IMPROVING MEASUREMENT ACCURACY DURING SCIENCE EXPERIMENTS

WITH FOURTH GRADE STUDENTS

by

Samantha Eichner

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 2018

TABLE OF CONTENTS

LIST OF TABLES

iv

LIST OF FIGURES

v

ABSTRACT

This action research study measured fourth grade students' length measurement accuracy on performance assessments and science experiments before and after treatment. The purpose of the treatment was to determine whether meaningful practice would improve students' accuracy. Students' measurement accuracy on science experiments and performance tasks was recorded along with observations, interviews, and surveys. To determine which system they were most accurate in, students measured in both the English and metric systems. The results suggest that with practice students can make notable gains and record fewer unreasonable answers.

INTRODUCTION AND BACKGROUND

The purpose of this study was to improve length measurement accuracy during science experiments. During science experiments students are preoccupied with the procedure, and I have observed their length measurement accuracy being negatively impacted. Since measurement accuracy is not their central focus during science experiments it must be natural and practiced to transfer. Directly teaching how to measure and get an accurate measurement to the half and fourth of an inch is in the Common Core Standards for third grade. Measurement is required to complete many of the standards in fourth grade, but there are no standards for teaching it because the standards presume students already know how to measure from third grade. Unfortunately, many students still struggle to measure accurately, especially during experiments.

The sub-questions came from my own curiosity on this topic. After observing students struggle with their measurement accuracy during science experiments, I wondered if it was purely the science experiment that was causing them problems or if they had trouble with accuracy in any situation. I also wondered if students made more or fewer mistakes based on which system they were asked to measure in. In elementary school the standards are to recognize relative size and measurement accuracy in both systems. However, if they have significantly fewer errors in one system, teachers can use that information to decide which measurement data to use for experiments. Finally, I was curious about the impact this study would have on me as a teacher.

Measurement is a skill that has many real-life implications. Improved measurement skills help students in math and science, as well as in numerous measuring situations throughout their lives. For teachers, especially in elementary schools, inaccurate data collected during science experiments is a deterrent to collecting data or taking the time to analyze it. Offering ideas of how to improve this accuracy will benefit both math and science instruction. Teachers may be more motivated to make time in the schedule to practice if the benefit can be felt in so many areas.

<div align="center">Research Questions</div>

How does explicitly teaching and practicing length measurement affect students' measurement accuracy during science experiments?

    a. What impact do science experiments have on student length measurement accuracy?

    b. What is the accuracy level when students are measuring in the metric system as compared to using the English system?

    c. What is the impact on the teacher?

<div align="center">Support Members</div>

When I need expertise on anything I always go to my fourth grade team, specifically MaryLou Geraud and Sarah Cuevas. MaryLou has been mentoring me since I started teaching with her four years ago and always has time to discuss ideas with me. Sarah Cuevas and I share most of our teaching ideas and during this project it was no different. She helped me to develop many of the treatment lessons, discussed data with me, and shared results from similar activities with her class.

My friend Andres Valdepeña is an electrical engineer and helped me with data and graph creation. He assisted me in analyzing pre- and post-treatment data to find trends. He also helped me find the best way to display those findings. My sister was my editing team. My sister helps high school students edit their college essays.

CONCEPTUAL FRAMEWORK

Measurement is a skill that is often taught in isolation from its practical and scientific applications in school. Students are often shown how to measure in elementary school using worksheets. When asked to apply their measurement knowledge to other disciplines, such as measuring length during science experiments, I have observed students' accuracy being compromised. This begs the question, is there a way to teach measurement that will allow students to more easily transfer the skill to different measuring situations?

A review of the literature reveals a benefit to allowing students to use many types of rulers on a regular basis to practice measuring (Drake, 2014). This may help students apply their measurement skills to other situations, including science experiments. Piaget's theories of age and development of students are also important to consider when discussing acquisition of these measurement skills. The practical measurement tasks that were used in many of the studies discussed below were integrated into this research to assess students' measurement abilities. The amount of error students make in their measurement can be represented many ways including percentage error and standard deviation.

<u>Articles Providing Framework</u>

In his paper, Clements states, "Teachers should emphasize experiences and ideas of motion and distance" (Clements, 1999, p. 10). Clements describes his experience overseeing student teachers that were told by their cooperating teachers that students should measure at their seats using worksheets. Some cooperating teachers did not even allow rulers in their classroom citing that they were a distraction (1999). His research about using standard or non-standard units found that even young children preferred using rulers and were able to "use them meaningfully to develop an understanding of length measurement" (Clements, 1999, p. 8). This suggests that teachers should not shy away from the use of standard units of measurement and rulers. Teachers should instead promote measuring in as many subject areas as possible to create hands-on experiences for students.

Helping Children Learn Mathematics describes four main reasons to include measurement in a math curriculum (Reys, Lindquist, Lambdin, & Smith, 2012). Two of those reasons relate to this project: measurement is useful in everyday life, and measurement is essential for other areas of the curriculum, like science. One way to connect measurement to everyday life and other subjects is by integrating measurement into science lessons and activities. "Students who learn measurement skills in the context of scientific investigations experience greater success on high-stakes assessment test items than students who did not learn measurement skills in context" (Adamson, 2008, p. 40). Adamson conducted a large-scale study that trained teachers to integrate measurement into their science curriculum. The primary focus was the impact on English

Language Learners, but they also looked at the impact for all students compared to the schools that did not have the integration training. Since students measured more effectively when they learned measuring skills in context, this research suggests a positive impact of integrating measuring into science.

One potential problem students can come across unexpectedly during science experiments is having to iterate, or copy a unit repeatedly. Iteration refers to copying a unit repeatedly. In measuring longer distances, students who understand the concept of iteration, will mark their spot and move their ruler forward after they reach the maximum length of the ruler. Students should have this skill by fourth grade, as it is critical to their success in measurement. Kamii and Clark assessed and interviewed 383 students from first to fifth grades and found that 76% of students understood iteration by fourth grade (1997). Since students in this study were developing iteration it is important to include discussions about this topic when explicitly teaching measurement.

Students can be challenged to apply skills they learned in a math lesson to a science experiment. One study compared two groups of first through fifth graders. One group was taught using a traditional math program without an emphasis on manipulatives and the other used a "tactile, manipulative, and individual activity methodology" (Selman, Sell, Michael, & Martois, 1976, p. 1003). Through many different measures, the findings indicated that the test group receiving the instruction with an emphasis on manipulatives might be "associated with a greater degree of transfer and generalization of skills and concepts from one task to another" (Selman, Sell, Michael, & Martois, 1976, p. 1012). This study had a small sample size but the strength of this claim made it reliable.

The finding that using manipulatives increases a students' ability to transfer skills

supports the idea that practicing measuring makes it more likely that skill will help

students transfer this skill to other situations, such as a science experiments.

<center>Theoretical Articles</center>

Piaget's development and learning theory, "to know an object is to act on it"

(1964, p. 176), agrees with this project's course of action. Within the stages of

development Piaget argues that students need the physical experience of measurement

before they can progress to the logical math section. Teachers tend to feel rushed to get to

the math portion of measurement because the Common Core Standards require students

to know the relative size of standard units, convert from larger to smaller units, and solve

problems involving measurement (2010). In third grade, the Common Core Standards

require students to "generate measurement data by measuring lengths using rulers marked

with halves and fourths of an inch" (National Governors Association Center for Best

Practices, 2010, p. 25). No other grade requires students to practice or refine this skill.

Piaget's theory argues that the physical experience is essential to learning a new skill,

like measuring.

<center>Research Articles on Methodology</center>

Assessing students' measurement skills takes a variety of techniques. In a 2012

study, Sisman and Aksu's used four different tests to determine the length measurement

abilities of seventh graders in Turkey. To assess students' ruler skills the researchers

asked students to measure a thirteen-centimeter line segment. Only 66.4% of the seventh

graders were able to come up with the correct answer (Sisman & Aksu, 2012). Based on

Sisman and Asku's simple assessment, for the performance task in the action research project, students were asked to measure the length of a physical object slightly larger than thirteen-centimeters to determine their competency when only focusing on measuring.

Students' measurement accuracy for this action research project was assessed using an activity based on *The Basics of Data Literacy* (Bowen & Bartley, 2014). The original activity was designed to help students make bar graphs with authentic data and involved cups and marbles. It was modified for use in this study and the data was analyzed for error in measurement. This data analysis technique was conducted in a similar manner to a study by Schiffman (1966). The errors that two groups made when estimating length in metric and English systems were analyzed using the mean deviation (without regard for sign), was taken and calculated as a percentage error (1966). Schiffman's study found that participants were not significantly more accurate at estimating in one system than the other using this analysis. Since this study was looking at amount of error students make in both a science experiment and a classroom environment, this was an important analysis tool for this action research project.

<u>Conclusion</u>

Measurement is a complicated and difficult skill for students. There are many aspects that can be challenging, but providing time to practice with a meaningful context may help students become more accurate. The Common Core math standards only require some practice in third grade when even seventh grade students still find it difficult to measure accurately. With multiple practices it allows for "skills that are crucial for

students to understand linear measurement meaningfully to be explicitly taught" (Sisman & Aksu, 2012, p. 1906).

METHODOLOGY

The purpose of this project was to improve student length measurement accuracy in long and short distances during science experiments. I have observed students struggle and make errors during science experiments with many classes; therefore, I wanted to see if teaching measurement more explicitly and giving students more time to practice would improve their accuracy. Similarly, I wanted to see if the science experiment caused the errors, or if students just had trouble with any measurement accuracy. Another variable explored was the impact of having to measure in both metric and English units had on students' accuracy. Since fourth grade standards require students measuring in both unit systems, I wanted to see if they were more successful in one system over the other. Finally, I wanted to see what impact conducting an action research project would have on me as a teacher.

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. The exemption is included in Appendix A.

Sample

For the last four years, I have been teaching fourth grade at Willow Creek Elementary School in Nampa, Idaho. Willow Creek is a public elementary school that serves students from kindergarten to fifth grade with four teachers per grade level and about 650 students. School-wide about 60% of the students are Caucasian and 40% are

Hispanic. More than two-thirds (68.5%) of students qualify for free or reduced lunch at Willow Creek, which is higher than the district and state averages.

I am responsible for teaching my homeroom students all subjects. The sample for this research was my homeroom class because they are the only students I had long enough to implement the treatment. This made the sample twenty-seven students without the extended resource student with severe autism. Although it would have been great to include this student, he was not in my classroom for enough instruction for it to be accurate information. There were twelve girls and fifteen boys in my class. One student was on an Individual Education Plans (IEP) and six students were identified as English Language Learners (ELL). None of the ELL students were new to the country, but they received support from a Spanish-speaking co-teacher for forty-five minutes per day.

Overall, this group of students was average in math and below average in reading and writing compared to other classes I have taught. The high percentage of ELL students contributed to many of them struggling in reading and writing, however, they mostly higher achieving in math. They were also a great group of kids and had above average collaborative skills for working in groups on projects and challenges. One of the other fourth grade teachers commented that my class was, "the best class for S.T.E.M." based on a challenge she had them complete. There were a few students that struggled when working in groups, a typical challenge in any of my former classes.

<div align="center">Treatment</div>

The treatment of explicitly teaching and practicing measurement was completed with approximately fifteen thirty-minute lessons over the course of the eight-week

treatment period, starting January 9 and ending February 22, 2018. Students participated in seven hours of measurement lessons during that period, three of which were purely practice with either a ruler or measuring tape. The aim of all the lessons was to give students engaging practice in measurement. All lessons had an equal amount of practice with the metric and English system. The first four lessons concentrated on estimating with the objective that students could identify when their answers were incorrect. For example, students played a measurement version of The Price is Right where they worked with a team to come up with the closest estimate. Then there were seven lessons where students were able to practice in a controlled environment. These lessons included things like playing measurement board games or using the measuring tape with a group to measure things around the room. The remaining lessons students participated in were science experiments, one using short distances and one using long distances. Students were given clear directions that emphasized measurement accuracy for these science experiments.

For example, to practice measuring short distances during science experiments we did an activity called "The Great Gummy Worm Stretch". This experiment required students to measure a sour and regular gummy worm in both systems before and after it had been stretched to see which could stretch the furthest. After the students submitted their initial results I reviewed them for accuracy and presented the range of answers. We had a group discussion about how many of the answers were unreasonable, discussed important things to consider when taking measurements, and then completed the same activity again with a partner to check their work for errors.

During treatment students were also taught their regular math curriculum. The district purchased Pearson's Investigation 3 Curriculum for this school year. Before the treatment began, students had completed two weeks of measurement-centered math lessons. During treatment, math lessons focused on other topics, so no lessons were about measurement.

<div align="center">Instruments</div>

Six different instruments were used to answer the defined research questions. Several of the instruments answer more than one question and were combined for ease of analysis. Which question each assessment helps answer is shown in the research matrix, Table 1.

Table 1
Research Matrix

Research Questions	Data Collection Methodologies					
	Pre and post science experiment assessment	Pre and post performance assessment	Interviews	Classroom Observations	Surveys	Journals
How does explicitly teaching and practicing length measurement affect students' measurement accuracy during science experiments?	✔			✔	✔	
a. What impact do science experiments have on student length measurement accuracy?	✔	✔		✔	✔	
b. What is the accuracy level when students are measuring in the metric system as compared to using the English system?	✔	✔	✔	✔	✔	
c. What is the impact on the teacher?			✔			✔

The first student assessment was the measurement survey (Appendix B) to record students' initial thoughts on measurement. The survey consisted of ten questions with yes/no items, Likert questions, and several free response follow-ups. Many of the questions were about students' confidence in their measurement skills, but there were a few general measurement accuracy and exposure questions that helped answer the main question. This instrument was piloted with another fourth grade class in my school, and was modified to more adequately address the research questions and group like items together so that one question leads into the next.

Interviews were conducted with six students. Three boys and three girls were picked at random to participate in these questions. The interview questions are included in Appendix C and were increasingly more specific as the interview was conducted. This was the other piloted instrument, which helped to improve validity and reliability. At the end of the interview there was a performance assessment where students were asked to measure the length and width of a popsicle stick in centimeters and inches. Although only six students were interviewed, all students were given an initial performance task to compare their accuracy in a non-science experiment setting. The audio was recorded during the interview and students were filmed during the performance assessment portion so that their answers and methods could be reviewed.

The science experiment measurement assessment activity came from *The Basics of Data Literacy* (Bowen & Bartley, 2014). During the experiment, each student was given the setup shown in Figure 1 and asked to measure three trials with three different cups. The focus of the experiment was testing different variables so it could be repeated

at the beginning and the end of the study without students noticing they were being

assessed. Students completed the experiment in December and March. After each

experiment, I measured each of their lines and recorded both their measurement and the

actual measurement to the nearest fourth of an inch. This activity was piloted last year

with a different group of students, however each measurement was not checked and
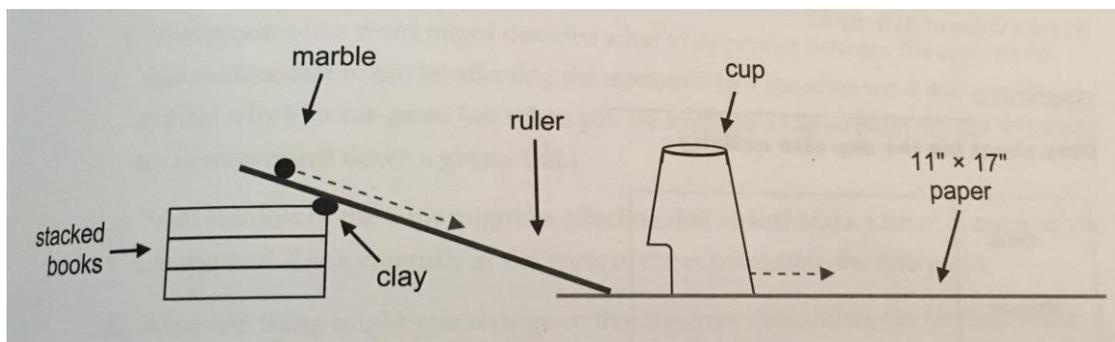
calculated like it was for this study.

Figure 1. Experiment set up.

To collect data on measuring long distances students also completed a

performance task and a science experiment. Students measured a designated section of

our hallway with a long measuring tape for the performance assessment. The distance

science experiment was a paper airplane trial with two types of planes, one plane type

measured in metric, and the other in English. During the experiment students threw their

paper airplanes, marked the spot it landed, and labeled it with their name and

measurement. Similar to the cup experiment these dots were checked later for accuracy

and repeated after treatment.

The classroom observation and teacher journal were combined into one document

and can be found in Appendix D. I chose to combine these instruments because I wanted

all areas that I need to record during an assessment or treatment on one document to

quickly record information. I collected data on my energy level, logged time during the

lesson, observed student comments and actions, and reflected on areas of strength and

weakness. Part of the formatting for the journal and observation sheet came from a

colleague, which help to make it more valid and reliable. These observations and

reflections helped answer the main question because I recorded what kind of errors

students were making. It will also answer sub-question C about the impact on the teacher.

The treatment period and the timeline of data collected are described in Table 2.

Table 2
Timetable

December	• Pre-surveys • Pre-interviews and performance tasks • Pre-assessment science experiment
January	• Treatment lessons • During treatment observations and teacher reflection
February	• Treatment lessons • During treatment observations and teacher reflection
March	• Post-assessment science experiment • Post-treatment interviews • Post-Survey

DATA AND ANALYSIS

To answer the main research question about improving students' measurement

accuracy during science experiments, I compared accuracy during four science

experiments before and after the treatment. Students made fewer errors after treatment in

all four experiments: short English, short metric, distance English, and distance metric.

The average error for all four experiments decreased after treatment. Since the

performance task was a set amount and the distances for the science experiments

changed, I did not use pure error numbers in the analysis. Instead I used the accuracy-

ranking rubric, shown in Table 3, to rank their error into advanced, proficient, marginal,

and unreasonable. The distances for each ranking were determined based on the

percentage error, trying to be as consistent as possible. For the science experiment section

the average distance for each experiment was used to determine the percentage error.

Table 3
Accuracy Ranking Rubric

Science Experiment	Advanced	Proficient	Marginal	Unreasonable
Short English	0 - .5 inch (5%)	.75 – 1 inch (10%)	1.25 – 2 inches (20%)	2 + inches
Short Metric	0 - 1 cm (4%)	2 cm (8%)	3 – 5 cm (20%)	6 + cm
Long English	0 – 8 inches (5%)	9 – 15 inches (10%)	16 – 30 inches (20%)	31 + inches
Long Metric	0 – 20 cm (5%)	21 – 38 cm (10%)	39 – 75 cm (20%)	76 + cm
Performance Task	Advanced	Proficient	Marginal	Unreasonable
Short English	0 - .25 inch (4%)	.5 – .75 inch (13%)	1 – 1.25 inches (21%)	1.5 + inches
Short Metric	0 - 1 cm (7%)	2 cm (13%)	3 cm (20%)	4 + cm
Long English	0 – 30 inches (5%)	31 – 60 inches (10%)	61 – 125 inches (20%)	126 + inches
Long Metric	0 – .8 m (5%)	.81 – 1.60 m (10%)	1.61 – 3.2 m (20%)	3.21 + m

The largest improvement of average error was in the distance metric experiment.

However, in Figure 2 it shows that this experiment also maintained the highest average

error after treatment of 293%. Students had the most error during the long distance

experiments to begin with and had not used a measuring tape to conduct experiments

before this study. With practice and explicit instructions on how to use the measuring

tape, they were able to bring their averages down by more than 20%. Many of the errors,

like measuring with the zero at the wrong end or looking at the English side when the

directions were to report in metric, were not as prevalent after treatment.

In the distance experiment using the English system, students started out with

substantially less error than the distance metric experiment, only 33%. Their average

error improved to 4% after treatment, which is translates to an average of 24 fewer inches

of error. This is also the only science experiment where no students were considered

marginal or unreasonable after treatment based on the accuracy-rating rubric, Table 3.

Their success in this area I would attribute to their familiarity with how small one foot is

compared to one meter. One student reported feet being easier than meters by saying, "It

is super easy because like 1.5 feet that's easy but big ones like 2.5 meters or 2.25 is not

easy because of the half." Both of these measurements have halves but I believe that this

student felt more comfortable with feet because they were not "big ones" like meters and

more manageable for them.

The short experiments were more accurate to start with and students only made

small improvements. The average for the English system error stayed nearly the same

(.77 inches to .71 inches) but was also the lowest to start with. The biggest barrier for

students to improving their short English accuracy seemed to be their understanding of

determining the fractional amount on a ruler. Students were taught to measure to the

nearest fourth of an inch, as per the Common Core Standards, but during treatment there

was still a lot of confusion. Students asked questions like, "Is that the half point?" or

"Can the nearest ¼ be a whole inch?" Even as we worked on fractions in math during treatment this still caused confusion for many students.

For the short metric experiment, fractions were not an issue since I only measured to the nearest centimeter. In Figure 2 it shows that there was only an average error reduction from 2.1 cm before treatment to 1.5 cm after treatment. After treatment only one student gave an unreasonable answer, compared to five before treatment. This was determined using the accuracy rubric in Table 3. The one student who remained unreasonable was less accurate after treatment but after reviewing her experiment paper, I noted that she made the lines after she recorded the measurement instead of prior to measuring as instructed. After treatment that same student reported, "I do not think that I measure accurately" but did not give specifics of what was difficult. Accuracy was clearly not a priority for her and treatment did not impress upon her the importance of it.

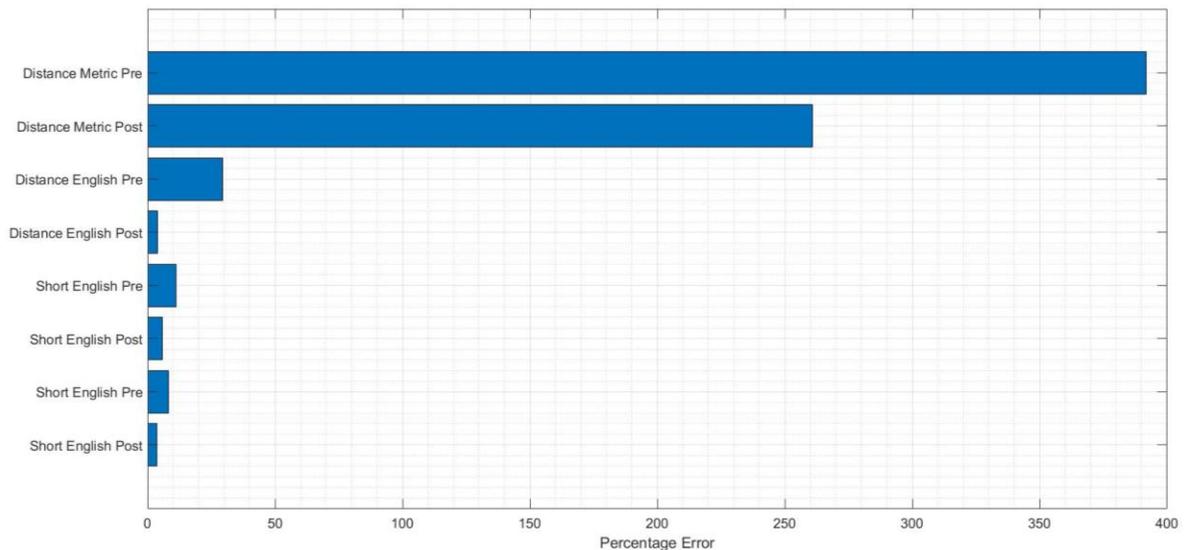

Figure 2. Average error during science experiments pre and post, (*N*=27).

The standard deviation similarly shows a tightening of scores only for the long distance science experiments. The short distance science experiments had essentially no

difference in their standard deviation before and after treatment. The experiment that had

the largest standard deviation improvement was the distance English science experiment

going from 34.7 to 3.8. Many students had the measuring tape backwards on the initial

experiment, causing the high standard deviation. This was the students' first time using

the measuring tape and many of them did not realize the importance of starting at zero.

One student did catch this error but most students just recorded what the tape said

without thinking about whether their measurement was reasonable. The post treatment

experiment did not have this large error because the same student who noticed the tape

was backwards in the initial experiment spoke up immediately when students started

putting the zero at the wrong end again. This time enough students knew that he was right

that they corrected it.

Survey data is shown in Figure 3 and confirms that after treatment students were

able to better articulate what they needed to do to measure accurately. Before treatment

when asked the question, "Explain what you do to measure accurately," the most

common response was that they already knew how to measure with no details about the

process. After treatment twelve more students mentioned either lining up the zero or

keeping the ruler straight in their answer. For example, before treatment one student

responded to that question, "Just measure the best you can and hope you got it correct."

After treatment they said, "Start at the zero and keep it straight." This was an important

finding because during the pre-assessment many of the students' errors were due to not

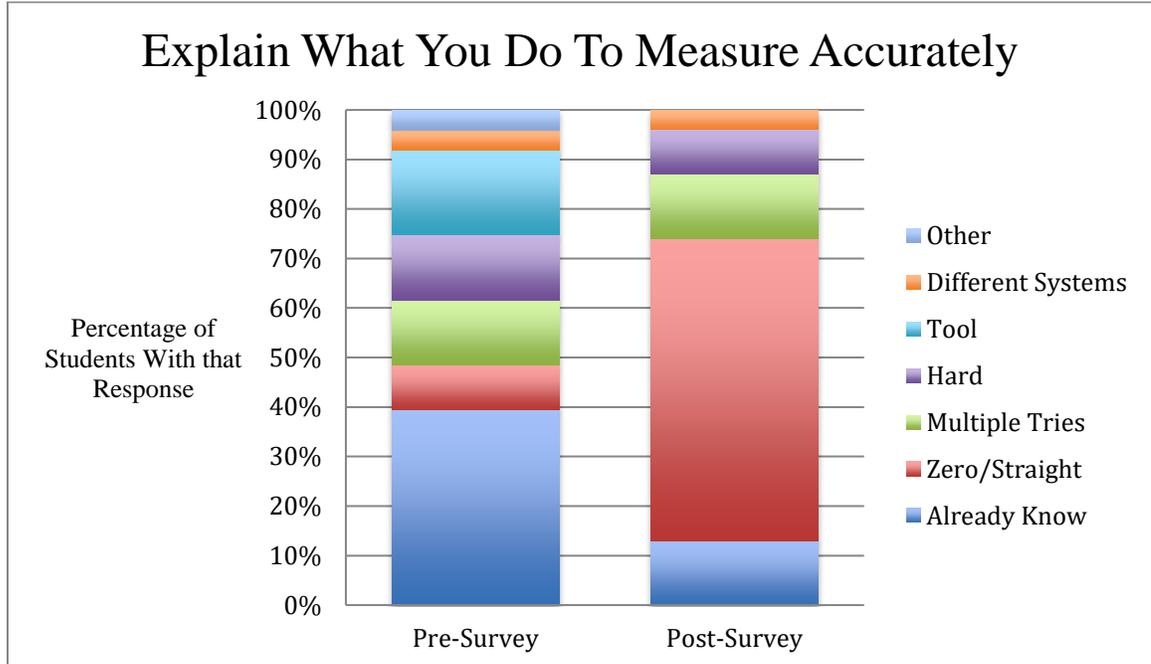lining up the zeros with either tool.

Figure 3. Categorized survey results for question: explain what you do to measure accurately, (*N*=23).

The improvement in student accuracy after treatment suggests that practice is beneficial. During the pre-interviews one student recommended practice as something that the teacher could do to help students measurement more accurately. She said, "We could measure more often. It would help me know how to measure better." Many math curriculums leave measurement until the last unit of the year and have minimal practice but focus more on conversions. Science experiments are one way to provide students with more time for practice throughout the year. With this study intentionally putting length measurement into lessons and science experiments, my students received an additional seven hours of practice that they would not have had with traditional measurement instruction. By continually integrating measurement into projects and experiments, it may improve students' length measurement.

Science Experiment Impact Analysis

Comparing science experiment accuracy to performance assessment accuracy after treatment showed that students were only slightly more accurate during the performance tasks. For the short performance tasks in both systems, only one student did not achieve the advanced mark after the treatment. Table 4 shows that in the corresponding short science experiments six or seven students missed the mark of advanced. Since all but one student was able to measure in the advanced range during the performance task, it seemed that at least some students are being challenged by the experiment, even after treatment. Short distances provided unique accuracy challenges and even students who are above average in math were reporting difficulties like: "Do you need to use more than the wholes?" or "I think it isn't easy to measure accurately because rulers have extra plastic." Students realized that these little things mattered and got them correct on the performance task but many did not during the experiment.

Table 4
Short Distance Performance Task and Experiment Comparison, (N=27)

	Advanced Students	Proficient Students	Marginal Students	Unreasonable Students
Post-Short English Performance Task	26	1	0	0
Post-Short English Experiments	21	0	0	6
Post-Short Metric Performance Task	26	0	0	1
Post-Short Metric Experiments	20	3	3	1

The distance lengths had very different results depending on which system students were measuring in. The distance English science experiment was the only experiment that students performed better on than its similar performance task. For this

experiment no students were labeled marginal or unreasonable (Figure 4). This is an interesting finding given that many of the possible explanations should also apply to the metric equivalent, which had very different results. This could be because students are more comfortable with the English system and have more experience with it. As a result, during the experiment it was easier for them to use the English system. It could also be that the size of a foot was more manageable, while a meter was so long they did not know exactly where it ended.

In the two metric experiments, students made significantly more error, which can be seen with their lower rankings in Figure 4. The same number of students were unreasonable or proficient in the performance task and science experiment. However, the science experiment had five students who had a ranking of marginal and the performance task had none. This meant that the performance task had eight more students ranking as proficient compared to the science experiment.

In all but one comparison students performed better on the performance task, where their only assignment was measuring. This suggests that students generally perform better when they are only focusing on measuring. However, students do have the capacity to measure accurately in completing science experiments as well. Making these more complicated science experiment situations familiar by keeping the set up and expectations the same may help improve accuracy in the long term.
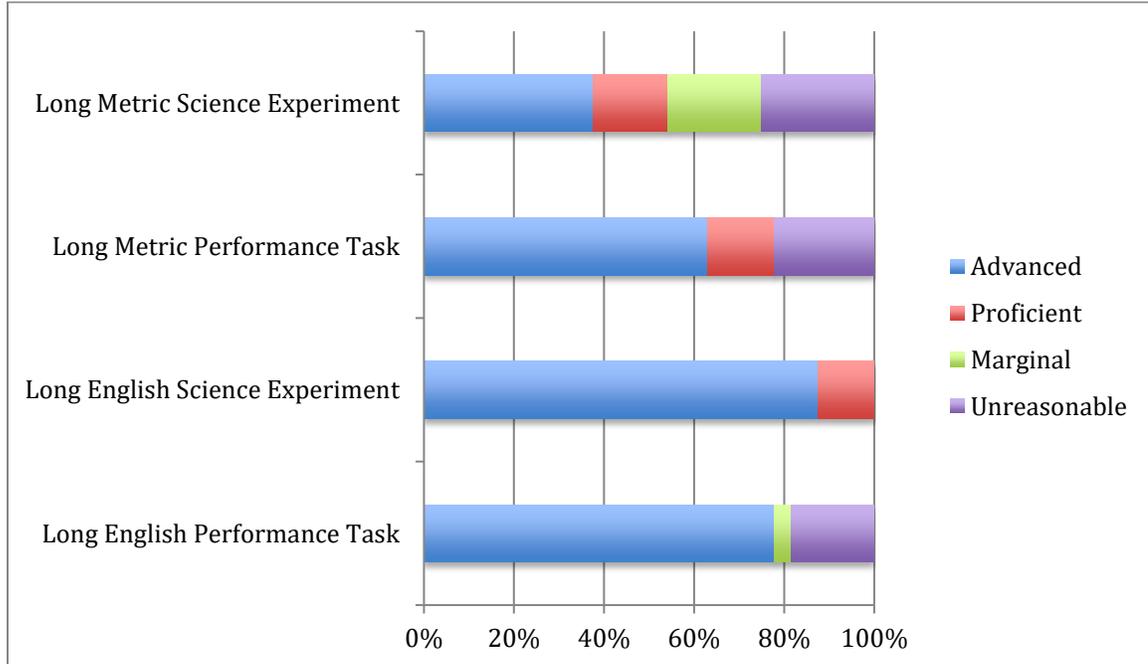
Figure 4. Long performance task and science experiment proficiency comparison, (*N*=27).

Metric and English Accuracy Analysis

Students completed science experiments and performance assessments in both the metric and English systems. More students had higher accuracy in the English system on six out of eight comparable situations, including both pre and post treatment. The only situations where students did better using the metric system were the short performance task and science experiment before treatment. Since students were struggling with measurements that were not whole numbers, especially before treatment, it makes sense that they would fare better when the centimeters were only being measured to the nearest whole. The lack of fractions and smaller unit size in the metric system may have contributed to higher accuracy before treatment. However, the treatment, along with familiarity with the units, caused students to be most accurate in the English system post treatment.

Observations, surveys, and interviews revealed that students were confused by the two systems. One of the two gifted and talented students said, "I think it is hard with the metric system and English system trying to not get them mixed up." He wasn't the only one; on the surveys seven students mentioned something about measuring being difficult because they did not know which system to use. This was also observed when students of all ability levels would report answers that combined both systems like "1 inch 1 cm". During treatment the students and I had several conversations about how those types of measurements do not make sense, but students continued to answer that way because they didn't understand the systems. After treatment only one of those seven students still answered that the different systems were confusing them. I attribute most of the improvement to repetitive practice and exposure. We may be doing students a disservice by asking them to juggle two systems with a task that is already new and difficult for them.

<center>Impact on the Teacher Analysis</center>

In his initial survey, one student perfectly summarized why measurement has been so difficult in the past: "I don't get it right at all! Help!" Before treatment began only two students answered "no" to the survey question, "I have had a lesson at school that taught me how to measure." However, most of the students' experience has not been extensive enough for the level of accuracy many teachers expect. Over the course of this study I learned how complicated a skill measuring is and because of the complexity it makes sense that students would not have it mastered by fourth grade. The difficulty does not mean however, that students cannot improve tremendously with some specific

instructions and practice. I learned not to demand perfection but to get as many students to make reasonable measurements as possible. In the past I was overly frustrated by all of the error in students' measurements, but throughout this process I have become more focused on growth and reasonable measurements.

The largest time commitments during this process was developing the lessons and checking each of the assessments for error. Each lesson took about an hour to find, adjust for the studies purpose, and prepare the materials, for a total of about six hours. Although these lessons took considerable time to develop all of them will be used in the future and were shared with other teammates to develop measurement skills for the whole fourth grade. The larger time commitment was checking each student's answers for error. Each assessment took more than an hour to re-measure and record. That was about six hours for before and after treatment. This is equivalent to the lesson planning time commitment and integral to this study but it is not something you would spend time doing on a regular basis. Instead, I might take time after the experiment to discuss and spot check errors as a class, pair students up to check each others answers and then have them discuss and correct themselves, or use a trustworthy student aid to recheck and record all the measurements.

INTERPRETATON AND CONCLUSION

Students can improve their accuracy, but it is more difficult for students to avoid error during science experiments. On average students maintained their original amount of error or improved after they had time to practice during treatment. Common mistakes that were addressed during treatment were less prevalent after treatment and helped to

lower their error, especially during the distance experiments. This is consistent with the findings of Drake, Sisman and Aksu that hands-on meaningful practice is helpful and a way to get a skill to transfer from one situation to another (2014 & 2012).

This means that any time a teacher can include measuring in any classroom activity it is beneficial to do so. Since practice does improve student accuracy and many math curriculums leave little time to practice, teachers need to get creative. Science is a great place for students to get practice measuring with a purpose but it can also be included in social studies projects or even art lessons. Many of these subject areas are being cut in elementary school in favor of reading or math but integrating them with measurement could help create time for both subjects. For example, I plan to add measurement elements to our Oregon Trail unit and seek out art projects that have measurement requirements.

This study revealed that science experiments have mixed impacts on students' length measurement accuracy. In three of the four comparisons students were more accurate during the performance tasks, but for the long English science experiment they did better than the performance task. This suggests that although students usually measured more accurately when they were focused on the sole task of measurement, they have the ability, especially in a system they are familiar with, to be just as accurate during a science experiment. Teachers might believe that science experiments negatively impact measurement accuracy, but I wanted to establish that through analysis before using it to explain student errors. This study confirms science experiments are obstacles for students in measuring accurately. The fourth comparison where students did better during the

science experiment suggests that there are lots of factors to be considered, not just the situation.

Another important finding of this study was that students were more accurate in the English system. They had less average error and lower rankings on the accuracy rubric in all but two experiments. After treatment students were more accurate on all the English measurements even though treatment had been equally split between use of the English and metric systems. Based on these findings I plan to spend more time instructing and practicing the metric system to hopefully make up for these deficits. The metric system should be used for all science related activities but should also be emphasized at school to counteract any practice students get in the English system at home.

Measuring long distances in the metric system was especially hard for students. The average error after treatment was about ten meters for both performance task and science experiment. The same tool was used for distance measurements in both systems so the familiarity with the tool should not have been an issue. All the measuring tapes were not the same though and some had the zero for the metric system in a different spot than they were for the English system. We discussed this issue during treatment, however, this could account for some of the differences between the two systems. Continued practice and setting up of the measuring tape could help students look for the zero and get used to using that tool. An example of repeated practice could come during the jumping events at a track meet.

Familiarity with the unit sizes for the English system may be causing the increased accuracy in that system. Students are likely to only use the metric system at school and may not have realized there are two systems until taught in a classroom. The treatment time spent measuring using the metric system did not seem to be enough to make up for their previous familiarity with the English system. If students are more familiar with the unit size they should be able to tell if their answer is reasonable or not. The purpose of the "Price is Right" estimating activity was to help students improve their estimating skills. Just like the rest of treatment it was equally split between the English and metric systems. However, based on these findings, it may be helpful for students to concentrate only on the metric system for these estimating activities. The metric system has simple estimates for students to remember for short distances like a centimeter is about the width of your pinky finger, but I have never found a simple one for a meter. That longer distance is harder for students to visualize compared to a foot that is about the length of their foot. Coming up with an easy estimate as a class may be beneficial.

Requiring students in the United States to learn and become proficient in both systems is doing them a disservice. Teaching the metric system should come first as failing to do so "may limit their participation in the global society," especially in areas like math and science (Monroe & Nelson, 2000, p. 23). However, needing to teach both systems, as mandated by the Common Core Standards, means that students get half the practice that they are allotted during school hours because they must practice both systems. There are some common mistakes that can be addressed in both systems, but for

this difficult skill we should not be asking students to keep two systems separate because adults are not willing to change (Monroe & Nelson, 2000).

The impact on me as an educator was limited to the area of measurement instruction. After spending so much time watching and rechecking student's measurement, I have a better idea of the many different errors that are common for fourth graders to make. The knowledge of these common errors and ways to combat them can be used in the future. Analyzing the before and after data and seeing improvement also reinforced to me the importance of practice in any and all forms. Since confronting error is something I have shied away from in the past, this is significant and will change my practice in both math and science.

VALUE

As an elementary school teacher, I used to avoid or heavily modify measurement during science experiments because students' answers tended to be so inaccurate. This was shown through the high errors students exhibited in the first round of science experiments in this study. In the past, I thought that is as far as I would get thinking there was nothing I could do to rectify the issues students were having. Over the course of this study, I was not given that option to quit and the persistence paid off. The increased accuracy students demonstrated after treatment show that students can improve with more instruction and practice measuring.

I also had to wrestle with the question, "How accurate is enough in fourth grade?" This came late in the action research process when I was trying to analyze the results. Initially, accuracy was getting the answer exactly with no error but looking at the data,

especially for the longer distances, that did not seem appropriate. Students were just learning and practicing measurement, so expecting perfect measurements was not reasonable. The initial frustration that led to this project had to do with students' unreasonable measurements, so focusing on reducing those types of errors was much more important than errors of one or two centimeters.

Finally, knowing where students' errors were coming from was vital. The use of observations and interviews showed where students were getting mixed up. This is a technique that I had never used, and it was efficient to get to the root of the problem. Previously, I had watched students while they measured but never formally recorded it to look at later. I will simplify the observation tool from this study so that I can record common mistakes in a streamline process while still getting this important information to drive instruction.

<div align="center">Next Steps</div>

Science experiments will return to being only measured in the metric system in my classroom. Science experiments were done in both systems during the action research process to compare student accuracy in the different systems. The results showed that students are less familiar and accurate in the metric system and therefore need as much practice as possible. Being consistent with this addresses another finding that students had a difficult time switching back and forth between systems. Maintaining consistency should be helpful in improving accuracy.

Some of the most valuable conversations I had with students during this study were around the idea of reasonableness. Many times after a treatment activity I showed

the range of answers and we discussed which we considered unreasonable. If time allowed I would have them measure again, sometimes with a partner, to hopefully come up with a reasonable answer. Based on the mostly positive findings I will continue to have these conversations with my students. Many of the treatment lessons I developed for this study will also be used in years to come and shared with teammates.

REFERENCES CITED

Adamson, K. (2008). The impact of an integrated math and science curriculum on third grade students' measurement achievement (Unpublished doctoral dissertation). University of Miami, Coral Gables, Florida.

Bowen, M. & Bartley, A., (2014). *The basics of data literacy: helping your students (and you!) make better sense of data.* Arlington, VA: NSTA Press.

Clements, D. (1999). Teaching length measurement: Research challenges. *School Science and Mathematics*, 99(1), 5-11.

Drake, M. (2014). Learning to measure length: The problem with the school ruler. *Australian Primary Mathematics Classroom*, 19(3), 27-32.

Investigations Curriculum Center. (2017). *Investigations 3 in number, data, and space: Fourth grade*. Glenview, IL: TERC.

Kamii, C. & Clark, F. (1997). Measurement of length: The need for a better approach to learning. *School Science Mathematics*, 97(3), 116-121.

Monroe, E. & Nelson, M. (2000). Say "Yes" to metric measure. *Science and Children*, 38(2), 20-23.

National Governors Association Center for Best Practices. (2010). Common Core State Standards (Mathematics). *National Governors Association Center for Best Practices*, Washington D.C.

Piaget, J. (1964). Part I: Cognitive development in children- Piaget development and learning. *Journal of Research in Science Teaching*. 40(1), 176-186

Reys, R., Lindquis, M., Lambdin, D., & Smith, N. (2012). *Helping children learn mathematics* (10th ed.). Hoboken, NJ: John Wiley and Sons, Inc.

Schiffman, H. (1966). A comparison of the English and metric system in a length estimation task. *Psychonomic Science., 4*(12), 399-400.

Selman, D., Sell, P., Michael, W. B., & Martois, J. (1976). Interrelationships among several variables reflecting quantitative thinking in elementary school children with particular emphasis upon those measures involving metric and decimal skills. *Educational and Psychological Measurement, 36*(4), 1003-1014.

Sisman, G. & Aksu, M. (2012). Seventh grade students' understanding of linear measurement: What has been learned about seventh grade measurement in seven years of schooling process? *Procedia Social and Behavioral Sciences*, 46. 1905-1909.

APPENDICES

APPENDIX A

IRB EXEMPTION

INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 00000165

MONTANA STATE UNIVERSITY

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

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

MEMORANDUM
....

TO: Samantha Eichner and Walter Woolbaugh

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

DATE: November 8, 2017

RE: *"Improving 4th Grade Measurement Accuracy During Science Experiments"* [SE110817-EX]

The above research, described in your submission of November 8, 2017, 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

MEASUREMENT SURVEY

Name:_____

Measurement Survey

1. I am a confident person.

   Strongly Agree   Agree   Disagree   Strongly Disagree

2. I am good at math.

   Strongly Agree   Agree   Disagree   Strongly Disagree

3. When measuring I am confident in my answers.

   Strongly Agree   Agree   Disagree   Strongly Disagree

4. When measuring I get an accurate measurement **every time**.

   Strongly Agree   Agree   Disagree   Strongly Disagree

5. It is easy to measure accurately.

   Strongly Agree   Agree   Disagree   Strongly Disagree

   Explain why you answered that way about measuring accurately:

   _____

   _____

6. I need to measure in science.   YES   NO

7. Measuring accurately is important in science experiments.

   Strongly Agree   Agree   Disagree   Strongly Disagree

8. It is easy to measure accurately during a science experiment.

   Strongly Agree   Agree   Disagree   Strongly Disagree

9. I have had lessons at school that taught me how to measure.   YES   NO

   If YES please explain one measurement lesson you were taught:

   _____

   _____

10. I know what to do to measure accurately?   YES   NO

   Explain why you answered that way:

   _____

   _____

APPENDIX C

INTERVIEW AND PERFORMANCE ASSESSMENT QUESTIONS

Interview and Performance Assessment Questions:

Warm-up Question- What do you like most about school? The least?

1. Do you like math?
   Probe- What about math do you like or dislike?
   Probe- What parts of math do you like?

2. Do you like measuring?
   Probe- What about measuring do you like or dislike?

3. When do you first remember a learning experience using measuring?
   Probe-What type of activity do you remember?
   Probe-What did you learn about measuring from that experience?
   Probe-Do you ever have to measure in subjects other than math?

4. On the survey you answered that you are (confident/not confident) about measuring accurately. Can you explain why you answered that way?
   Probe- How confident do you feel in your measurement skills/answers?

5. What could a teacher do to help you measure more accurately?
   Probe- How do you make sure that you measure accurately during science experiments?

6. What is the length, width, and height of this popsicle stick? Let me know when you have an answer for each so I can record it.

APPENDIX D

TEACHER JOURNAL AND CLASSROOM OBSERVATION TEMPLATE

Date _____   Circle One: Pre-assessment   Treatment

<u>My energy:</u>

Total Time	Instructional Time	Skill Taught	Practice Time

Student Comments	Observed Behavior

What Went Well?	What Didn't Go Well?