THE EFFECTS OF INSTANT LAB FEEDBACK THROUGH THE USE OF EXCEL

SPREADSHEET ON STUDENT LEARNING IN

HIGH SCHOOL PHYSICS

by

John L. Wilkie

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

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2016

ii

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In my physics classes, I found students disengaged from laboratories. They work through the procedure and analysis, but upon receiving back graded labs they often disregard the feedback since the grade has already been given. I believed this put them at a disadvantage on assessments since they would never correct their misconceptions. I wanted to know if students would learn better if they received instant feedback during the labs so that they could correct their misconceptions when it was relevant to their grades. My focus question was to see the effect would be of immediate feedback during laboratory exercises via Excel spreadsheet. As a subquestion, I wanted to know specifically what effect this immediate feedback would have on their summative assessment scores.

Research was performed on two sections of physics students over four units of study. Two of the units had laboratories in which feedback would be given in the traditional manner in which students complete a lab and it is them graded and returned to them to review. Two of the units had laboratories in which feedback is automatically given via a preprogrammed Excel spreadsheet which would return a smiley face for correct cell values and a frowny face for incorrect cell values.

To study the effects on student attitudes I had students take a survey on their feelings about physics, laboratories, and laboratory feedback both before and after the action research. Students were also chosen at random and asked to participate in a short interview to probe deeper into student attitudes. To measure student performance on summative assessments, unit concept inventory test was given before and after each unit. Both survey responses and test scores were analyzed to see if there was any significant difference due to the instant feedback given during laboratories.

The results of the analysis showed that though students found the laboratories enriching in their visualization of the concepts, there was inconclusive evidence to suggest and difference in student attitudes or in student test performance. During interviews students expressed that they found laboratories not very useful for the purposes of studying for quizzes and tests, as laboratories seem to require a different set of skills than those normally required for a classical multiple choice or written assessment.

INTRODUCTION AND BACKGROUND

I currently teach at West Anchorage High School (West High) located in

Anchorage, Alaska. The 2015-2016 enrollment of West High is approximately 1,905

students with racial demographics of 37% Caucasian, 22% Asian or other Pacific

Islander, 15% Hispanic or Latino, 8% Alaska Native, and 7% African American.

Anchorage is Alaska's largest city with approximately 300,000 people (Anchorage

School District, 2015; Anchorage: Economy, n.d.). The largest source of employment in

Anchorage stems from government jobs, since the city is where most Alaska senators and

representatives have their offices; international shipping though Ted Stevens

International Airport and the port of Anchorage due to their being waypoints to and from

Asia and the lower 48 states; and the oil industry whose offices largely reside in the city.

Finally, it is worth noting that just north of the city is Joint Base Elmendorf-Richardson

with over 5300 active military personnel, and about one third of the population of

Anchorage are veterans (Joling, 2014).

In my years as a teacher, especially my time here at West High as a physics

teacher, I have been frustrated by students who make the same mistakes over and over

again on lab reports and assignments. Furthermore, when these students complete a high

stakes summative assessment, they make the same mistakes and further lower their

grades. This may be due to difficulty connecting feedback on the labs with memories of

a lab that happened a week prior. However, the lab coordinator here at the University of

Alaska at Anchorage has been experimenting with Excel lab spreadsheets that self-grade

and offer immediate feedback to students. The spreadsheets can be programmed with

cells that look at what values the student inputs or how they calculate a value. Then

feedback is given about whether their results are correct or incorrect, what the correct

response should be, or even an explanation of why their input is incorrect. I find the most

useful characteristic of the feedback presented by this spreadsheet is that it is

instantaneous. Students do not have to receive a low grade before they realize they have

a misconception, and given the proper feedback, they can correct their mistake and learn

the concept correctly the first time. I believe this is of great benefit to student learning

and can save both the student and teacher valuable time that would have been spent

relearning and re-teaching key concepts.

There are conflicting studies to support the benefits of both immediate feedback

and delayed feedback, though both gave feedback in different amounts of time as well as

over different types of assessments. Since my situation and assessment format is

different, I wanted to conduct my own study to see what sort of feedback would be more

effective in increasing student learning. This lead to the focus question of the study,

What is the effect of immediate feedback during laboratory exercises using Excel

templates? The subquestion is, *What is the effect on student performance of summative*

laboratory skill assessments, i.e. midterm and final exams?

CONCEPTUAL FRAMEWORK

Many studies support the idea that instructor feedback is beneficial to student

learning compared to no feedback at all (Bälter, 2013; Bangert-Drowns, 1991;

Sassenrath, 1965). Elliot, Murray, and Ward (2003) state that feedback must inform the

student whether they are right or wrong and provide information so that they may correct

themselves. However, there are different types of feedback, different time frames in which it may be delivered, and different methods of delivery. The performance level of the student population, as well as their exposure to the concepts being assessed, determines the type of feedback that should be used to assess those students (Shute, 2007).

Feedback can be categorized into five types: no feedback, verification, correct response, correct response with explanation, and correct response with interactive teaching. The *no feedback* category is the complete lack of feedback after an assessment or exercise is given. Verification feedback, otherwise known as *knowledge of results*, is a limited form of feedback in which the students only know of the correctness of their responses or their overall grade but are not informed of the correct responses. Correct response feedback, otherwise known as *knowledge of correct response* feedback, is feedback in which the student knows the correct response. The last two types of feedback that Roper lists fall in a general category that Shute would call *elaborated feedback*. In elaborated feedback, students are given explanations about why a certain response is correct or incorrect and provided with an opportunity for reteaching. Each type of feedback has its benefits. In general, students perform better when the feedback is more comprehensive (Roper, 1977).

In a study by Bälter, Enström, and Klingenberg (2013), students were surveyed about their opinions of short, computer-based quizzes with minimal feedback indicating only whether responses on short assessments were correct or incorrect (*knowledge of results* feedback). The results of the study revealed nearly all students surveyed

appreciated the feedback and about 20% reported adjusting their study habits as a result. It was concluded that even *knowledge of results* feedback, though limited, was better than no feedback at all.

Bangert-Drowns, Kulik, Kulik, and Morgan (1991) conducted a meta-study of dozens of published papers to analyze the effects of feedback and yielded two important conclusions. The first is that feedback of any kind ended up producing a greater effect than no feedback at all. The second conclusion was that in studies where *knowledge of results* feedback was compared to *knowledge of correct response*, the latter was found to be more effective in helping students correct themselves. In a more recent article, Epstein, Lazarus, Calvano, Matthews, Hendel, Epstein and Brosvic (2002) found that students' active involvement in the assessment process played a crucial role in learning new concepts. Students who had to correct each quiz item immediately after answering were able to more successfully acquire new information, incorporate new information into what they had already learned, and retrieve the correct information in post assessments.

A meta-study done by Shute (2007) for the Educational Testing Service found studies concluding that *knowledge of correct response* and elaborated feedback are more instructive to students who are still trying to acquire new information (Bangert-Drowns et al., 1991; Corbett & Anderson, 2001; Gilman, 1969; Mory, 2004). If the feedback is too limited, Shute found studies indicating that it can cause students to lose confidence and overthink the feedback (Kluger & DeNisi, 1996; Sweller et al., 1998) and perhaps even lose motivation to fix their misconceptions (Ashford, 1986; Corno & Snow, 1986).

Another study (Hanna, 1976) indicated KR can be beneficial to higher achieving students who will search for the correct answer themselves or to average students reviewing previously taught concepts.

Another factor upon which feedback seems to be heavily based is the time frame in which feedback is given. There are two main types of timing concerning feedback: immediate and delayed. Immediate feedback can mean that the feedback is given to the student immediately after assessment, or with modern technology, students can receive feedback after each question is answered without compromising the rest of the assessment. Delayed feedback can mean anything from 20 minutes of delay after the assessment to as much as a few days (Shute, 2007).

In a study by Sturges (1978) on a computer-assisted assessment, retention 1-3 weeks later was found to be more improved in cases when the feedback was delayed 20 minutes or 24 hours than with immediate item-by-item feedback. The study cites another by Kulhavy and Anderson (1972) which posits an interference-perseveration hypothesis which states that students' memories of earlier responses cloud their acquisition of new knowledge. There was no difference in retention between the 20 minute and 24 hour delay. Therefore, it seemed unlikely that this hypothesis would be correct nor could the study say for sure whether delayed feedback was the lone cause of the increased retention. The study also showed that student confidence on new responses increased with delayed feedback.

Conversely, a meta-study by Kulik and Kulik (1988) stated that delayed feedback might only be effective in laboratory type situations, e.g. when someone is given a list of

words or symbols to remember. In real world, classroom environments, immediate

feedback was more effective than delayed feedback. In a study on the effect of timing of

feedback, Dihoff, Brosvic and Epstein (2003) gave students five quizzes and then

reassessed them two weeks later in order to measure improvements. The researchers

found that students showed the highest recall, the highest identification of their initial

quiz responses, the most confidence in their answers, and the least bound to their

previous misconceptions when immediate feedback was provided after each test item as

opposed to when feedback was given at the end of the quizzes or 24 hours later.

More recent findings summarized by Shute (2007) found that it may not be timing

that matters over all, but rather the instructional purpose of the task being completed and

the academic level of the student completing it (Schroth, 1992; Corbett and Anderson,

2001). In these studies it was found that delayed feedback might be better for transfer

and concept building, but that immediate feedback is better suited for learning steps and

procedures. Similarly Clariana (1999) stated that difficult tasks might require immediate

feedback, but simple tasks could be better served with delayed feedback.

The method in which feedback is given can also affect a student's ability to

assimilate the new information and correct their misconceptions on future assessments.

In a study about an automated assessment program by Wang, T., Su, Ma, Wang, Y., and

Wang, K., (2011), four requirements of an assessment program were outlined. The study

stated that an automated assessment program should be

- thorough in its assessment

- check not only for the correct response but also the correct method

- should be able to point out errors and assess partial credit so as not to dishearten students, and

- provide "timely and corrective feedback" (p. 221)

Qualitative analysis of the assessment program revealed positive approval from both students and teachers.

In a study of the use of interactive spreadsheets in physics education, Wagner (2007) found that the formative feedback capabilities of the spreadsheet were very versatile. The article noted that, although the immediate feedback provided by the spreadsheet might shift student focus from learning the concept to merely getting the right answer, the spreadsheet restrictions along with proper instruction motivated students to seek understanding over mere points on an activity. A study by Lai and Hwang (2015) observed the effects of a spreadsheet based program which allowed elementary geography students interactive manipulation of factors in a simulated climate. The researchers found that one of the major benefits was that the cognitive load the students had to endure in the exercise was reduced, allowing them to focus less on procedure and more on investigating higher order concepts.

The use of technological tools and assessments are becoming more widely used in the classroom. In two studies by Hennessy et al. (2007), the authors studied four teachers using multimedia education in their classrooms, and eleven teachers from eight schools using interactive educational tools. The findings of these studies suggest that teachers are starting to move away from exclusively traditional hands on laboratories and activities and are dabbling more into multimedia techniques. The researchers also found

that the use of this technology can surpass simple demonstration and simulation. Proper application of these tools can help students explore *What If?* questions and reinforce the real world concepts they are attempting to learn.

As technology evolves the methods by which educators assess and give feedback to their students will undoubtedly change. Though there are several papers studying the benefits of different types any methods of feedback, the reality was that some methods of feedback were less practical than others. More comprehensive methods of feedback that were laborious and time consuming for educators in the past are becoming more accessible as different tools are developed for classroom use. The timing of feedback is also shifting as computer programs and online applications allow for instantaneous feedback. As our ability to assess students develops, it is necessary to continue to research the effectiveness of these methods and how they can be best used to enhance student learning.

METHODOLOGY

I began my research by first asking my high school physics students to complete the Pre-Study Attitude Survey. The survey measured student attitudes toward my physics class, labs, and feedback on labs and how they use the feedback to prepare for summative assessments such as tests and quizzes (Appendix B). Before each unit a Concept Inventory Test was given to assess students' grasp of unit concepts and practical abilities such as solving for a physical property using equations (Appendix C). The research methodology for this project received an exemption by Montana State University's

Institutional Review Board and compliance for working with human subjects was maintained (Appendix A).

Due to the small number of students in each of my two physics classes (12 students and 18 students) I decided both classes would undergo nontreatment and treatment methodologies. So as not to adversely affect student attitudes toward periods of nontreatment methodology after exposure to the treatment methodology, I studied student gains across two units utilizing nontreatment methodology first, and then across two units utilizing the treatment methodology. Nontreatment and treatment units were both similar in the style of lectures, difficulty of homework, laboratory procedures and data analysis the students had to perform, and the style and difficulty of quizzes that were periodically administered. During laboratories, students were required to input data and perform calculations using an Excel template that was provided to them.

During the nontreatment units, the Excel template provided no feedback and was simply a method of recording data and performing calculations. During the treatment units, the Excel template was much more informative, offering students a "smiley face" in the adjacent cell if they calculated an appropriate value or a "frowny face" along with feedback if their value was incorrect or not calculated using Excel (Figure 1).

		Trial 1	Trial 2	Trial 3	Average	Symbolic Equation		Trial 1	Trial 2	Trial 3	Average	Symbolic Equation
Part A:	Finding Focal Length of the Converging Lens Using a Far Away											
	Position of Converging Lens on Meter Stick:	15	10	13	12.67	=AVERAGE(Trial 1 + Trial 2 + Trial 3)		x	x	x	:-D	:-D
	Position of Screen on Meter Stick:	25	19.5	22.5	22.33	(Trial 1 + Trial 2 + Trial 3) / 3		:-D	:-D	:-D	:-D	:-D
	Object Distance:	1000	Far??	100	550.00	Directions say to enter 1000 cm for the distance		:-D	:-(	:-(	:-D	:-D
	Image Distance:	10	9.5	9.5	9.67	Position of the screen - position of the lens		:-D	:-D	:-D	:-D	:-D
	Calculated Focal Length (f):	9.90	#VALUE!	8.68				:-D	#VALUE!	:-D		Don't forget a symbolic equation
	Result of Covering Up Half the Lens?										Try it out!	

Figure 1. Treatment unit excel template.

The Excel template also awarded points to students as sections of the spreadsheet became more complete. At the end of each unit, the students were given same Concept Inventory Test to compare normalized gains. Throughout the course of the nontreatment and treatment units, I kept a journal of observations and reflections that I recorded shortly after the lab was completed.

After both nontreatment and treatment units were completed, students were asked to complete the Post-Study Attitude Survey. In this survey I asked many of the same questions regarding their attitudes toward physics, labs, and feedback on labs and how they use it to prepare for quizzes and tests. I also asked students if they thought the instant feedback provided by the Excel template helped. Finally, I then chose six students at random and asked them to participate in a Post-Study Interview in which I asked questions to gain more elaborate feedback about students' attitudes towards the physics class, laboratories, and the Excel spreadsheets used during the treatment phase (Appendix E).

The normalized gains of this Concept Inventory Test were calculated and analyzed using a Wilcoxon rank-sum test to determine if there was significant difference in median student performance. These results of the survey were analyzed with a chi squared goodness of fit test to determine if there was a significant shift in students' agreement with the statements due to the treatment in the study. Due to the varying attendance of students throughout the semester, the normalized gains for each unit were calculated with varying numbers of students. The sources of data used to help answer my focus question diagramed below in a Data Triangulation Matrix (Table 1).

Table 1.
Data Triangulation Matrix

Focus Question	Source 1	Source 2	Source 3
What is the effect of immediate feedback during laboratories using Excel templates?	Concept Inventory Test	Pre-Study and Post-Study Attitude Survey	Post-Study Interview Teacher Journal
What is the effect on student performance of summative laboratory skill assessments, i.e. midterm and final exams?	Concept Inventory Test	Pre-Study and Post-Study Attitude Survey	Post-Study Interview

DATA AND ANALYSIS

The normalized gains of the Concept Inventory Test were calculated for each unit, and only for students who were present at the time of both assessments. Due to inconsistent student attendance, the normalized gain distributions for the first and second comparison and first and second treatment units were comprised of different numbers of students, $N=28$, $N=20$, $N=24$, $N=24$ respectively. The median normalized gains from first and second comparison and first and second treatment units was 10%, 15%, 0%, and 15% respectively (Figure 2). The treatment distributions had a larger overall range and interquartile range than the comparison units such that the treatment distributions contained the two highest normalized gains of 50% and 70%, but also the lowest normalized gain which was -40% during the first treatment unit.

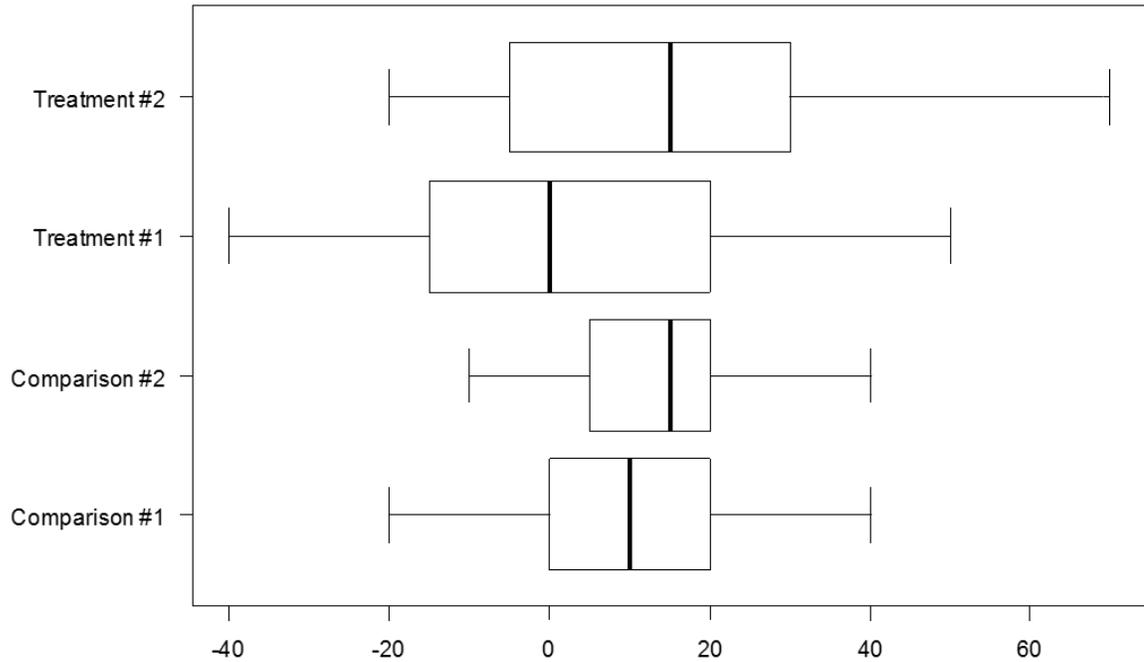*Figure 2*. Distribution of normalized gain percentages from student pretest vs posttest for the first and second comparison units and first and second treatment units, (*N=28, N=20, N=24, N=24* respectively).

To further condense the data, normalized gains were averaged for students who completed both comparison pretests and posttests and also for students who completed both treatment pretests and posttests (Figure 3).
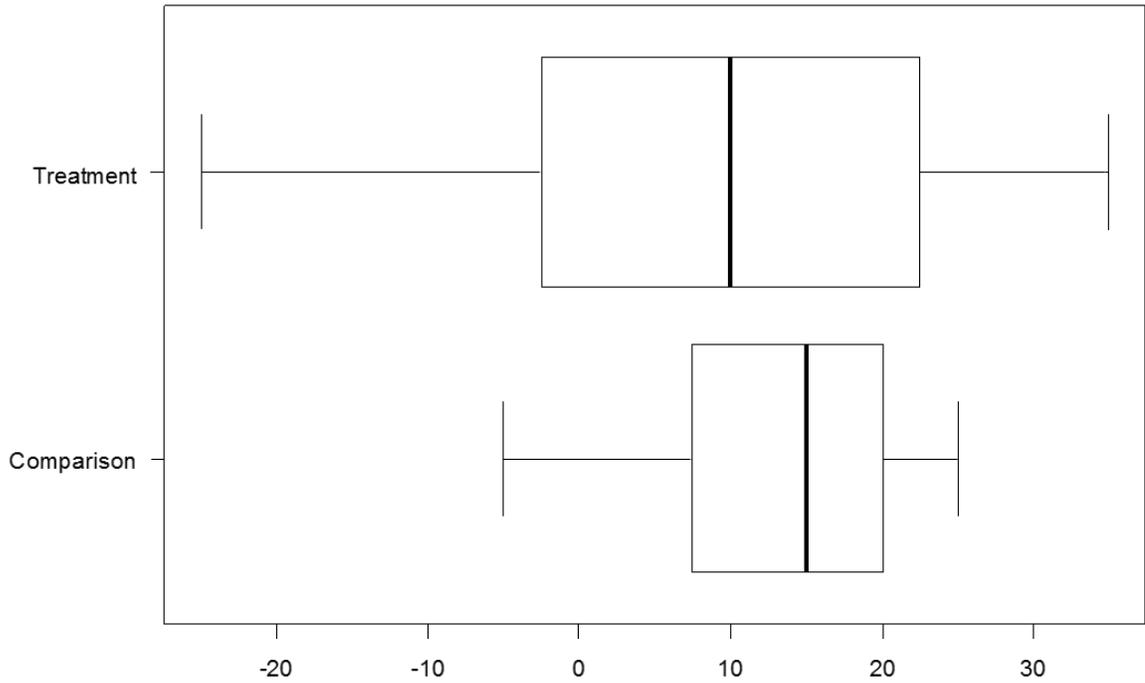
Figure 3. Average of normalized gain percentages of students who completed both units of comparison or treatment pretests and posttests, (*N=20 and N=*19 respectively).

The range of the treatment gains was larger than that of the comparison gains, such that

the treatment gains yielded both a higher maximum and a lower minimum value. Though

the median comparison gain of 15% was higher than the treatment gain of 10%, a

Wilcoxon-Rank Sum test was performed on the two distributions, yielding a p-value of

0.765, which suggests that there is insufficient evidence to conclude that these two

medians are statistically different from each other.

Before and after the research began students were asked to complete the Attitude

Survey to gauge their attitudes towards physics, labs, and particularly lab feedback before

and after the action research. When students were initially asked if *the feedback (they*

received) on labs helps (them) understand the topics they are learning in physics, 48% of

students reported that they disagreed or strongly disagreed that labs helped them learn

physics better and 52% agreed or strongly agreed that labs helped them. After the research was conducted 36.7% of students disagreed or strongly disagreed that labs helped them and 67.5% agreed or strongly agreed that labs helped them learn physics concepts better (Figure 4). A chi-squared analysis yielded a p-value of less than 0.01, which indicated that there was strong evidence of a significant difference in the distributions of responses from before to after the treatment.

Figure 4. Distribution of student responses to survey question *The Feedback I Receive on Labs Helps Me Understand the Topics I Am Learning in Physics*, (*N*=25 for pretreatment and *N*=19 for posttreatment).

Chi-squared analysis of all other Attitude Survey items yielded p-values too high to conclude if there is any significant difference in the distribution of responses.

Students were asked if they *currently like physics* and from before and after the research and it would seem to the eye that student attitudes toward physics are more negative after the action research. Before research 4% of students *strongly disagree* that they like physics and 32% *strongly agree* they like physics, and after treatment *strongly disagree* that they like physics and 15.8% *strongly agree* they like physics and 31.6%

disagree they like physics (Figure 5). A chi squared goodness of fit test was done to see

if the distributions were significantly different and was returned a value of p-value of

0.57, indicating that I cannot conclude they are significantly different.

Figure 5. Distribution of student responses to survey question *I Currently Like Physics*, (*N*=25 for pretreatment and *N*=19 for posttreatment).

Students were asked after the research was done if *(they) found the automated*

spreadsheet helpful with learning physics concepts. The response to that question

showed that 37.8% of students disagreed or strongly disagreed and 63.2% of students

agreed or strongly agreed that the automated spreadsheets helped them learn physics

concepts (Figure 6). A chi squared goodness of fit test was run with the null distribution

that each responses was equally likely, and p-value of 0.96 was returned indicating that

there is little evidence to show these responses are significantly different from a random
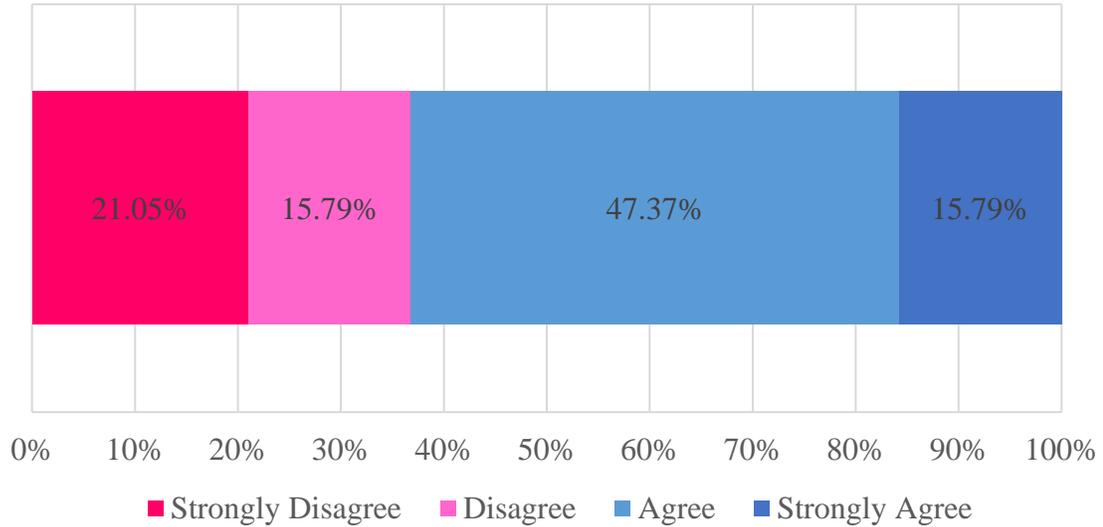
selection of responses by students.

Figure 6. Distribution of student responses to survey question *I Found the Automated Spreadsheet Helpful with Learning Physics Concepts*, (*N*=19).

Students were also interviewed at the conclusion of the research. When students were asked of their opinions of the automated spreadsheets one student responded positively that they were, "…very, very helpful. We didn't have to ask [the teacher] if something was wrong." That student added a caveat that, "…[the feedback] backfires if the spreadsheet itself was wrong, but in that case we could just ask [the teacher] and get help that way." Another student stated positively that the automated feedback, "…made me feel good when I got a smiley face."

INTERPRETATION AND CONCLUSION

The most telling details of the normalized gain distributions was the larger range with the treatment data. The use of interactive Excel spreadsheets would appear to have had very different effects on different students, with some finding them helpful in reinforcing concepts and some finding them confusing and counterproductive. This assumes that the comparison and treatment units were the same in every other respect,

save the feedback provided by the Excel spreadsheets, but that is not the case. In addition

to the units being given different times of the academic year, the concepts that were being

taught in each unit were very different. Some units, like the light unit and the lens unit

included some of the same ideas, but that is not so between the energy and momentum

unit and the light unit. This uncertainty was only amplified by the small sample sizes

available to me even before students missed effect of the labs due to abscesses or

lethargy. Statistically this uncertainly was confirmed with a relatively high p-value of

0.765 from a Wilcoxon Rank-Sum test. This prevents me from concluding if there was

any difference in student performance on summative assessments due to the use of

immediate feedback via interactive Excel spreadsheets. The only thing which might be

concluded is that the new approach worked better for some students than others.

Student attitudes appeared at first to be more negative after than before in some

cases. It is difficult to say whether this was caused by my treatment or whether it had to

do with it being the end of the year. It might also have been due to the transition into the

waves and electrostatics portion of physics, which is often a difficult transition for many.

Most of the chi squared analyses yielded p-values that were too high to conclude that the

pretreatment attitudes were different from the posttreatment attitudes. The exception

being that over the course of treatment students agreed more that labs helped them learn

physics concepts better. Students also has seemingly positive responses to the

spreadsheet with automated feedback, though the analysis failed to provide evidence that

the positive response was statistically significant.

Students were also interviewed at the conclusion of the research and the responses reflect the students' survey responses. Students generally responded that though labs helped them "understand concepts in general" they "hardly check" the feedback from the labs to study for assessments. One student responded that labs "make me confused. Like going over it in class is different from the feedback you give us." This might be because the assessments generally measured student ability to solve traditional physics problems that are either conceptual or algebraic in nature, and though labs might be an enriching activity, they are not as useful when students have to perform on quizzes or exams.

Despite these apparent shortcomings of laboratories, students seemed to enjoy the automated spreadsheets, stating that they were more efficient and freed them from having to check with me if their answers were correct. Though there were a few programming errors in some cells of the spreadsheets, students felt they could then ask me for help and still rectify their misconceptions before the end of the class period. Students also seemed to enjoy the achievement aspect of satisfactorily completing parts of the procedure, adding an element of gaming to their laboratory experience.

VALUE

My focus question about the effect of immediate feedback via automated Excel spreadsheet was answered with both qualitative and quantitative data. Qualitatively, most students seemed to enjoy the validation they received when performing labs with the automated spreadsheets. The instant feedback allowed them to know immediately whether they were doing that particular section correctly and gave them affirmation. At the same time, it removed that uncertainty that struggling students must feel as they make

their way through a lab procedure and analysis, never being sure whether they are doing

it correctly or whether they are falling further down a rabbit hole of misconception.

Students also reported that the labs helped them visualize the concepts we have been

learning better, which to me is the original purpose of laboratories in the first place and

would have occurred with or without the automated spreadsheets as a guide.

The answer to my subquestion of how automated feedback on labs affects

students' summative assessment scores could not be answered with any certainty. There

was no statistically appreciable gain or loss in student performance on quizzes.

Considering the larger variation in normalized gains of the treatment units, it would seem

that the automated spreadsheets might have helped some students but might have further

confused others. This could also have been due to the material being covered, as sound

waves and optics were very different concepts from the mechanics units students had

been used to learning up to that point. It must be always noted that any experiment's

conclusions can only be supported through repetition, and in the future I would like to do

further research on instant feedback via the automated spreadsheet for my other units.

For myself, there was a logistic value using the automated spreadsheets. The

automated spreadsheets were wonderful tools in the hectic lab environment, as it took

some of the pressure off of me as the teacher to help confused students and put that onto

the students. With students correcting themselves, I was able to focus on students who

had true misconceptions, and were not just making mathematical errors. I also found

grading the labs much easier, as students had already corrected their mathematical

calculations by the end of the lab. The strongest feedback from students was that of the

usefulness of the lab and how students actually study. In general, almost all students

interviewed found that the labs were not useful for assessments.

The greatest value that was gained during this research was to my own

development as a physics teacher. Students in the interviews echoed that though the labs

were valuable to helping them visualize the concepts they were learning, they would not

sure them to study for quizzes and tests. The most frequent reason given was that the

labs did not make them practice the same set of skills that they needed for summative

exams. This lead me to ask myself important questions: should I be adjusting my

laboratories to better reflect the summative assessments, or should I be adjusting my

assessments to better reflect my laboratories? What kind of skills do I want my students

to possess when they have successfully completed my class? Should they be able to

calculate physics problems or should they be able to perform lab measurements and

calculations? The answers to these questions were the key to deciding what kind of

teacher I wanted to be.

When I reflected on those questions, it became clear that I wanted my students at

least to understand how scientific knowledge is synthesized and how the technological

developments had made and were still making come to be a reality. At best I wanted my

students to be able to enter the world with the ability to consider an unknown, to design

an experiment, to gather data, to analyses the data, and finally to determine if there is a

relationship, in short, to be able to synthesize new knowledge. I found it no coincidence

that for me the clearest path to that end would be teaching my students science through

inquiry, the pedagogy that was emphasized in the MSSE program at Montana State.

Though some of my lessons and laboratories were open ended, my curriculum itself was

not dedicated to having students learn through inquiry. When students were asked to do

such activities without preparation, they were unprepared with how to proceed and some

would just shut down. My next classroom experience will be to transition more fully into

an inquiry teacher and have a curriculum that teaches inquiry from the very start so that

students are prepared for lesson and labs that use inquiry to develop concepts. Though

this transition will be a long process, it will take me closer to the goal of the teacher I

would like to become

REFERENCES CITED

Anchorage: Economy. (n.d.) Retrieved April 11, 2015, from http://www.city-data.com/us-cities/The-West/Anchorage-Economy.html

Anchorage School District. (2015, October 30). Ethnicity Report. Retreived November 30, 2015, from http://www.asdk12.org/media/anchorage/globalmedia/documents/demographics/Ethnicity%202015-2016.pdf.

Holmes, E., Dong, Y.F., Brubaker, B., Mongeau, G., Paquette, J., Bayne, E., (October 1, 2014). *UNIVERSITY of ALASKA ANCHORAGE Fact Book 2013-14*. Anchorage: Office of Institutional Research. Retrieved April 10, 2015, from http://www.uaa.alaska.edu/ir/publications/upload/2014-Fact-Book-PDF-October-1-2014.pdf

Joling, D. (2014, January 8). Army plans to reduce number of Alaska soldiers; 780 from JBER. Associated Press. Retrieved April 10, 2015, from http://www.adn.com/article/20140108/army-plans-reduce-number-alaska-soldiers-780-jber

Kulik, J.A., Kulik, C.C. (1988). Timing of Feedback and Verbal Learning. *Review of Educational Research*, *58*, 1, 79-97.

Mullet, H.G., Butler, A.C., Verdin, B., von Borries, R., Marsh, E.J. (2014). Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately. *Journal of Applied Research in Memory and Cognition*, *3*, 222–229.

U.S. Census Bureau: State and County QuickFacts. (n.d.) Retrieved April 10, 2015, from http://quickfacts.census.gov/qfd/states/02/02020.html

Ashford, S. J. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management Journal, 29*, 465–487.

Bälter,O., Enström, E., Klingenberg, B. (2013). The effect of short formative diagnostic web quizzes with minimal feedback. *Computers & Education, 60*, 234-242.

Bangert-Drowns, R.L., Kulik, C.C., Kulik, J.A., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research, 61*(2), 213-238.

Clariana, R. B. (1999, February). *Differential memory effects for immediate and delayed feedback: A delta rule explanation of feedback timing effects.* Paper presented at the annual Association of Educational Communications and Technology convention, Houston, TX.

Corbett, A. T., & Anderson, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of ACM CHI 2001 conference on human factors in computing systems* (pp. 245–252). New York: ACM Press.

Corno, L., & Snow, R. E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 605–629). New York: Macmillan.

Dihoff, Roberta E., Brosvic, Gary M., Epstein, Michael L. (2003). The Role of Feedback During Academic Testing: The Delay Retention Effect Revisited. *The Psychological Record*, 53, 533-548.

Elliot, A.B., Murray, S.D., Ward, R.A. (2003). Spreadsheet-Based Exercises for Introductory Laboratories. *The Physics Teacher, 41*, 34-37.

Epstein, Michael L. , Lazarus, Amber D., Calvano, Tammy B., Matthews, Kelly A., Hendel, Rachel A., Epstein, Beth B., Brosvic, Gary M. (2002). Immediate Feedback Assessment Technique Promotes Learning and Corrects Inaccurate First Responses. *The Psychological Record*, 52, 187-201.

Hake, R.R. (1998). Interactive-engagement versus traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*, 64-74.

Hanna, G. S. (1976). Effects of total and partial feedback in multiple-choice testing upon learning. *Journal of Educational Research, 69*(5), 202–205.

Hennessy, S., Wishart, J., Whitelock, D., Deaney, R., Brawn, R., la Velle, L., McFarlane, A., Ruthven, K., & Winterbottom, M. (2007). Pedagogical approaches for technology-integrated science teaching. *Computers & Education, 48*, 137-152.

Gilman, D. A. (1969). Comparison of several feedback methods for correcting errors by computer-assisted instruction. *Journal of Educational Psychology, 60*(6), 503–508.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Kulhavy, R. W., & Anderson, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology, 63*(5), 505–512.

Kulhavy, R. W., White, M. T., Topp, B. W., Chan, A. L., & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology, 10*(3), 285–291.

Lai, Chiu-Lin & Hwang, Gwo-Jen (2015). A spreadsheet-based visualized Mindtool for improving students'learning performance in identifying relationships between numerical variables, *Interactive Learning Environments*, 23:2, 230-249.

Mory, E. H. (2004). Feedback research review. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). Mahwah, NJ: Erlbaum Associates.

Murthy, S. (2007). Peer-assessment of Homework Using Rubrics. (Unpublished professional paper). Massachusetts Institute of Technology, Boston, Massachusetts.

Peeck, J., Tillema, H.H. (1979). Learning from Feedback Comparison of Two Feedback Procedures in a Classroom Setting. *Perceptual and Motor Skills, 48*, 351-354.

Roper, W.J. (1977). Feedback in Computer Assisted Instruction, *Innovations in Education & Training International*, *14*(1), 43-49

Sassenrath, J.M., Garverick, C.M. (1965). Effects of Differential Feedback from Examinations on Retention and Transfer. *Journal of Educational Psychology, 56*(5), 259-263.

Schroth, M. L. (1992). The effects of delay of feedback on a delayed concept formation transfer task *Contemporary Educational Psychology 17*(1), 78–82.

Shute, V.J. (2007, March). Focus on Formative Feedback. Princeton, NJ: Educational Testing Service.

Sturges, P.T. (1978). Delay of Informative Feedback in Computer-Assisted Testing. *Journal of Educational Psychology, 70*(3), 378-387.

Sweller, J., Van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296.

Wagner, J.E. (2007). Using Spreadsheets to Assess Learning. *The Physics Teacher, 45*, 34-37.

Wang, T., Su, X., Ma, P., Wang, Y., Wang, K. (2011). Ability-training-oriented automated assessment in introductory programming course. *Computers & Education, 56*, 220-226.

APPENDICES

APPENDIX A

INSTITUTIONAL REVIEW BOARD EXEMPTION

INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 00000165

MONTANA STATE UNIVERSITY

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: John Wilkie and John Graves

FROM: Mark Quinn, Chair *Mark Quinn CH*

DATE: December 9, 2015

RE: *"Effects of Instant Feedback via Excel Template on Student Learning"* [JW120915-EX]

The above research, described in your submission of December 9, 2015, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

_X_ (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.

_X_ (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.

APPENDIX B

PRE-STUDY AND POST-STUDY ATTITUDE SURVEY

NOTICE: "Participation in this research is voluntary and participation or non-participation will not affect a student's academic standing in any way. You can choose to not answer any questions you do not want to answer and/or you can stop at any time."

1) I currently like physics.

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

2) Participating in labs helps me to better understand the topics we are learning in physics

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

3) Participating in labs helps prepare me for tests and quizzes

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

4) The feedback I receive on labs helps me to better understand the topics we are learning in physics.

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

5) The feedback I receive on labs helps prepare me for tests and quizzes.

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

6) How often to you review the comments and feedback on labs.

Almost Always   Sometimes   Once and a While   Rarely   Never

7) How do you use comments and feedback given on labs?

   a) I never look at the comments and feedback

   b) I look at them, but don't really try to see what I did that was incorrect.

   c) I look at them, and use them to correct my misunderstanding.

8) (Post Survey Only) I found the automated spreadsheet feedback helpful with learning physics concepts.

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

APPENDIX C

SAMPLE OF UNIT CONCEPT INVENTORY TEST

Name:_____

"Sound and Waves" Pre-assessment (10 pts):

1) What affects the speed of sound?

    a) wavelength

    b) frequency

    c) period

    d) medium

2) Paul Po from the police department measures the noise from a party at one part of a property line to be at a volume (power level) of 20 dB (decibels) and at another part of the property line to be 40 dB. How does the second volume compare to the first?

    a) 20x greater

    b) 10x greater

    c) 100x greater

    d) 2x greater

3) A spectrum analyzer shows a sound #1 to be composed of one tone at 20 Hz and 30 W and sound #2 to be composed of a tone at 40 Hz and 5 W. Which of the component tones will sound "higher pitched" and which will sound louder?

    a) Sound #1 will be higher pitched, Sound #1 will be louder

    b) Sound #2 will be higher pitched, Sound #2 will be louder

    c) Sound #2 will be higher pitched, Sound #1 will be louder

    d) Sound #1 will be higher pitched, Sound #2 will be louder

4) Tyrone Wills L. finds measures the loudest note on a train horn while it is standing still and finds it to be 800 Hz. Later he measures the loudest note on the same horn to be 850 Hz while the train is moving. Is the train going toward or away from Tyrone while he takes the second measurement?

    a) Toward him

    b) Away from him

    c) Neither

    d) He is on the train.

5) Simon Tulov is plays a note on a tuning fork and the corresponding note on his guitar and notices that the note seems to get loud and soft in "beats". Is his guitar exactly in tune? How do you know?

    a) Yes, beats indicate resonance, which is what you want in a guitar.

    b) No, beats indicate a different frequency than the tuning fork, which means the guitar is out of tune.

    c) Yes, the beats are a kind of instrumental "vibrato" which is appealing and don't indicate out of tune.

    d) No, the loud and soft beats indicate a loose piece in the guitar, and that it need repair to be in tune.

6) What is the wavelength of the A below middle C (440 Hz) if the speed of sound in air is about 343 m/s?

    a) 150,920 m

    b) 0.780 m

    c) 1.28 m

    d) 97 m

7) Alison Chang measures a tone's wavelength to be 3 m and its frequency to be 500Hz in water. What is the speed of sound for the water being tested?

    a) 1500 m/s

    b) 166.67 m/s

    c) 0.006 m/s

    d) 497 m/s

8) If a fundamental standing wave forms in a string of length 0.60 m, what is the frequency of the wave, assuming the speed of sound in the string is 400 m/s?

    a) 333.33 Hz

    b) 666.67 Hz

    c) 240 Hz

    d) 480 Hz

9) If a the third harmonic of an 0.414 m open tube oscillator (like a flute) is 1233 Hz, what is the frequency of the fundamental frequency, assuming the speed of sound in the air is 340 m/s?

   a) 411 Hz

   b) 0.28 Hz

   c) 510.5 Hz

   d) 960 Hz

10) If you are drilling holes in a home made flute, what should the spacing be between the different holes, i.e. what is the difference in open tube lengths to cause resonance to occur? Assume the frequency is 3430 Hz speed of sound is 343 m/s

   a) 0.025 m

   b) 0.10 m

   c) 0.05 m

   d) 10 m

APPENDIX D

POST-STUDY INTERVIEW

NOTICE: "Participation in this research is voluntary and participation or non-participation will not affect a student's academic standing in any way. You can choose to not answer any questions you do not want to answer and/or you can stop at any time."

1) What do you find to be the biggest benefits of doing labs, if any?

2) How do you normally study for quizzes and tests?

3) Do you think feedback on labs helps you learn physics better?

4) If you don't use feedback, why don't you?

5) Is there a type of feedback that you think is the most helpful?

6) How helpful was the automated spreadsheet feedback?