

## PERSPECTIVES ON FOUNDATION STATISTICS: SOME EXAMPLES FROM PROSPECTIVE SECONDARY MATHEMATICS TEACHERS

George Ekol  
University of Witwatersrand, Johannesburg  
george.ekol@wits.ac.za

*The qualitative case-study reported in this paper is framed by the ‘big’ ideas in introductory statistics, which emphasise statistical thinking, reasoning and literacy. We analysed secondary qualitative data from prospective secondary mathematics teachers’ statements, based on six themes in statistics: the influence of secondary-school experience and the teacher; difficult topics; avoidance of statistics; lack of coverage; teacher knowledge; reasons associated with the difficulty of statistics; and the contributions associated with school practice. We used five elements of the foundations of statistical thinking to enable us to analyse the prospective teachers’ statements from each of the six themes. Data analysis suggests that representing data to enable a clear understanding of statistical concepts emerged more frequently in the pre-service teachers’ statements than the other five foundation elements. The consideration of variability emerged the least number of times. Implications for statistics education are discussed.*

### INTRODUCTION

In his book, *Statistics on the Table*, published in 1999, Stigler stresses the importance of data for supporting claims made by individuals or by institutions. He cited a letter written by Karl Pearson to *The Times* of London in 1910, challenging scientists and non-scientists alike to do more than merely assert an answer, but to provide evidence: “If the question is important and one position has been advanced with well-considered supporting evidence, then it is incumbent upon a critic to ‘put statistics on the table’” (Stigler, 1999, p. 1). Stigler elaborated that putting “statistics on the table” is more than just putting forward numbers as evidence, because evidence could take any form of data, including qualitative data. To Stigler, a careful analysis of the “forces” that would affect any data should also be borne in mind, such as the methods by which the data were collected, methods of quantifying and expressing uncertainty and the method of agreeing on claims, for instance, when an assertion should be rejected or not. Stigler’s perspectives on data-based decision making are consistent with other research studies (for example, Wild & Pfannkuch, 1999; GAISE, 2005, elaborated by Ekol, 2013; and Ekol & Sinclair, 2016).

Wild and Pfannkuch (1999) proposed five elements which are fundamental to any introductory statistical enquiry. They include (i) considerations of variability, for example, activities such as noticing, estimating, predicting, and describing variations in a data set (Ekol, 2013; Ekol & Sinclair, 2016); (ii) reasoning with statistical graphs; (iii) integrating statistical and contextual knowledge; (iv) ‘trans-numeration’, representing data such that the ‘big ideas’ in introductory statistics, such as shape, centre, and dispersion, are clearly shown in the data without the need for numerical calculations; and (v) recognising the necessity of data for supporting claims. All five elements fundamentally involve the concept of variability.

Variability: statistical variability is considered by researchers and educators as the very foundation of statistics and without variability there would be no use for statistics. In introductory statistics, the terms *dispersion*, *spread* and *variability* are often used interchangeably to mean how stretched or squeezed a distribution is from the centre. Research studies, however, show that, while students are good at calculating the formal measures of variability, such as range, interquartile range and standard deviation, many are challenged by what these measures mean (Ekol, 2013; delMas & Liu, 2005; delMas, Garfield, & Ooms, 2005; Batanero, Burrill, & Reading, 2011; Wild et al., 2011). One of the challenges noted by delMas and Liu (2005) is the tendency of students to interpret graphical representations of data in terms of single measures, such as the range, the mean and standard deviation rather than first interpreting data in aggregate, in terms of the general shape of the graph, the positioning of its centre and the nature of its spread. Is the graph ‘fairly balanced’ about its centre (symmetrical) or is ‘somewhat unbalanced’ (asymmetrical or skewed). Our review of studies cited above leads to suggest that inadequate practice with aggregate reasoning with data at the introductory statistics level has a negative bearing on future data-based decision making. We view “data” more

broadly, integrating Stigler's perspectives as above, but also incorporating similar theoretical perspectives (for example, Wild & Pfannkuch, 1999; GAISE, 2005; Batanero, Burrill, & Reading, 2011).

From a pure statistics education stance (for example, GAISE, 2005; Batanero, Burrill, & Reading, 2011), much emphasis is put on developing the statistical thinking of the students and using authentic data that engage students in the processes of learning the concepts in statistics. Further underscored is integrating digital technology in statistics education, with the motive of minimising the burden on students of mechanical computations and focusing them on higher-level cognitive tasks. Among these are analysing data (for example, through activities such as comparing, differentiating, experimenting, testing and questioning) and evaluating claims (for example, through activities such as judging, selecting, arguing and defending).

The purpose of this paper is to identify and discuss, which elements of the foundations of statistical thinking (Wild & Pfannkuch, 1999) and GAISE (2005) framework emerge from pre-service teachers' statements about statistics before and after participating in the school placement. The specific research questions are (1) which elements of the foundations of statistical thinking are more likely to emerge from the pre-service mathematics teachers' statements about statistics before and after teaching statistics?; and (2) what are the implications for pre-service secondary statistics teacher education? Based on our review (for example, Wild & Pfannkuch, 1999; GAISE, 2005; Batanero, Burrill, & Reading, 2011), we assume that data-based decision making is enhanced if students are actively involved collecting, organising, analysing and interpreting the data. We further take it for granted that using digital technology to perform routine calculations releases students from routine processes and engages them in higher-level cognitive tasks, such as applying a statistical concept in a new situation, analysing data, and evaluating a decision based on data analysis. In the following paragraph, we briefly describe the background to the study and the source of the secondary data that we analysed to enable us to answer the research questions.

## METHOD

The design of this study is a qualitative case-study. Data analysed are secondary data obtained from Fitzmaurice, Leavy and Hannigan's (2014) research study. That study was carried out at the University of Limerick in Ireland in the 2010-2011 academic year. Participants were drawn from the two programmes which educate mathematics teachers in UL: a four-year undergraduate degree programme (leading to a BSc Physical Education & Mathematics degree), with 115 registered students and a one-year Professional Diploma in Education (PDE) in mathematics teaching. All the undergraduates had studied and passed higher-level leaving-certificate mathematics at the secondary school, which would put them in the category of academically high achievers. Fitzmaurice et al. (2014) first carried out a quantitative study of the prospective mathematics teachers ( $n=134$ ) using the Comprehensive Assessment of Outcomes in Statistics (CAOS) Test, to investigate levels of conceptual understanding (see delMas, Garfield, Ooms, & Chance, 2007) as opposed to procedural understanding of statistical concepts taught in introductory statistics. The authors also investigated the attitudes towards statistics of the prospective mathematics teachers using the Survey of Attitudes Towards Statistics (SATS) instrument (Schau et al., 1995). Findings from the quantitative study revealed that overall, the prospective mathematics teachers' attitude towards statistics was positive. However, a perception that statistics was a difficult subject to learn was evident in their comments.

It was the unresolved perception of the difficulty of statistics in the prospective teachers' comments that motivated Fitzmaurice et al. (2014) to carry out a follow up qualitative study to understand the underlying factors. Students who had participated in the quantitative study were invited and nine students volunteered to participate in one-on-one interviews. Four of the participants were second-year male students who had not taught statistics on school practice, but had done a course in introductory statistics three months before the interview. Five final-year undergraduates (three female and two males) students who had all taught statistics on school practice were also volunteered and were interviewed.

We chose Fitzmaurice et al.'s (2014) secondary data in our case study because of their detailed coverage of pre-service mathematics teachers' perspectives of foundation statistics, which we thought we could use as examples in our study. Even though we have highlighted results from the quantitative data, our primary interest is in the qualitative data. In the qualitative data, we identify five

elements of the foundations of statistical thinking that emerge from the prospective mathematics teachers' statements about statistics (Wild & Pfannkuch, 1999). In the next paragraph, we present the interview data adapted from the original study. We then discuss the data based on the theoretical framework. We conclude with some reflections on the teaching and learning of statistics.

## DATA PRESENTATION AND ANALYSIS

Interview data were obtained from the nine prospective mathematics teachers who were introduced in the Method section above. Each prospective teacher is identified by his or her year of study, gender and the serial number of his or her small group. For example, a male student who is in his second year of study is presented in the data as "second-year male 3". Similarly, a female prospective teacher who is in her fourth year of study and is assigned number two in her group is presented in the data as "fourth-year female 2". The square brackets [...] represent a pause in the student's sentence. In this paper, we only use part of the interview data from six pre-service teachers (two female and four male). Fitzmaurice et al. (2014) categorised the interview data into five inter-related themes, namely (i) the influence of secondary-school experience and the teacher; (ii) avoidance of statistics as a focus of study; (iii) teacher knowledge; (iv) perception of difficulty; and (v) the influence of teaching practice.

### *The influence of secondary-school experience and the teacher*

A second-year male 3 student said statistics "is a very important part of mathematics but, it is something that we very much skimmed over [...] we never did statistics and probability at the junior cert level. We did very little at the leaving level." He also said "statistics needs to be taught well...it is important to have a really good foundation [...] like everything with maths, if you don't understand the basics you can't develop [...] you can't build on it." By mentioning the importance of statistics, the student seems to relate with the many applications of statistics in real examples, which fits in our framework of integrating statistical data with context. This student also seems to be aware of the need for statistics to be taught well, which we categorised under trans-numeration – representing data to enable a clear understanding of concepts.

### *Difficult topics*

Some concepts, such as the outcomes of a random process, confidence intervals, and hypothesis testing, were perceived by the pre-service teachers as difficult. The second-year male 3 student said this about a random process: "am ... outcomes of a random process ...we haven't ... I have never done them before". About confidence intervals or hypothesis testing, the same student said, "to be honest... not my strongest area ... it is something I would have to look up ... I would definitely have to practice". The difficulty of the concepts seemed to be associated with the newness of the topic to the students, as fourth-year female student 4 stated, "It was just because it was new I think that I was a bit overwhelmed." We believe that difficult concepts such as confidence intervals and hypothesis testing can be mediated through the use of graphs (e.g. Wild et al., 2011), to enable students understand the concepts involved. We thus put random process, confidence intervals and hypothesis testing hypothesis under two foundation elements. First is representing data to enable a clear understanding of concepts (trans-numeration). For example, introducing random process through the concept of random sampling. Simple practical examples involving the use of random digits may lead learners to appreciate a random process. Second is reasoning with statistical graphs. The examples can be presented through graphs on the computer, so that students can clearly observe the simulations and perhaps begin to reason about them.

### *Lack of coverage*

Prospective teachers who had no experience teaching statistics said they would avoid teaching it if they had a choice. They did not feel confident that they would teach statistics during school practice, but they would rather teach subjects such as algebra, geometry and trigonometry. Second-year male 3 student stated this about his teaching preference, "My preference? ... def [initely] algebra number one [...] definitely geometry and trigonometry. Numbers and functions and then statistics and probability till last [...]". We are unable to categorise male 3 student's statement about lack of coverage under any of the five elements of our framework. We, however, note that the student did not

feel confident to teach statistics, probably because it was not compulsory and thus he did give as much thought to teaching it as he would to the other subjects. We propose that the student did have an interest in statistics, but, due to the influence of secondary-school experience and the teacher, he became discouraged with statistics and developed an interest in other areas of mathematics. This was the very same student who earlier in the data under influence of secondary-school, said “statistics needs to be taught well...to have a really good foundation.”

#### *Teacher knowledge*

The prospective teachers who had not taught statistics on school practice said their knowledge of statistics was not sufficient for teaching. Second-year male 1 student stated, “I wouldn’t say [my knowledge of statistics] would be brilliant”. Another student (second-year male 4) was concerned with interpreting information from graphs:

*actually knowing what the figures mean when you see them. And stuff like standard deviation as well [...] like seeing a figure [...] whatever 1.3 and saying that’s leaving it off. Yeah, maybe there is a grey area there where people don’t really know, including myself, what these things actually mean.*

Second-year male 1 student’s statement suggests a lack of knowledge of concepts in foundational statistics. Although his statement would fall in all the five foundation elements, we categorised it specifically under *representing data to enable a clear understanding of concepts*. As for the second-year male 4 student, his statements clearly indicate the elements of reasoning with graphs and considerations of variability.

#### *Reasons associated with the difficulty of statistics*

Some students perceived statistics as difficult because of the uncertainty of the topic and the language used. “You could think you were right for the whole question and then it could end up that you weren’t at all. So we learned the other topic more [...] at least we will know if we were right or wrong” (second-year male 4). We were not able to categorise this student’s statement in any of the five elements of our analytical framework. However, we think that the student was referring to questions, for example in algebra, such as, solve for  $x$  in the equation  $\log_3 243 = 3^{2x+1}$ . The answer to this question is more predictable than a question in introductory statistics, such as, “Briefly describe what is meant by *standard deviation*”. A student who answers the latter question, for instance, by simply stating the formula for standard deviation would probably feel confident that his or her answer is correct, the same way he would feel about obtaining the correct values of  $x$  in the former question. He, however, he or she would be disappointed to find out that a formula of standard deviation was not the answer expected from the latter question. Perhaps this simple example illustrates why many, certainly not all students, think that statistics questions are generally more difficult to pass than the “other” [topics] as claimed above by the second-year male 4 pre-service mathematics teacher.

#### *Contributions associated with teaching statistics during school practice*

The prospective teachers who taught statistics during school practice showed positive statements in the interview. For example, fourth-year female 2 said:

*statistics has been the one that I have liked the most.... I just liked it because it’s, it’s a lot more engaging and you don’t have to follow the book.... It’s easier to teach statistics—the students can relate to it.*

According to her, statistics is “something that’s valuable both in life and when you move on.” Fourth-year male 3, who also taught statistics during school practice, said he would still be interested in teaching statistics if it was an option, because statistics “is something that you can see everyday anyway, ...like ... seeing graphs, seeing bar charts...” Both students mentioned graphs in their statements, which we categorised under the element of *reasoning with statistical graphs*. Both students also affirmed the value of statistics in real life. We categorised those statements under *integrating statistical concepts with context*. The female student, in particular, stated that statistics is

“more engaging” [to teach] and “you don’t have to follow the text book”. We interpret this statement to include in general all the five elements of the foundation statistics, in particular, elements such as reasoning with statistical graphs, integrating statistical data with context by including real-life examples, and representing data to enable a clear understanding of the underlying concepts.

## DISCUSSION

We make no claim that our data analysis and the theoretical claims are watertight. Moreover, the secondary data we analysed were originally not meant for our theoretical framework. Some readers may even be inclined to think that we merely re-stated Fitzmaurice et al.’s (2014) research study without adding much to it. We do not think so. As stated before, the secondary qualitative data included many perspectives, which motivated us to analyse the data again, but using another theoretical lens. Our analysis did not in any way negate the results of the original study; rather, it complemented it. Thus, we believe that we have contributed to the original study by extending it, which we did by analysing the qualitative data using a different theoretical perspective.

Our study set out to answer two research questions: (1) which elements of the foundations of statistical thinking are more likely to emerge from the pre-service mathematics teachers’ statements about statistics before, and after teaching statistics during school practice?; and (2) what are the implications for pre-service secondary statistics teacher education? We counted how many times the pre-service teachers’ statements on six interview themes (i.e., *the influence of secondary-school; difficult topics; lack of coverage of statistics; teacher knowledge; reasons associated with difficulty of statistics; and contributions of school practice*) included keywords from the five elements of statistical thinking in our framework (i.e., *considerations of variability; reasoning with statistical graphs; integrating statistical data with contextual knowledge; ‘trans-numeration’, representing data to enable a clear understanding of concepts; recognising the need for data for supporting claims*). On the first research question, ‘representing data to enable a clear understanding of concepts’ emerged the most frequently in four out of the six themes (*the influence of secondary-school, difficult topics, teacher knowledge, and contributions of school practice*); whereas ‘representing data with statistical graphs’ emerged in three themes out of six (*difficult topics, teacher knowledge, and contributions of school practice*). ‘Integrating statistical data with context’ was identified in two themes (*the influence of secondary-school and contributions of school practice*), whereas ‘consideration of variability’ emerged only once under one theme (*teacher knowledge*). We expected the five foundation elements of statistical thinking to emerge from the students’ statements in all the six themes. However, we were not able to identify any foundation elements in the students’ statements under two themes (*lack of coverage; and reasons associated with difficult topics*). It is interesting to note that, ‘consideration of variability’, appeared less frequently in the pre-service teachers’ statements than the other elements yet it is a very important concept in introductory statistics. The explanation could be that the concept of variability is not well understood by many students, including senior undergraduate prospective mathematics teachers. Our data analysis also implicates teacher knowledge.

On the second research question, we suggest that statistics educators pay more attention to the seemingly dominant negative perceptions of statistics among high-school students joining university and other tertiary institutions. Addressing perceptions may be difficult, but applying some elements of the foundations of statistics, such as engaging students with interesting data, reasoning with statistical graphs and representing concepts in interesting ways to enable clear understanding, might contribute in changing perceptions of statistics. For prospective mathematics and statistics teachers, encouraging them to teach both mathematics and statistics topics during school practice can be one way of building their confidence in teaching both subjects, particularly teaching statistics.

## CONCLUSION

“Out of the [subjects] I have taught, statistics has been the one I’ve liked the most.... I don’t know. I just liked it because it’s, it’s a lot more engaging ... and you don’t have to follow the book...” (Fourth-Year Prospective Female Mathematics Teacher, cited in Fitzmaurice et al., 2014, p.12). Our analysis suggests that it is possible for pre-service mathematics teachers to change their perceptions about statistics and become active contributors to the discipline, if they are encouraged by their educators and mentors.

## REFERENCES

- Batanero, C., Burrill, G., & Reading, C. (Eds.) (2011). *Teaching Statistics in School Mathematics- Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study*. Springer Science & Business Media.
- Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy Reasoning, and Thinking* (pp. 295–318). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, 4(1), 55–82.
- delMas, R., Garfield, J., & Ooms, A. (2005). Using assessment items to study students' difficulty with reading and interpreting graphical representations of distributions. In K. Makar (Ed.), *Reasoning about Distribution: A collection of current research studies*. Proceedings of the *Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, Auckland, 2-7 July, 2005, Brisbane, Australia: University of Queensland.
- delMas, R.C., Garfield, J., & Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28–58.
- Ekol, G. (2013). *Examining constructs of statistical variability using mobile data points* (Unpublished doctoral dissertation, Simon Fraser University, Canada).
- Ekol, G., & Sinclair, N. (2016). Undergraduate students' conception of variability in a dynamic computer-based environment. In D. Ben-Zvi & K. Makar (Eds.), *The Teaching and Learning of Statistics – International Perspectives* (pp. 193-203). Springer International Publishing: Switzerland.
- Fitzmaurice, O., Leavy, A., & Hannigan, A. (2014). Why is statistics perceived as difficult and can practice during training change perceptions? Insights from a prospective teacher. *Teaching Mathematics and Its Applications: An International Journal of the IMA*, 33(4), 230–248.
- GAISE (2005). Guidelines for assessment and instruction in statistics education (GAISE) college report. Alexandria, VA: American Statistical Association.
- Stigler, S.M. (1999). *Statistics on the table: The history of statistical concepts and methods*. Boston, MA: Harvard University Press.
- Schau, C., Stevens, J., Dauphiner, T. L. & Delvecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement*, 55(5), 868–875.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–248.
- Wild, C. J., Pfannkuch, M., Regan, M., & Horton, N. (2011). Toward more accessible conceptions of statistical inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(2), 247–295.