United States Middle School Students' Perspectives on Learning Statistics

Jerry Dwyer, Department of Mathematics and Statistics
(jerry.dwyer@ttu.edu)
Kim Moorhouse, College of Education
Malinda J. Colwell, Department of Human Development and Family Studies
Texas Tech University, MS 1042, Lubbock, TX 79409

Abstract

This paper describes an intervention at the 8th grade level where university mathematics researchers presented a series of lessons on introductory concepts in probability and statistics. Pre- and post-tests, and interviews were conducted to examine whether or not students at this grade level can understand these concepts. Students showed a significant gain in knowledge as evidenced by the improvement from pre to post-test. However, some question responses were limited and the interviews indicated some of the barriers to learning that students encountered. Acknowledging these barriers will enhance the learning experience that teachers provide their students and will promote an awareness of the diversity of learning styles and preferences present in any given group of students. In addition, appropriate background material should be provided and assimilated by the students in the early stages of an intervention.

Introduction

The understanding of statistics among non specialists is often limited and misguided. Newspaper articles commonly describe misinterpretations by the general public of data that are presented in either numerical or graphical format. These trends in public knowledge can often be traced to deficiencies in elementary and middle school mathematics and statistics learning. At the same time the study of probability and statistics at these levels has attracted critical attention and there is some debate regarding how much can be presented and at what grade level.

This paper describes an intervention at the 8th grade level where university mathematics researchers presented a series of lessons on introductory concepts in probability and statistics. The purpose is to examine whether or not students at this grade level can understand these concepts and whether or not they possess sufficient mathematical background to understand and implement routine statistical techniques. It was hoped that in-depth analysis of problem solutions together with student interviews would offer insights into the learning of statistics that would prove useful to other instructors.

The next section reviews some relevant literature. This is followed by a description of the participants and the lessons that they received. Students' solutions to assigned problems are analyzed in the next section. Student interviews are then described in detail and the impediments to their learning are discussed. Finally recommendations for future work are suggested.

Literature Review

The teaching of introductory statistics courses has not kept up with changes in society, technology and communications. Utts (2003) discusses changes, and the way the introductory syllabus should change to reflect them. In particular, several ideas are discussed that every student who takes elementary statistics should learn and understand in order to be an educated

citizen. These include cause and effect, statistical significance, probable coincidences, biases in surveys, and confusion between normal and average. Misunderstanding these topics leads to misuse of study results by policy-makers, physicians, and others. Most people, including reporters, are not able to critically read statistics that are found daily in newspapers and magazines. Given that people encounter statistics on a daily basis, knowledge of statistics is relevant in helping people to make a multitude of decisions. However, the area that often receives the least attention in present mathematics curriculum is statistics. According to Mitchell (1990), high school graduates seem to have no understanding of what constitutes experimental design and have difficulty grasping the concept of random sampling. A second area sorely neglected in the curriculum is probability and a related part of the curriculum not taught in a relevant manner is the concept of percentages. Failing to teach students the skills necessary to interpret and understand these concepts lowers their potential to be informed consumers of nutritional information on food items, news reports, and weather information.

Even with compelling evidence that students need a solid understanding of basic statistical concepts, the relevance of statistics in the K-12 curriculum is often questioned. Supporters will point out that probability can be used to illustrate the concepts of proportion and percentage. Detractors will suggest that more time should be spent on these concepts alone without spreading the curriculum too thin with too many topics. A recent article by Martinie (2006) discusses the importance of statistics at the middle school level. Probability and statistics are an important component of recommendations from the National Council of Teachers of Mathematics in the United States (NCTM, 2000). These recommendations suggest the need to find creative and effective ways to incorporate these concepts into the K – 12 curriculum. Innovative methods will help to capture student interest in statistics and allow students to appreciate its relevance in their everyday lives. Russo & Passannante (2001) describe a project that allowed students to generate their own survey instrument, collect data, analyze data, and present their findings. A mathematics teacher and statistician worked together to develop the project, which was part of an eight-week unit on statistics for seventh-grade students. Teachers acted as facilitators, guiding students through the learning process, while students worked toward finding answers to questions they had posed. The project provided a meaningful context for teaching valuable statistical concepts, made use of important resources, and gave students the opportunity to experience firsthand the value of mathematics in a real-world context and the connections among various areas of mathematics, science, and technology.

Another innovative approach was taken by Goldsby (2003) who believes that statistics is a vital part of the curriculum and it can be introduced in a captivating way that will stimulate students to want to learn. The 6^{th} grade students in their study worked with the concepts of counting, frequency tables, and graphing data. It was assumed that students understand concepts better and work better when material is presented that is meaningful, relevant, and that is an area of interest. To get the students interested, they utilized music in the lessons. The work of Russo and Passannante (2001) and Goldsby (2003) show that by using creative approaches, students can be encouraged to actively engage in the learning process and come to appreciate the practical applicability of statistics in many everyday contexts.

In addition to using hands-on projects to teach statistics, general methodologies used by teachers are also important to consider. Wulff & Wulff (2004) investigated one teacher's evolution from formal lecturing to interactive teaching and learning in a statistics course. Student perception data were used to demonstrate the instructor's use of communication to align the content, students, and instructor throughout the course. Results indicated that the students

learned, that communication in the alignment process played a role in the learning, and that the instructor used four key categories of communication: encouraging open communication, demonstrating examples interactively, structuring opportunities for application through problem-solving, and engaging students in reflection about their learning. They conclude with the instructor's reflections on how communication and the approach to teaching and learning changed as a result of conducting the scholarship of teaching and learning in the course. These findings suggest that more general approaches to teaching and the way in which a teacher interacts with his or her students is important for the students' experience of learning statistics. The timing of students' exposure to learning statistical information also is likely to be important for understanding what students know about statistics. Scepansky & Carkenord (2004) describe an empirical investigation of whether performance on a senior-level assessment of methodological and statistical knowledge was related to elapsed time since students took prerequisite methods and statistics courses. In a senior-capstone course, 50 students completed a 50-item multiple-choice assessment measuring their knowledge of methods and statistics. They also measured elapsed time (number of months) since students completed Research Methods and Quantitative Methods (statistics). Results revealed that performance on the assessment showed no significant correlation with elapsed time from the Research Methods course or with elapsed time from the Quantitative Methods course. On a self-report survey, however, students reported they believed the elapsed time negatively affected their performance. These findings suggest that not only do major concepts need to be introduced early in curriculum, but they need to be reinforced throughout related classes to help students maintain and build on what they have learned.

A related international study by Nooriafshar and Maraseni (2005) reports that geographical location and cultural background have no major effect on students' preferences in terms of learning method or style. However, university students prefer learning the concepts first and then find out about the application, in contrast to high school students. Furthermore, the percentage of students seeing the greater practical use of statistics increases according to their level of education.

Intervention

Seven lessons, of about 45 minutes each, on probability and statistics were presented to an 8^{th} grade class in a private middle school. Lessons occurred about once a week over a two month period towards the end of the spring semester. There were seven students (four female and three male) in the class and most were from a white middle class background. They had been exposed to a little Algebra I prior to this intervention. Students completed a short problem set before and after the intervention in order to determine if their understanding of the concepts had improved through the course of the intervention. Three students were interviewed at the conclusion of the study in order to provide further insights into their learning.

Lessons

The lessons are briefly described hereunder, together with some comments on student reactions. It was interesting to observe that several elementary arithmetic concepts had not been mastered prior to these lessons. Overall, students were most engaged when an activity (active learning) was taking place.

The first lesson was on elementary probability. This involved concepts of proportions and percentages. While introducing the idea of probability, something the students should already be familiar with, it was found that the students couldn't multiply two fractions. Before the lesson

continued, the students were given a review of fractions. It was interesting to note that the students did not know what the term "equivalent fraction" meant, but that they knew what it was when it was explained to them.

The second lesson extended to the concepts of conditional and multiple probability. This required some review of percentages. The use of pie graphs was useful in explaining proportions and relating them to percentages. A coin flipping activity generated more enthusiasm than the theoretical explanations. During this lesson, the students often gave answers without understanding the reasoning behind why they do certain steps; rather, they only went through the motions.

The third lesson introduced the ideas of Mean, Median, and Mode. Examples of temperature and real world applications, such as student height and weight, and exam scores, were used. Students worked in two groups to find the mean of a set of data. Some students failed to obtain the answer due to careless mistakes. Also, during this lesson, the girls did not seem very interested in the lesson. This may be due to their being bored with something that they already knew how to do.

The fourth lesson reviewed the concept of the Mean and students didn't seem to pay attention. Students were then asked to engage in an activity with a spinner and probability of hitting a certain number over several trial spins. They were awake and engaged during this activity. When the activity was over the students lost interest again and were not paying attention.

The fifth lesson was on Standard Deviation. This was introduced by noting that the Mean doesn't give all the information we need from a data set. The calculations here were long and tedious and students easily forgot some of the steps. A practice problem had been erased from the board, so perhaps they would have done better if it had been left on the board because none of them were taking notes. Two of the girls seemed totally disinterested during the lesson.

The sixth lesson was on the Normal Distribution. At the beginning students seemed to have forgotten what they did in the previous lesson. Students constructed a table of data and were told that when the sample increases, the data graph looks more like a normal curve. The concepts of proportion and percentage were reinforced here. Examples of percentages in everyday life, such as relative humidity and blood concentrations, were discussed. Students were given an average and SD and sample size, and asked to draw their own Normal Distribution curve. There wasn't enough time to complete their work and the solution was presented hastily. It was interesting to note that while the students were working on their own, one of the students mentioned to another student that sometimes when the teacher's lips are moving, she is disconnected and doesn't really hear what the teacher is saying. This same student asked another student how he got his answer and he did the work for her. This indicated that she may have not been engaged during this lesson. Also, during the lesson, two students passed notes the entire time, one student often laid his head on his desk, and another tried to read a book. There was clearly some disengagement from this lesson.

The seventh and final lesson was on Hypothesis Testing. It was recognized that this concept may be beyond the students' capacity and so only a general idea was be presented. Because the explanation was somewhat abstract, with no visuals, the students did not seem engaged.

Problem Set Results

In order to ascertain whether or not the students benefited from the statistics lessons, a pre- and post-test was given. The test, or survey, included four questions that covered probability, mean, median, standard deviation, normal distribution, and hypothesis testing (see Appendix 1). Each of these was covered in the lessons.

The results of the surveys show that each student's score improved when comparing the pre- and post-survey (see Table 1). The average increase of the scores was 4.5. This indicates that the students did in fact learn something during the intervention period.

Question 1 on the survey asked, "A box contains 4 red balls and 6 green balls. If a ball is drawn randomly from the box, what is the chance that it is red? If the ball is not replaced and a second ball is drawn, what is the chance that the second ball is red? If the first ball is replaced what is the chance that the second drawing gives a red ball?" On both the pre- and post-survey, the students got an average of 3.86 out of a total of 6 points.

Question 2 on the survey, which was worth 7 points, asked, "What are the mean, median, and standard deviation of the following set of numbers: 10 15 15 11 14." The students got an average of 1.64 out of a total of 7 points on the pre-survey and an average of 2.86 out of 7 on the post-survey.

Question 3 on the survey, which was worth 8 points, asked "If a set of data is normally distributed with mean = 100 and standard deviation = 16, what is the probability that a randomly chosen data item from the set has a value less than 100? Less than 116? Less than 84? None of the students received any points for this question on the pre-survey. On the post-survey, the students received an average of 2.57 out of 8 possible points. One students did not attempt to answer the question on the pre-survey and two of the students did not attempt to answer the question on the post-survey.

Question 4 on the survey, which was worth 5 points, asked, "A report states that women aged 25-35 drive an average of 7,000 miles per year. A survey of 500 women in a certain city finds that the average distance driven is 6,500 miles with a standard deviation of 2000 miles. Would it be reasonable to conclude that the average for **ALL** women in that city is less than the national average of 7,000? Explain your reasoning." Like the previous question, none of the students received any points for this question on the pre-survey. On the post-survey, the students received an average of .71 points out of 5 possible points. One student did not attempt to answer the question on the post-survey.

Interviews

Three students were interviewed in order to provide further insight into their learning and their perception of the lessons. It is interesting to note the correlation between their responses and what was observed during the classroom lessons, as well as the common themes that emerged from students' responses.

Each of the students found that the topic of Standard Deviation was the most difficult. Student 3 commented "Standard deviation; it was definitely hard to find it because there is a really long process so it is hard to remember how to do it. I have a bad memory." They may have grasped the concept but it seemed that the lengthy calculation process was too involved to hold their attention. It is no surprise that they were not particularly engaged during this lesson.

Students seemed to grasp ideas more easily when they were presented in multiple ways. It seemed there were different learning styles and that students benefited from different forms of presentation ranging from traditional learning and seat work to group work and hands-on

activities. Each student was asked "How would you say that you learn best?" Student 1 said, "I am kind of a visual learner; if I see it I'll know it." Student 2 said, "By being taught over and over again" and Student 3 responded "probably if I read something because it's like talking to myself and I kind of remember better when I see something." Repetition of topics in other courses also helped to reinforce the concepts. In stating what concepts were the least challenging, Student 3 said "Pretty much everything else [other than standard deviation] because I knew most of it and it's kind of like going back to simple stuff. Similarly, Student 1 said "probability, because we are learning that in science."

One issue that was very clear was their affirmation of the value of activities and games in the classroom. All of the students enjoyed being active during the lessons. In response to the question "What gets you engaged or interested in what is being taught?", student 2 said, "if they did something with an activity with it, like a game. Something like that so I can remember it." This also confirms the observation that students were more engaged during the activities.

Students were asked "What do you think students need to know before they can study and learn statistics?" Two of the students noted that they needed to know percentages or fractions before they could learn statistics. The third student displayed a level of higher order thinking when she suggested that one needed to know more of the process in order to carry out the procedures. She said, "They probably should learn processes, like how to go by steps and learn how to memorize going in order and which orders to do because that is my difficulty with memory and all, so I'll do one thing first and then it will end up being the wrong answer because it is not negative because I did it wrong."

Discussion

Seven lessons of about 45 minutes each were given to seven students covering the statistical concepts of probability, mean, median, standard deviation, normal distribution, and hypothesis testing. Two lessons were devoted to probability, two to mean and median, and one each to standard deviation, normal distribution, and hypothesis testing.

While observing the students during the lessons, it was noted that they did not seem interested and would often only partially pay attention. This may be due to the fact that they were not being graded on the material taught during these lessons. Nonetheless, the students would become engaged in the lesson when they were asked to perform a practice problem on their own or when they participated in an activity related to the lesson. It was interesting to note that student responses to interview questions correlated with the researchers' perceptions of student engagement during the lessons. In other words, when the researchers' observed that students were not focusing on the material, the students themselves reported they were not learning what was being presented and they felt the concepts were difficult. Although not a particularly new pedagogical idea, these findings reinforce the notion that if students seem disengaged, perhaps redirecting their focus or modifying how the material is presented is the best way to facilitate student learning.

Throughout the lessons, it was surprising to the researchers to find that the students had forgotten how to perform simple operations involving fractions, percentages, and decimals. Much time was devoted to reviewing these concepts that the students should have already been proficient in. The lack of foundational knowledge illustrates a commonly found issue in the teaching of mathematics (Mitchell, 1990). Without providing students with a clear understanding of basic concepts, more complex principles can not be easily taught, no matter how innovative or engaging the teacher is when presenting the information. One implication of these findings is

that more emphasis needs to be placed on the foundational mathematical concepts in order to prepare students for more complex learning. This may suggest that depth of foundational knowledge is more important than breadth of general mathematical knowledge, particularly early on in a student's academic career. By ensuring that students have an early firm foundation and working knowledge of these concepts, they are equipped to eventually excel in more advanced mathematics courses. However, a student's potential is limited when material is presented in such a way that exposes them to many concepts, but does not encourage them to master the basics before moving on.

A common theme that arose during the interviews was that the students had the most trouble with standard deviation. One of the students complained that there were too many steps in computing the standard deviation and another stated that it was hard for him to understand it. Unlike probability and mean, the students had not ever heard of standard deviation and perhaps one lesson was not enough to introduce it. It is of interest that there was consensus between the three students who were interviewed that standard deviation was the most challenging concept. In fact, only 2 students attempted to provide an answer to the standard deviation question on the posttest and neither was correct. Incorrect problem responses and lack of student engagement also indicated that hypothesis testing was a step too far for this grade level. These findings serve as a reminder that the developmental level of the students needs to be considered when determining what concepts need to be covered.

Conclusions

This study provides an insight into student learning of statistics at the middle school level. Students showed a significant gain in knowledge as evidenced by the improvement from pre to post-test. However, there were certain concepts that were problematic as shown in individual question responses and in student engagement during the lessons. The interviews indicated some of the barriers to learning that students encountered. Acknowledging these barriers will enhance the learning experience that teachers provide their students and will promote an awareness of the diversity of learning styles and preferences present in any given group of students. Although it is unreasonable to expect a teacher to cater to each student's preferences for every lesson, awareness of the barriers to learning and of different learning preferences can lead to using a variety of pedagogical techniques and creating a learning environment that is respectful of a diverse group of students.

In addition to implications for teaching mathematics, the current study also has implications for research in this area. Appropriate background material should be provided and assimilated by the students in the early stages of an intervention. The study could then proceed with more hands-on activities. When designing future problem sets to assess this type of intervention, items could be delineated more clearly with an emphasis on selected topics to specifically align each lesson with a particular question(s) in the problem set. This likely would enrich the interpretations made of the data gathered.

Larger class sizes would yield a more statistically valid sample than the one used in the current study. Information from a larger and more diverse student population, including that within a public school setting and including economically disadvantaged students, would increase the generalizability of these findings. Overall, these findings are important for enriching our understanding of how middle school students learn and understand statistical concepts and for guiding teachers and researchers to a better understanding of how to most effectively prepare students to use their mathematical knowledge and skills in the real world.

Authors

Jerry Dwyer is a faculty member in the Department of Mathematics and Statistics at Texas Tech University. He is an applied mathematician with an interest in how students at all levels learn mathematics. Malinda Colwell is a faculty member in the Department of Human Development and Family Studies at Texas Tech University. Her interests include developmentally appropriate practices for children and training pre-service teachers. Kim Moorhouse is a K-12 special education teacher in Dallas Texas.

References

Goldsby, D. S. (2003). Lollipop statistics. *Mathematics Teaching in the Middle School, 9,* 12-15.12-15.

Martinie, S. (2006). Data analysis and statistics in the middle school. *Mathematics Teaching in the Middle School, 12,* 48-49.

Mitchell, C. E. (1990). 7 Areas neglected in math. *Education Digest, 56,* 64-68.

National Council of Teachers of Mathematics, Principles and Standards for School Mathematics, Reston, Virginia, 2000.

Nooriafshar, M. & Maraseni, T. N. (2005). A comparison of learning preferences and perceptions of students for statistics concepts and techniques. *International Journal for Mathematics Teaching and Learning. Available online at* http://www.cimt.plymouth.ac.uk/journal/nooriafshar2.pdf.

Russo, L. M. & Passannante M. R. C. (2001). Statistics Fever. *Mathematics Teaching in the Middle School, 6,* 370-376.

Scepansky, J. & Carkenord, D. M. (2004). Senior year retention of methods and statistics concepts. *Learning of Psychology, 31,* 9-11.

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician, 57,* 74-79.

Wulff, S. S. & Wulff, D. H. (2004). "Of course I'm communicating; I lecture every day": Enhancing teaching and learning in introductory statistics. *Communication Education, 53,* 92-102.

Table 1 Problem set results

Student #	Pre-Survey	Post-Survey	Amount Increased
1	2/26	13/26	11
2	8/26	14/26	6
3	6/26	7/26	1

4	9/26	12/26	3
5	4/26	10/26	6
6	5/26	8/26	3
7	5.5/26	7/26	1.5

Appendix 1 Survey of statistics knowledge and skills

Please explain and show all of your work as fully as possible.

1. A box contains 4 red balls and 6 green balls. If a ball is drawn randomly from the box, what is the chance that it is red? If the ball is not replaced and a second ball is drawn, what is the chance that the second ball is red? If the first ball is replaced what is the chance that the second drawing gives a red ball?

2. What are the mean, median, and standard deviation of the following set of numbers?
   10 15 15 11 14

3. If a set of data is normally distributed with mean = 100 and standard deviation = 16, what is the probability that a randomly chosen data item from the set has a value less than 100? Less than 116? Less than 84?

4. A report states that women aged 25-35 drive an average of 7,000 miles per year. A survey of 500 women in a certain city finds that the average distance driven is 6,500 miles with a standard deviation of 2000 miles. Would it be reasonable to conclude that the average for **ALL** women in that city is less than the national average of 7,000? Explain your reasoning.