THE EXPERIENCE OF TEACHING STATISTICS TO NON-SPECIALIST STUDENTS IN SAUDI UNIVERSITIES: THE ROLE OF TECHNOLOGY AND LANGUAGE

Abdullah ALOMIR
University of Glasgow
a.al-omir.1@research.gla.ac.uk

John H. MCCOLL
University of Glasgow
John.McColl@glasgow.ac.uk

Catherine BOVILL
University of Glasgow
Catherine.Bovill@glasgow.ac.uk

ABSTRACT: The importance of statistics is not limited to statisticians but also impacts on non-statisticians who have to use statistics. One important issue is how statistics is best taught to, and learned by, non-specialist students. The pervasive use of the English language causes additional challenges to learners whose first language is not English, especially when technological resources that use English language, such as statistics software packages, are an integral part of the course. This paper presents research into the current position in Saudi universities, where there has previously been a lack of research into this topic. Mixed methods research has been used: a questionnaire survey of 1,053 students and qualitative interviews with 16 teachers of statistics from all colleges within all six universities where statistics is taught to non-specialist students in Saudi Arabia's Eastern Region. This presentation will discuss differences between the experiences of learners taught in the Arabic and English languages.

Key words: Language in Statistics, Introductory statistics course, Non-specialist students, Statistical software packages; Saudi Arabia.

INTRODUCTION

With the continuing development of statistics as an independent discipline, the number of researchers and those interested in this discipline continues to increase. The need to learn statistics is not limited only to statisticians but also includes non-statisticians in various disciplines who use statistics within their fields, and this has made the teaching of statistics more complex. Therefore, one important issue that requires attention by researchers and those interested in the field of educational statistics is the teaching of statistics to non-specialist students (Garfield and Ben-Zvi 2007, Simpson 1995).

Research and studies related to the teaching and learning of statistics to non-specialist students are only beginning to emerge. Some of the most important topics in this research area which will be discussed in this paper are the impact of English language in learning and teaching statistics to non-specialist students whose first language is not English and how language can affect students' use of technological resources such as statistics software packages in statistics courses.

English Language in Learning and Teaching Statistics

Language plays a crucial role to connect parts of the educational process but sometimes it can be an obstacle for students. Usually, disciplines use specific terminologies which are unknown to junior students or those who are unfamiliar with the discipline. Also, many students who learn subjects in a second language have usually faced some difficulties with the language. Furthermore, printed and online reading resources and software tools which are used in the study of statistics are often not available in languages other than English. Students may face difficulties in their courses in trying to overcome these problems.

The language which is used in teaching statistics is one of most important factors that may help students to understand statistics courses. The use of technical words in teaching statistics to beginner students, particularly students from a non-English Speaking Background (NESB) who have a limited level of language skills, may lead these students to believe statistics courses are more difficult than they really are (Kaplan et al. 2009). Some

technical words used in statistics do not exist in all languages and some of them are not fully explained in the dictionaries (Abdelbasit 2010; Hubbard 1990). Moreover, in statistics there is often more than one meaning indicated by one word and in addition, there are several words used to imply one meaning, and this can add to student confusion. For instance, teachers often use "normally" and it does not necessarily relate to "normal distribution" but may relate to another meaning. Also teachers use "average" and sometimes "mean" and both can relate to the same concept. Kaplan et al. (2009) mention that "The use of domain-specific words that are similar to commonly used English words, therefore, may encourage students to make incorrect associations between words they know and words that sound similar but have specific meanings in statistics that are different from the common usage definitions" (p. 2). This can delay or reduce students' understanding of lectures.

Many difficulties that may face students when they use the English language in their studies need to be considered by teachers when designing statistics courses for non-specialists. One of these difficulties is that students need more time in lectures to understand the subject and this may cause the loss of students' attention. Students in many countries also rarely use English outside the classroom, so their level of English language does not improve as much as those students living in English-speaking countries (Abdelbasit 2010).

Technology in Learning and Teaching Statistics

In recent years, technology has become a major influence upon student learning of statistics because of the rapid development of a range of different technologies. At the moment, we cannot imagine that the process of learning and teaching statistics does not include technology because it has become an integral part of teaching. Statistical problems and calculations used to be calculated by hand and by calculator, but technology has changed this and thus approaches to teaching statistics have also needed to change (Chance et al. 2007). In previous years, teachers needed more time to explain new statistical concepts to students in the classroom and some statistical problems and analyses took more time to complete. This situation has changed because technology has provided us with many modern methods and tools to demonstrate these concepts better (Good 2006).

Technology has helped many teachers to teach statistics and it has also helped many students to learn statistics. Chance et al (2007) argue that the use of technology can benefit students through them being able to: collect data on a large-scale and deal with it more accurately; undertake statistical procedures and operations automatically; obtain the results and statistical interpretations more easily and display them with more clarity. These benefits can contribute to building confidence among students during their study of statistics and can encourage and motivate them to learn statistics (Good 2006).

This study examines the role played by English language in teaching statistics to non-specialist students (especially NESB students), and how the language used in teaching affects the use of technology in statistics in Saudi Arabian universities. This paper will discuss differences between the experiences of non-specialist students in Saudi universities taught in the Arabic and English languages.

METHODS

Mixed methods research is used in this study. Quantitative data allows the researcher to gather data from a large number of participants while qualitative data provides in-depth information from a smaller group of participants. Both quantitative and qualitative data complement each other in this study.

The first method of data collection (questionnaire) was used to collect data from non-specialist students who had completed at least one statistics course in different disciplinary colleges of the six universities in the Saudi Arabian Eastern Region. The instrument was designed to identify the current state of learning and teaching statistics for non-specialist students in Saudi universities. The instrument has thirty-one closed questions using yes/no and Likert scale responses, Strongly Disagree (SD), Disagree (D), Neutral (N), Agree (A), Strongly Agree (SA), as well as some open ended questions. The second method of data collection (interview) was used with teachers of statistics who teach statistics courses to non-specialist students in these colleges.

Because of limitations of cost and time affecting the data collection phase of the study, the researcher decided to collect the data from the eastern region of Saudi Arabia as a sample for this study. The eastern region is considered the largest province and covers 33% of the entire area of the country. Also, the population of the region represents 15.6% of the country's total population and, because of its location on the coast of the Arabian Gulf, the people who live in the eastern region represent a mixture of all of the groups in Saudi Arabia. Through the previous justifications, non-specialist students and teachers of statistics courses in the six universities in eastern region, which represent 18.2% of universities in Saudi Arabia and 100% sampling of universities in this

region, are the research study population and also provide the researcher with a realistic representation about other universities in Saudi Arabia. The researchers have chosen not to identify the specific institutions at this stage of the analysis.

Sampling

The 16 disciplinary colleges within the six universities in eastern region and 39 academic departments which form most of the academic departments within these colleges where statistics courses are taught to non-specialist students were identified. Then, one suitable class from each academic department was randomly selected and the study aimed to obtain 100% response rate from the study population (students) in these classes for two reasons; first, in order for this study to be more realistic and credible. Second, the last 15 minutes of lectures in these classes were used by the researcher to distribute the questionnaires to students participating in the study, so it was considered unjustified to include some students and not others. Therefore, questionnaires were distributed to every student present in the classes.

When the researcher (AA) entered the classrooms, he introduced himself to the students and briefly explained four important points: students who participated in the study would have to have completed at least one course in statistics, their participation would be voluntary and they were to make sure that they answered the questionnaire as realistically and honestly as possible. Finally, he informed them that all of the information in the questionnaires would be treated confidentially and would be used for research purposes only. He then distributed the questionnaires to every student in each classroom, regardless of the number of students, and so this is the main cause of variation in the number of participants in the various academic departments. The researcher stood at the front of the room until the students had finished filling in the questionnaires, which he then collected from them. The total number of responses from students in this study was 1053 non-specialist students.

To conduct the study of statistics teachers who teach non-specialist students from the six universities being studied, 16 statistics teachers who teach non-specialist students were randomly selected from most of colleges within these universities where statistics is taught to non-specialist. The interviews took 20-35 minutes for the teachers to answer all the questions.

Data analysis methods

The questionnaire contained questions asking respondents to rate their own English language abilities for four skills – reading, writing, listening and speaking – each at one of three levels – Excellent (scored 1), Good (2) or Poor (3). Cross-tabulations and chi-square tests were conducted to identify the level of association between these skills. The associations were very strong and the chi-square tests gave highly significant results, so the mean value of the four scores was calculated and rounded to give an integer-valued overall score for English language skill of 1, 2 or 3 with 3 representing the poorest level of ability.

Several questions had a ‘not applicable’ response option; ‘not applicable’ responses were discounted from the analysis. Seven questions using Likert scales phrased in a negative direction were reverse coded before being analysed.

Because there are many individual questions in the questionnaire, with substantial associations expected among them, factor analysis was conducted for dimension reduction. It was anticipated that the large number of observed items would represent a much smaller number of underlying factors, including the influences of English language and technology use. Cross-tabulations and chi-square tests were then used to explore the relationships between factor scores and background variables, including the main language used for teaching the statistics course (Arabic or English) and the student’s level of English language skill.

RESULTS AND FINDINGS

The-Kaiser-Meyer-Olkin measure verified the sampling adequacy for the factor analysis, KMO=0.82 overall and > 0.6 for individual items, which is well above the acceptable limit of 0.5. Bartlett’s Test of Sphericity, chi square = 4718.69, $\rho<.001$, indicated that correlations between items might be sufficiently large to make factor analysis useful. Five factors were identified with eigenvalues over Kaiser’s criterion of 1, which in combination explained 57.28% of the variance. These factors were interpreted as: (1) Effectiveness of Teachers, (2) English Language, (3) Relevance of Course, (4) Students Engagement and (5) Using Technology. This paper will focus on Factor 2 (English Language) and Factor 5 (Using Technology). Items contained in these factors are listed in

Table 1 and the frequencies of responses for these items after reverse coding (as necessary) are shown in Table 2 below. For each factor, an overall score was obtained by taking the mean value of all the completed items involved in that factor, then rounding the mean to an integer value of 1, 2, 3, 4 or 5 with 5 being the most positive response.

Table 1. Items Contained in Factors 2 and 5

Factor 2 (English Language)	Factor 5 (Using Technology)
Q34. Some statistical technical words which are not translated into my language make it difficult for me to learn statistics. (reverse coded) Q35. Most statistical software programs are in English and this makes it difficult to learn statistics. (reverse coded) Q36. My lack of English language skills is an obstacle to me using statistical software programs. (reverse coded) Q37. My level of English language skills has prevented me from accessing statistics resources and references in the English language. (reverse coded)	Q26. The teacher used technology (e.g. PowerPoint, the internet, statistical packages) effectively to help teach this course. Q27. Using a statistical software program led me to concentrate on the technology and not the statistical concepts being taught. Q28. Using a statistical software program to compute the statistical problems caused me misunderstanding of some steps of statistical procedures. Q29. Enough time was spent in learning and using a statistical software program. Q30. I could have used technology to learn statistics just as effectively without the teacher.

Table 2. Frequencies of Responses for Items in Factors 2 and 5

Factor 2 (English Language)						
	SD	**D**	**N**	**A**	**SA**	**All**
Q34	111	218	227	257	163	976
Q35	86	180	203	215	198	882
Q36	119	203	166	240	188	916
Q37	130	194	177	242	194	937

Factor 5 (Using Technology)						
	SD	**D**	**N**	**A**	**SA**	**All**
Q26	214	122	83	264	241	924
Q27	190	212	233	140	57	832
Q28	143	236	210	160	57	806
Q29	189	202	208	177	47	823
Q30	169	194	236	248	77	924

Comparison of Universities

In this section we summarize the abilities of students in the English language and describe the overall picture of their responses for Factors 2 and 5, depending on which university they were attending. Within the universities (1, 5 and 6) that used English as the main language for teaching, most students classified themselves in the first and second level (excellent and good) of English language skills. The highest rates were for universities 5, 1 and 6 respectively. After that came universities 2 and 3, where some of the colleges used English and the others used Arabic, then university 4, which used Arabic as the main language and where most students claimed to have poor English skills. Scores for Factor 2 showed that the majority of students in universities 1 and 5 had a positive experience of the way English was used in their statistics courses. After that came students in universities 6, 2 and 3 respectively, but the majority of students in university 4 felt that the use of English impacted negatively on their study of statistics. Furthermore, the scores for Factor 5 reveal that universities 5 and 1 also had the highest rates of satisfaction among students towards the use of technology in studying statistics, then came universities 2 and 3. In universities 6 and 4, the majority of students were dissatisfied with the use of technology. Despite the results for the other universities, it is perhaps surprising that in university 6, although English is the main language of instruction, the majority of students were dissatisfied with the use of technology in studying statistics.

English Language (Factor 2)

Figure 1 compares the distribution of Factor 2 (English Language) scores for students who were taught statistics in Arabic and English. 34% of students who were taught in Arabic had a negative experience of the impact of English on their study of statistics (factor scores of 1 or 2) whereas 32% of them had a positive experience (scores of 4 or 5). However, 58% of students who studied statistics in English had a positive experience and only 18% of them a negative experience. A chi-square test of association between language of instruction and Factor 2 score was significant: $x^2(4, N = 1010) = 78.96, p < .005$. The association was of moderate strength, Cramer's V = 0.28.

Figure 2 compares the distribution of Factor 2 (English Language) scores for students who reported different levels of skill in English. Within level 1 (Excellent) in English language skills, almost 75% of students had a positive experience of the use of English in their study of statistics (scores of 4 or 5), while 18% were neutral and 7% of them had a negative experience. On the other hand, pupils whose English skills are at level 3 (Poor),

49% found that the English language negatively affected their study of statistics. A chi-square test of association between level of English language skill and Factor 2 score was significant: $x^2(8, N = 1010) = 227.57, p < .005$. The association was of moderate strength, Cramer's V = 0.34.

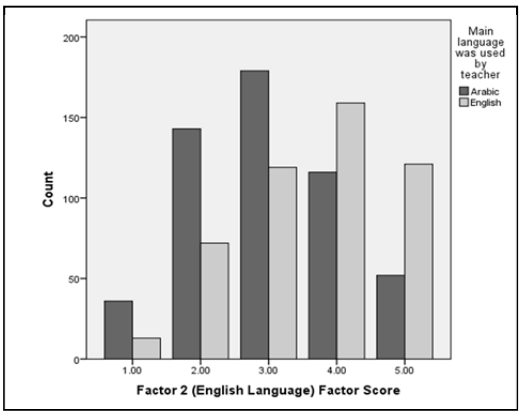

Figure 1. Frequencies of Factor 2 (English Language) Scores for Students Taught in Arabic and English

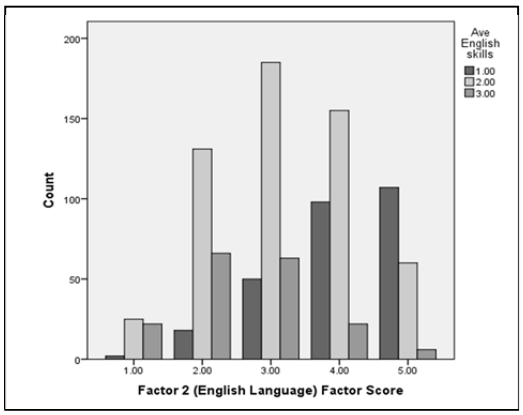

Figure 2. Frequencies of Factor 2 (English Language) Scores for Students with Different Levels of English Language Skills

Using Technology (Factor 5)

Figure 3 compares the distribution of Factor 5 (Using Technology) scores for students who were taught statistics in Arabic and English. When English was the main language of the course, 20% of students were dissatisfied with the use of technology (scores of 1 or 2) and 26% of them had positive satisfaction (scores of 4 or 5). Of students who studied the statistics course in Arabic, 46% were dissatisfied and 18% positively satisfied. A chi-square test of association between language of instruction and Factor 5 score was significant: $x^2(4, N = 1002) = 76.29, p < .005$. The association was of moderate strength: Cramer's V = 0.28.

Figure 4 compares the distribution of Factor 5 (Using Technology) scores for students who reported different levels of skill in English. Among students whose English language skills were excellent (level 1), 21% had negative responses (factor scores of 1 or 2) regarding the use of technology in their statistics course compared to 25% who had a positive response (scores of 4 or 5). On the other hand, 46% of students whose English level was poor (level 3) were not satisfied and just 17% were positively satisfied. A chi-square test of association between level of English language skill and Factor 5 score was significant: $x^2(8, N = 1002) = 33.18, p < .005$. The association was of weak strength, Cramer's V = 0.13.

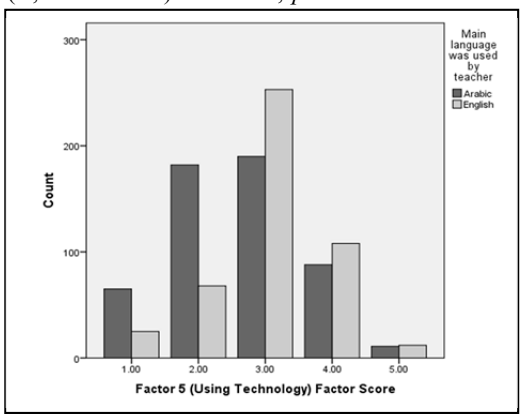

Figure 3. Frequencies of Factor 5 (Using Technology) Scores for Students Taught in Arabic and English

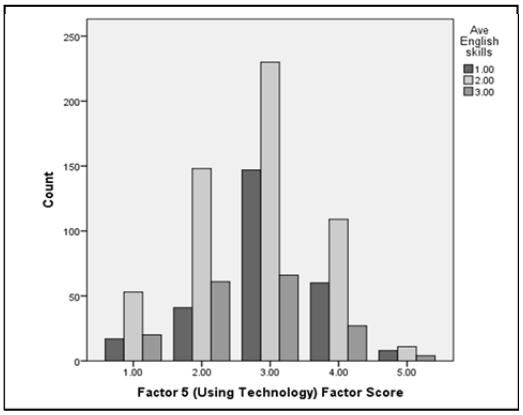

Figure 4. Frequencies of Factor 5 (Using Technology) Scores for Students with Different Levels of English Language Skills

Teachers' Interviews

Analysis of the staff interviews is at an early stage, but some themes related to this paper are already emerging. In using technology, almost all the teachers are using PowerPoint and the Internet for teaching statistics. Most of the interviewees are using statistical software in their teaching (3 are not) and the following packages are mentioned: Excel, SPSS, Matlab, Minitab, Statistica and SAS. Lack of technical facilities and support staff and large class sizes are the main challenges that face teachers who try to use technology in teaching statistics.

Concerning the influence of the students' limited English language skills on their study of statistics, teachers were divided into two groups based on their opinions. Teachers in group one said there are no problems relating to this issue. For example, teachers 1, 2 and 12 find that there is no effect of English in studying statistics; as Interviewee 1 said, "There is no significant effect because Statistics does not require students to have a good command of English language". The student questionnaires on the other hand revealed that the majority of students in the same colleges as these members of staff have negative responses regarding to the influence of English in their study of statistics. The rest of the teachers in this first group explained that the lack of impact of the English language on studying statistics is due to the good English skills of their students; as Interviewee 8 said, "This problem does not affect students' learning of Statistics in this university because the course is 100% taught in English. The fact that the teacher's mother tongue is Arabic may drive him to spontaneously use some Arabic words in the lecture, though the use of Arabic is not encouraged". The second group of teachers argue that the limited English language skills of students, does affect their study of statistics, whether they used Arabic or English as the main language. For instance, Interviewee 15 said, "I often find in exams that the most common errors in the answer sheet result from students' incorrect understanding of the questions, which affects their final results. The students themselves admit that their level in the English language is not good enough to understand the exam questions". Also, Interviewee 16 mentioned that, "Sometimes, I imagine that were the teaching of Statistics in Arabic, students' understanding would be much better. In general, the English language stands as a stumbling block to students' learning."

CONCLUSION

Most statistical software packages and their commands are in English. Text books, websites, lectures on YouTube and other available statistics resources are mainly in the English language and significantly outweigh resources available in Arabic. Consequently, as we can see from the data presented here, many of the non-specialist students in Saudi Arabia's Eastern Region who are taught statistics courses in Arabic have encountered challenges and difficulties in using technology in statistics courses, and most of the students taught in Arabic agree that the English language is an obstacle to their study of statistics. Conversely, the majority of students who are taught a statistics course in English have found that English language was very useful for them and supported them in learning statistics and in dealing with technology. High percentages of students who reported that there were no challenges in using technology were also those students who described their English language ability as excellent. This suggests that English language skills play a key role among non-specialist students in their studying of statistics courses. Our future analysis will include examining whether the type of university is related to the decision about what language is used to teach statistics teaching, and whether student capabilities in particular institutions are related to language skills.

To maximise the benefits of studying statistics for non-specialist students in Saudi Arabia's higher education institutions, the administrators and teachers of statistics in these institutions need to consider several things. First, for institutions that use English language in teaching statistics, they have to be sure that their students' English language skills are sufficient in order for English not to be an obstacle to them. In these institutions as well as in institutions that use Arabic for teaching statistics courses, institutions need to consider teaching some supplementary English language covering relevant statistical terminology to assist students to use more statistics resources that can enhance their use of technology, support student learning through making studying statistics easier. It is also necessary to help students by teaching them the basic skills to use statistical software programmes (Simpson 1995). In designing assessments in English and Arabic, it is important that questions are stated as clearly and unambiguously as possible, and with cognisance of the potential for students to misunderstand statistical terms that are not clearly stated. Where students have good English language skills, teachers should be encouraged to use the good range of statistics resources available in the English language to benefit students. Finally, it is very important for institutions to encourage statistics specialists and researchers to write and translate statistics books into Arabic and create online resources to provide a better range of references for non-specialist students studying in Arabic.

Abdelbasit, K. M. "Teaching Statistics in a Language Other Than the Students'." *Presented at Data and Context in Statistics Education: Towards an Evidence-Based Society. Proceedings of the Eighth International Conference on Teaching Statistics. Voorburg, The Netherlands: International Statistical Institute. http://www. stat. auckland. ac. nz/~ iase/publications. php*.

Chance, B., Ben-Zvi, D., Garfield, J., and Medina, E. (2007). "The Role of Technology in Improving Student Learning of Statistics."

Garfield, J., and Ben-Zvi, D. (2007). "The Discipline of Statistics Education." *http://www.ugr.es/~icmi/iase_study/BackgroundpaperGarfield.pdf*.

Good, P. (2006). *Resampling Methods*: Birkhauser.

Hamadu, D., Adeleke, I., and Ehie, I. (2011). "Using Information Technology In Teaching Of Business Statistics In Nigeria Business School." *American Journal of Business Education (AJBE)*, 4(10), 85-92.

Hubbard, R. (1990). "Teaching Statistics to students who are learning in a foreign language." *ICOTS 3*, 514 – 517.

Kaplan, J. J., Fisher, D. G., and Rogness, N. T. (2009). "Lexical ambiguity in statistics: What do students know about the words association, average, confidence, random and spread." *Journal of Statistics Education*, 17(3), n3.

Simpson, J. M. (1995). "Teaching statistics to non-specialists." *Statistics in medicine*, 14(2), 199-208.