Strongly Agree Agree Undecided Disagree Strongly
Disagree

3. The formative assessments are challenging.

Strongly Agree Agree Undecided Disagree Strongly
Disagree

4. The formative assessments are a key part of my learning in physics.

Strongly Agree Agree Undecided Disagree Strongly
Disagree

Explain your answer

5. The formative assessments bring up questions and situations that I wouldn't consider with the HW lesson alone.

Strongly Agree Agree Undecided Disagree Strongly
Disagree

6. The formative assessments help me realize what I don't know or need to work on.

Strongly Agree Agree Undecided Disagree Strongly
Disagree

7. The formative assessments help me understand physics definitions and concepts.

Strongly Agree Disagree	Agree	Undecided	Disagree	Strongly
----------------------------	-------	-----------	----------	----------

Explain your answer

8. The formative assessments help me successfully learn how to solve numerical problems.

Strongly Agree Disagree	Agree	Undecided	Disagree	Strongly
----------------------------	-------	-----------	----------	----------

9. I enjoy doing the formative assessments.

Strongly Agree Disagree	Agree	Undecided	Disagree	Strongly
----------------------------	-------	-----------	----------	----------

10. Seeing the response summary is a useful part of the assessment.

Strongly Agree Disagree	Agree	Undecided	Disagree	Strongly
----------------------------	-------	-----------	----------	----------

APPENDIX C

CORRESPONDENCE BETWEEN SURVEY AND RESEARCH QUESTIONS

Research Question	Survey Prompt
<p>How do conceptual formative assessments affect conceptual understanding? Computational problems-solving?</p>	<p>1. The formative assessments reinforce what I learn from the homework lessons.</p> <p>2. I am able to do the formative assessments successfully.</p> <p>3. The formative assessments are challenging.</p> <p>4. The formative assessments are a key part of my learning in physics.</p> <p>5. The formative assessments bring up questions and situations that I wouldn't consider with the HW lesson alone.</p> <p>6. The formative assessments help me realize what I don't know or need to work on.</p>
<p>How do conceptual formative assessments affect conceptual understanding? Computational problems-solving?</p>	<p>7. The formative assessments help me understand physics definitions and concepts.</p>
<p>How do problem-based formative assessments affect conceptual understanding? Computational problem-solving?</p>	<p>8. The formative assessments help me successfully learn how to solve numerical problems.</p>
<p>How do formative assessment affect student's attitude on learning?</p>	<p>9. I enjoy doing the formative assessments.</p> <p>10. Seeing the response summary is a useful part of the assessment.</p>

APPENDIX D

PRE-TREATMENT INTERVIEW QUESTIONS

Pre- Treatment Interview Questions

Participation in this research is voluntary and participation or non-participation will not affect your grade or standing in any way.

1. What are your thoughts on science in general? Physics in particular?
1. Do you think the formative assessments are an effective way to test your understanding in physics? Why?
2. Do you feel your performance on the assessments accurately reflect your understanding of the material? Explain.
3. Do you feel that the number of questions is optimal? Do you think more or less questions would be more helpful? Why do you think that?
4. Do you feel that the difficulty of the questions is at the right level? Why do you think that?
5. What are three ways you think they can be improved?
6. Do you think having the assessments make a difference on your learning of the material? Explain. Can you give an example?
7. Do you think having the assessments makes a difference on your performance on the quizzes and exams? Explain.
8. Do you change what you study or how you study as a result of your performance on the formative assessments? Elaborate.
9. Do you do similar assessments in other classes? If so which? How would you compare them? Strength and weaknesses in each?
10. Think about another class you are taking right now. Do you think having formative similar assessments in those classes would be helpful? Why?
11. There were different types of questions were asked in the assessments? Here are some examples:

(A conceptual question example and problem-solving example previously given will be provided)

Do you feel one type of question is more helpful than another? Why?

12. Are you comfortable doing the formative assessments?
13. How do you feel doing them compared to doing quizzes and exam? Explain.

APPENDIX E

CORRESPONDENCE BETWEEN INTERVEIW AND RESEARCH QUESTIONS

Research Question	Interview Prompt
<p>How do formative assessments affect learning physics?</p>	<ol style="list-style-type: none"> 1. Do you think the formative assessments are an effective way to test your understanding in physics? 2. Do you feel your performance on the assessments accurately reflect your understanding of the material? 3. Do you feel that the number of questions is optimal? Do you think more or less questions would be more helpful? 4. Do you feel that the difficulty of the questions is at the right level? 5. How do you think they can be improved? 6. Do you think having the assessments make a difference on your learning of the material? 7. Do you think having the assessments makes a difference on your performance on the quizzes and exams? 8. Do you change what you study or how you study as a result of your performance on the formative assessments? 9. Do you do similar assessments in other classes? If so which? How would you compare them? Strength and weaknesses in each? 10. Think about another class you are taking right now. Do you think having formative similar assessments in those classes would be helpful? Why?

<p>How do conceptual formative assessments affect conceptual understanding? Computational problems-solving?</p> <p>How do problem-based formative assessments affect conceptual understanding? Computational problem-solving?</p>	<p>11. There were different types of questions were asked in the assessments? Here are some examples:</p> <p>(a conceptual question example and problem-solving example previously given will be provided)</p> <p>Do you feel one type of question is more helpful than another? Why?</p>
<p>How do formative assessment affect student's attitude on learning?</p>	<p>12. Are you comfortable doing the formative assessments?</p> <p>13. How do you feel doing them compared to doing quizzes and exam?</p>

APPENDIX F
IRB EXEMPTION APPROVAL



INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 00000165

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

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

MEMORANDUM

TO: Jeffrey Noblejas and Walt Woolbaugh

FROM: Mark Quinn, Chair *Mark Quinn et al*

DATE: December 5, 2013

RE: "The Effectiveness of Conceptual and Problem-Solving Based Formative Assessments in the High School Physics Curriculum" [JN120513-EX]

The above research, described in your submission of December 5, 2013, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

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

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