An Exploration of Students' Statistical Thinking

Randall E. Groth
Salisbury University, USA.
e-mail: regroth@salisbury.edu

Summary
The statistical thinking exhibited by 14–19 year-old students during clinical interview sessions is described. The students' thinking with regard to fundamental statistics concepts is reported in order to help inform instructional practice.

◆ INTRODUCTION ◆

There are several types of knowledge one needs in order to be an effective statistics teacher. Content knowledge is a necessary, but not sufficient, part of the knowledge base for teaching (Shulman 1987). Teachers also need to have knowledge of students' thinking processes (Even and Tirosh 2002; Fennema and Franke 1992). Building one's knowledge base with regard to students' statistical thinking pays dividends in instructional design.

In this article, some highlights are provided of a study of the statistical thinking of fifteen students aged 14–19. Each student was interviewed for a total of approximately two to three hours. Several statistical thinking tasks were posed to each student during the interviews. The thinking that students exhibited about some of the foundational ideas of statistics is reported below. Extensive details about the study design and the patterns of thinking the students exhibited have been reported elsewhere (Groth 2003a). This brief report is anecdotal rather than exhaustive in nature. It is hoped that the thinking reported in this brief article will help teachers build and/or solidify their knowledge bases in the area of students' thinking with regard to foundational concepts in statistics.

◆ CONCEPT 1: THE LAW OF ◆ LARGE NUMBERS

A useful operational statement of the law of large numbers reads:

> Draw independent observations at random from any population with finite mean μ. Decide how accurately you would like to estimate μ. As the number of observations drawn increases, the mean $\bar{x}$ of the observed values eventually approaches the mean μ of the population as closely as you specified and then stays that close (Yates et al. 1999, p. 390).

The intuitive idea behind the law of large numbers is that the more data that accumulate, the more certain we can be about our conclusions. Jacob Bernoulli, whose work led to the modern statement above, felt that this intuitive idea behind the law of large numbers was obvious. He remarked 'For even the most stupid of men, by some instinct of nature, by himself and without any instruction (which is a remarkable thing), is convinced that the more observations have been made, the less danger there is of wandering from one's goal' (as cited in Stigler, 1986, p. 225). Bernoulli, although a great mathematician, was not a noted teacher. Hence, I put his strong claim to the test during my interview sessions.

The interview question shown in figure 1 was designed to uncover students' understanding of the intuitive foundation of the law of large numbers. This two-part question was towards the end of a set of questions centred on a fictitious standardized test that had been given to students in several different states in the USA. The first part of the question asked students if they expected to observe a difference between the mean of a random sample of size $n = 5$ and another one of size $n = 15$. The second part was designed to find out if they understood that the smaller random sample was more likely to have a mean far from the population mean than the larger random sample, along with the idea that both would be likely to produce fairly non-representative samples. This two-part approach to investigating understanding of the law of large numbers was inspired by similar tasks written by Konold et al. (1993).

Fig. 1. Interview question concerning the intuitive foundations of the law of large numbers

I found that, contrary to Bernoulli's claim, the intuitive foundations of the law of large numbers are not obvious to all. One of the college students, Paul, who had recently completed a year-long high-school statistics course, thought that there would be no difference between the means of the two samples. In response to the interview question, he stated:

> No. Um, because it is a simple random sample, so everyone has an equal chance of getting picked. So by trying to do that, the ones that were picked will be spread out and averaging out to be basically the same, and that would be the same for both groups.

When asked if one of the samples was more likely than the other to yield an average of 17, he stated:

> No. Because not . . . too few of students scored 17 or above. It's not going to average out to be 17, most likely below, because most students did score less than 17.

Paul was not alone in his incomplete intuitive understanding, since six other students interviewed experienced similar difficulties in answering the same interview questions.

We need to make sure that all statistics students have an intuitive understanding of the law of large numbers since it is such a foundational concept. Computer software gives us one means for doing so. The *Sampling Sim* program (Garfield et al. 2001), available for free download at http://www.gen.umn.edu/research/stat_tools/, is a helpful program for building statistical intuition. Students can be asked to predict the result of drawing progressively larger random samples from a population. They can then compare the sample means with the population mean.

◆ CONCEPT 2: LINEAR DATA ◆ TRANSFORMATIONS

One of the recommendations set forth in the *Principles and Standards for School Mathematics*

We have the following information on test scores for three different classes in a school building that took the same 20-point test (Class B has two less students than Classes A and C):

Class A	Class B	Class C
12	7	7
13	9	11
14	13	12
14	14	14
15	15	15
16	19	18
16	19	19
17	20	19
18		20
18		20

Look at the data for Class A. Suppose the teacher adds 15 points to each student's score. How is the typical score for the class affected? How is the spread of the scores affected?

Fig. 2. Interview question concerning the impact of a linear transformation

(National Council of Teachers of Mathematics (NCTM) 2000) is that all students in grades 9–12 (the age group for the study described herein) should 'recognize how linear transformations of univariate data affect shape, centre, and spread' (p. 324). I designed the interview task shown in figure 2 to examine students' thinking with regard to this topic. The question asked them to describe what would happen to the centre and spread of a data set if each of its values was increased by 15.

As with the law of large numbers, I found that there were gaps in some students' intuitive understandings. Brooke, a high-school student enrolled in beginning algebra, recognized that increasing each score by 15 would also cause the centre to increase by some amount. However, when asked about what would happen to the spread, she responded:

> I think the spread would probably be more evident if you had more numbers in between. Because the way you have it now, there's just one number difference in all of the scores. But if you add 15 on, then there would be a noticeable difference.

Her response indicated that she felt adding 15 on to each score would actually increase the distances between the points in the data set. Brooke was not alone in this type of thinking. Two other students thought that there would be some impact on the spread of the data set if each value was increased by a constant amount.

Technology affords teachers the opportunity to challenge students' incorrect intuitions about the impact of linear data transformation on the centre and spread of a univariate data set. The idea that the centre increases by the amount of the transformation and that the spread remains unchanged can be illustrated quickly and efficiently through before-and-after computations of each measure executed on a calculator. Such exercises have the potential to help students build more solid intuitive understandings.

◆ CONCEPT 3: AVERAGES ◆

One of the primary emphases of high-school statistics curricula is teaching students to interpret and apply measures of central tendency such as mean, median and mode. It makes sense to help students understand these measures, since they are frequently encountered in everyday life. Each of the three measures provides a type of average, but it is important to be aware that their numerical values usually differ, at least slightly, depending on the data set. The media and politicians sometimes exploit the fact that the public does not understand this by presenting the average that best helps them justify the case they wish to make (Huff 1954).

Since understanding how to measure the centre of a data set is an important component of statistical literacy, I designed an interview question to investigate students' understanding of the topic. The item contained the 20 annual incomes shown in figure 3. The $2,000,000 income was an outlier, so the median income of $42,500 was fairly far away from the mean income of $140,650.

Some students seemed to have a deeply held belief that the typical income could only be determined by computing the arithmetic mean. Jeff, a student who had recently completed a high-school statistics course, did not seem bothered by the fact that the $2,000,000 income was high above the others. To find the typical income, he simply took his

16,000	56,000
80,000	5000
25,000	54,000
24,000	53,000
37,000	45,000
110,000	38,000
46,000	2,000,000
26,000	64,000
28,000	40,000
14,000	52,000

Fig. 3. Annual incomes in clinical interview task (US$)

calculator and found the arithmetic mean. Rick, a beginning algebra student, tried an intuitive approach to finding the typical value at first, but his intuition seemed to be trumped by his belief that the typical income should be determined by calculating the arithmetic mean. The following exchange took place during Rick's interview.

> Rick: 'Um, typical income . . . well without calculating it out, I would say it is around 40,000 to around 50,000 . . . should I give a precise answer?'
>
> Interviewer: 'Sure, as best you can do.'
>
> Rick (picks up calculator and finds arithmetic mean): 'Typical answer . . . typical income I found was um, 140,000, no, yeah, 140,000. I found this because there was a couple of high incomes, but there was also some really low incomes. There's a guy with a 2 million income, I think that really raised it, and then there's a guy with eleven hundred thousand, and a 110,000 income. Those were high incomes. But, the average income that he collected out of these 20 people, was 140,000 dollars a year.'

His initial intuitive analysis of the data set, which yielded a value fairly close to the actual median income, was set aside after the arithmetic mean provided a different answer.

Students who think like Rick must have their intuitive strategies encouraged and fostered before they become students who think completely like Jeff. Rick, at least initially, used a reasonable intuitive strategy to find the typical income in the data set. Unfortunately, that strategy was discarded due to his belief that the arithmetic mean is the only formal measure that produces a typical value. If Rick had been aware that his intuitive estimate was fairly close to another formal measure, the median, he may not have been as likely to retreat from his initial strategy. As teachers, we must be careful not to communicate the message that the arithmetic mean is the only formal measure of centre or typicality. Students should not completely disregard their intuition in choosing a measure of centre to describe a data set. Instead, they should be encouraged to let correct intuitions factor into the choice of formal measure of centre.

Suppose that the governor of Florida puts you in charge of finding answers to the following questions:
(a) What is the typical income of adults in the state?
(b) Will I be re-elected in the election this fall?
(c) What percentage of the state is computer literate?
(d) Does the new drug for treating the West Nile virus actually work?
(e) How successful was the law which raised the minimum driving age from 16 to 18?
Describe a plan for gathering the information you would need in order to answer each of the questions, and how you would carry out each plan and report the results to the governor.

Fig. 4. Interview question regarding statistical study design

◆ CONCEPT 4: EXPERIMENTAL ◆ STUDY DESIGN

NCTM (2000) recommended that students should have experiences of designing simple statistical studies as early as pre-kindergarten and then gradually learn to build more sophisticated studies all the way through to grade 12. If students understand the components of statistical studies, they will have a foundation to interpret claims made by studies in realms such as medicine and politics. Unfortunately, many students do not experience instruction that aligns with NCTM's vision. As a result, many adults have difficulty critically evaluating claims made by statistical studies.

In order to uncover students' notions about statistical study design, I designed the interview question shown in figure 4. The question presents several different scenarios, and asks students to design studies to investigate them. Part (d) of the task lent itself to an experimental study design. However, no statistical study design was suggested to or imposed on the students. They were free to conduct their investigations in whatever manner they deemed most appropriate. An exhaustive description of the patterns of thinking they exhibited is given elsewhere (Groth 2003b).

Nancy, a high-school geometry student, did not suggest designing a study at all in response to part (d) of the task. Rather, she relied solely upon pre-existing studies and artefacts, stating:

> Well, I guess you'd have to go to like a direct source from . . . I mean if you couldn't get the answers in a book or a periodical or online or something . . . If you could actually find someone who had actually encountered the virus and was working on a drug for it, if you had that kind of access.

Hence, she did not engage in the process of study design at all, preferring to leave the task to others she perceived to be better qualified.

All of the students interviewed, other than Nancy, proposed some type of study design to determine whether the hypothetical vaccine in part (d) was effective. However, six of those students designed studies that were not experimental in nature. Jessica, a high-school geometry student, offered the following response:

> I would speak to doctors, see what their opinion is. But, probably more importantly, talk to people who have contracted the disease and have taken the drug to see how they felt, have their symptoms got better, and how it worked out for them.

Brooke, a beginning algebra student, also discussed a non-experimental design:

> For the drug one, I think that would be pretty easy, because you would just have to go to the hospital and see how many had been given the vaccine. And then (to see) if it works . . . you would just have to look at the comments on the chart and see if they had to come back or not. Then you could kind of just make up a little average, I guess, of how many times it did work, and if it was good enough. You would probably have to have a pretty high percentage for it to actually work for most people. So, maybe you would have to compare ages, maybe for like certain ages it works better, or what not.

Although experimental methods are often considered the gold standard in determining cause and effect, several students interviewed did not even seem to consider imposing a treatment upon a group of individuals in order to determine its effectiveness.

Choosing to conduct an experiment to gather evidence for the effectiveness of a treatment does not come intuitively to some individuals. Therefore, statistics courses and textbooks that treat experimental and non-experimental study design as two different topics do a great disservice to students. It is not sufficient to teach students *how* to design experimental and non-experimental studies. We must also pay attention to helping them determine *when* one type of design is preferable to another. The ability to evaluate the appropriateness of statistical study design is foundational to evaluating claims derived from studies disseminated to our society.

◆ CONCLUSION ◆

Our society is rapidly becoming more and more dependent upon statistics and the results of statistical studies. Hence, one of our primary responsibilities as mathematics teachers is to foster students' statistical thinking. One must understand students' thinking in order to help it develop. It is my hope that some of the insights I have shared will help other teachers better understand their own students' statistical thinking. I suspect that some of the points discussed above will resonate with other teachers' experiences. In any event, the hope is that this article will serve as a catalyst for conversation among teachers, further investigation of students' statistical thinking and cognitively appropriate instructional design.

References

Even, R. and Tirosh, D. (2002). Teacher knowledge and understanding of students' mathematical learning. In: L.D. English (ed.), *Handbook of International Research in Mathematics Education*, pp. 219–240. Mahwah, NJ: Lawrence Erlbaum Associates.

Fennema, E. and Franke, M.L. (1992). Teachers' knowledge and its impact. In: D.A. Grouws (ed.), *Handbook of Research on Mathematics Teaching and Learning*, pp. 147–164. New York: Macmillan.

Garfield, J., Chance, B. and del Mas, B. (2001). *Tools for Teaching and Assessing Statistical Inference*. Available at http://www.gen.umn.edu/research/stat_tools/

Groth, R.E. (2003a). *Development of a High School Statistical Thinking Framework*. Doctoral dissertation, Illinois State University, 2003. *Dissertation Abstracts International*, **64**(04), 1202.

Groth, R.E. (2003b). High school students' levels of thinking in regard to statistical study design. *Mathematics Education Research Journal*, **3**, 252–69.

Huff, D. (1954). *How to Lie with Statistics*. New York: W.W. Nornton.

Konold, C., Well, A., Lohmeier, J. and Pollatsek, A. (1993). Understanding the law of large numbers. In: J.R. Becker and B.J. Pence (eds), *Proceedings of the Fifteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Volume 2*, pp. 299–305. Pacific Grove, CA.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: The Council.

Shulman, L.S. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review*, **57**(1), 1–22.

Stigler, S.M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Harvard University Press.

Yates, D.S., Moore, D.S. and McCabe, G.P. (1999). *The Practice of Statistics: TI-83 Graphing Calculator Enhanced*. New York: W.H. Freeman.