Discussion: How Can Technology be Used to Teach Statistical Practice?

1. INTRODUCTION

Many established faculty in statistics have lived through the rapid changes in technology during the end of Twentieth Century and beginning of Twenty-first Century. Those of a certain age will remember the punch cards used for computing, and how we were not allowed to get particularly close to the computers. Even if we could see what the computers looked like, more than half of us at that time did not understand how they worked. The closest we got to technology in classrooms all over the world was a calculator, and even then calculators were usually prohibited in examinations as many students were unable to afford them. Nowadays, in developed countries we cannot contemplate examinations without calculators, and some exams are even undertaken in computer laboratories. It is obvious that technology is an undeniable part of education, especially statistical education. The question now is "*how can technology be used to teach statistical practice?*"

This discussion will summarize the two papers presented (Stern, Coe, Stern & McDermott, 2012; Baglin & Decosta, 2012) in 2012 IASE Roundtable Conference – *"Technology in Statistics Education: Virtualities and Realities"* – in Cebu, Philippines and the following discussions that took place after the presentations. In the last section a list of recommendations on learning and teaching and research will be provided.

2. OVERVIEW OF PAPERS

In this session we had two papers. The available teaching resources and educational environments for the authors of these papers were very different from each other. The first paper (Stern et al. 2012) presented a solution to a problem at Maseno University in Western Kenya, although the problem is not unique to them. In Kenya, both the students and the university have limited resources compared with the resources available in Western world. The second paper presented a theoretical framework, and the authors were located at Royal Melbourne Institute of Technology (RMIT) in Australia.

2.1 MSc Training in Research Methods Support

Stern et al.'s paper describes two contributions made to the MSc degree at Maseno. Their degree is different from those in other universities in Kenya because of how the degree is taught and also what is taught in the degree. The authors utilize freely and publically available resources such as Moodle as the Learning Management System (LMS) to enable them to offer e-learning to their students, YouTube videos to help students understand different aspects of the course, free online text book(s) to support students outside the class and free software (i.e. GenStat) to enable applications of what had been learnt in class.

The curriculum is designed with standalone modules and mimicks the process of statistical consulting. Students start their learning with a topic in data ownership (module 1), and the second module is planning the data flow (module 2). These initial modules are followed by planning the data collection (which elements, how and which platform) (module 3); and actual data collection if required (data collected on paper) (module 4) then data entry and organization of the data base, cleaning the data and preparing the data for analysis (module 5). These five modules are usually not part of the statistical degrees, but they are highly required skills by the employers. Since students at Maseno are already working in the industry, the value of these skills is recognized by the students and their employers. Module six covers various statistical analyses, which would be the main module in any statistical degree. Interpretation of the results from the analyses and writing up statistical reports form module seven, which is again often not prominent in statistical learning within many universities. The course finishes with two modules on data storage and access after the completion of a statistical problem solution (module 8), and dissemination and feedback to data originators (owners) (module 9).

There were some initial concerns regarding the implementation of this degree, both from the way it was taught and due to its name. It was not clear whether the students would participate in online discussions created in Moodle (LMS) and be able to work in groups in an online environment, given that they have not met face to face. The first iteration of the course helped the authors to overcome these fears, as they proved to be unfounded. Secondly, the academics were worried about the degree name 'research methods': whether students would apply for this degree and if employers would support their employees who wanted to study this degree. This has also been found unproblematic; conversely it made the course more popular because both the employers and the students recognized the importance of research methods skills due to their experiences in their workplace.

What was presented in this paper is parallel to what was presented in an earlier session by Finzer (2012) "*The Data Science Education Dilemma*" in the Roundtable. In a nutshell the aim of both the Stern et al. (2012) and Finzer (2012) papers is to show the real statistical analysis procedure to the audience. Statisticians, research analysts and data scientists are not just individuals who know and apply statistical methods to available data, they also design studies from the beginning to the end of the project including dissemination of the results. Therefore statistical education should also cover the required skills from the beginning to the end, not just the techniques and mathematics related to the techniques.

2.2 A Theoretical Framework for Developing Statistical Package Competence as an Outcome of Statistical Literacy

Baglin and Da Costa's paper argues that development of technological knowledge and statistical package competence should become a focus of statistical education. They alert us to the dangers of concentrating on a specific statistical package, teaching instead statistical packages that support the understanding of statistical theories. The ease of calculations and applications of statistical techniques to data sets by using a statistical package should not blind students to the theories and assumptions of the statistical analysis they are using.

The employers demand certain statistical software knowledge from their future employees. This is evident from job advertisements. Depending on the industry, students might be required to know different statistical software such SAS, SPSS and R. Therefore, we, the educators should answer two important questions before we can incorporate any statistical software to our classroom. Firstly, which package? Secondly, how should we teach it? The authors recommend that the decision as to which package should be used depends on the degree the students are completing, for example, in psychology they would be more likely to need skills in SPSS, and maturity of the students because it would be difficult, for example, to incorporate R without a GUI package such as RCommander to the first year statistics courses. Then the question remains, how should we teach?

Step by step instructions for each exercise (guided training) might help students to move forward quickly on their learning, but may also discourage thinking about what they are doing and why. Active exploratory learning is desirable if students have all the time in the world and nothing else to do, and they are very enthusiastic about statistics (we all know that would be very unlikely for majority of students who have to study statistics, whether they want it or not). Therefore, a balance between guided and active learning is required.

3. RECOMMENDATIONS

This section will provide a summary of the recommendations based on these two papers and the discussions following the presentations under two headings: learning and teaching recommendations and research recommendations.

3.1 Recommendations on Learning and Teaching

1. Although there might be resistance from the colleagues with whom we teach, we need to ensure that the soft skills, as well as skills that we all developed outside the education paradigm through our own work with clients (statistical consulting) or working on research projects are incorporated into statistical education.
2. Statistical software packages are an indispensable part of statistical work, and therefore, one of the learning outcomes for each statistical unit should be "competency of statistical software". This needs to be taught in a way that the skills learnt in one statistical software package can be easily transferable to another. The best analogy for this would be the skills required to drive a car. When someone has a driver's license, it is assumed that the person can drive any car, and it is not required that they sit another exam to drive a different car. The skills required to use one statistical software package should be as transferable as driving skills.
3. Statistics has become an even more important and prominent subject over the course of the last Century since data are accumulating rapidly due to technological improvement. We need to educate our students to use the available data and resources (statistical software) to answer questions which might help their community and improve their standard of living.

3.2 Recommendations on Research

1. The inquiry into statistical software competency needs more research. We need to investigate and document our research on this area to enable better learning environments and outcomes for our students. We should document not just successful outcomes, but also failures to help us to move to a better future together.
2. We need to have ongoing and updated research on the needs and wishes of the industry (i.e. employees) for our students. Of course courses should not be dictated purely by what employers want, although their views might help our students.

REFERENCES

Baglin, J. & Da Costa, C. (2012). A Theoretical Framework for Developing Statistical Package Competence as an Outcome of Statistical Literacy, *Technology Innovations in Statistics Education, 7*(2).

Finzer, W. (2012). The Data Science Education Dilemma. *Technology Innovations in Statistics Education, 7*(2).

Stern, D.A., Coe, R., Stern, R.D., McDermott, B.M. (2012). MSc Training in Research Methods Support, *Technology Innovations in Statistics Education, 7*(2).