Gender Differences in SEC Mathematics 2015

Introduction

The aim of this report is to analyse any existent gender differences in candidates' attainment in the SEC Mathematics 2015 assessment. The candidates' average attainment in parts of Paper 1 were computed and statistically significant differences between the marks obtained by male and female candidates sought. The report looked at the content and structure of the questions where such differences were observed and at those questions where they were not. A sample of candidates' responses was also analysed. The aim was to try and assess what causes these differences. By so doing, one can assess whether there are questions which favour a particular gender. Then, one can see whether these differences in construction, if existent, are part of the syllabus or exterior artefacts and assumptions.

The table below shows the number of candidates who sat for SEC Mathematics in 2015 stratified by paper choice and gender. Candidates, who were absent for any paper were omitted from this research, are not counted in the table below.

Table 1: Candidates Sitting for SEC Mathematics in 2015

Candidates			
IIA		IIB	
Males	Females	Males	Females
928	970	1073	1161

Like most other subjects studied at this level, more female than male candidates registered for SEC Mathematics in 2015. Contrary to other SEC subjects, however, a greater percentage of candidates registered for Paper IIB rather than Paper IIA (MATSEC Support Unit, 2016).

For this report, only Paper I will be analysed. This is a common paper for all candidates. SEC Mathematics Paper I consists of two sections:

- Section A is a non-calculator section consisting of 20 questions each carrying one mark. 20 minutes are allowed for this section. This is often called the 'mental paper'.
- Section B is a calculator section. This section carries 80% of the marks for Paper I and has a time allocation of 1 hour 40 minutes. In 2015 it consisted of eleven questions with different mark allocations.

Statistical Analysis

Independent samples t-test with equal variances assumed were conducted whereby the average difference between the marks of males and females for items were analysed. The analysed items include all components of Paper 1 and the final mark obtained by candidates. Candidates were stratified according to paper choice. Statistically significant differences at α=0.05 are shown in bold.

Table 2: Males and Females Average Attainment

Item		IIB Candidates			IIA Candidates		
		Males Average Mark	Females Average Mark	Average Difference (Male – Female)	Males Average Mark	Females Average Mark	Average Difference (Male – Female)
Paper 1	Mental Paper	9.8	8.4	**1.4**	16.3	15.4	**0.8**
	Question 1	2.6	2.7	-0.1	7.2	7.1	0.1
	Question 2	5.2	5.2	0.0	9.6	9.6	-0.1
	Question 3	3.9	3.6	**0.3**	6.7	6.6	0.1
	Question 4	2.2	2.3	-0.1	6.8	7.1	**-0.4**
	Question 5	2.3	1.9	**0.4**	5.1	4.6	**0.5**
	Question 6	2.8	2.9	-0.1	4.6	4.7	**-0.2**
	Question 7	2.5	2.7	-0.2	6.2	6.2	0.0
	Question 8	0.7	0.4	**0.3**	2.8	2.0	**0.8**
	Question 9	2.4	1.8	**0.6**	5.3	4.7	**0.5**
	Question 10	2.6	2.3	**0.3**	6.0	5.4	**0.6**
Total (Paper 1 + Paper 2)		90.3	87.2	3.1	128.5	125.7	**2.8**

In 2015, male candidates outperformed female candidates in most questions. In nearly half of these questions, the difference was statistically significant. There are only two instances where female candidates outperformed male candidates with a difference that is statistically significant, and both these instances are for Paper IIA candidates. There is no instance where Paper IIB female candidates outperformed male candidates with a difference that is statically significant. The difference between Paper IIA and Paper IIB candidates is, as one would expect, quite large. However, this was not investigated for this report.

In Paper 1 questions 5, 8, 9 and 10, males statistically outperformed female candidates irrespective of paper choice. Thus, this report analyses question construction and candidate responses to these items to seek what might be causing these differences. Males also statistically outperformed female candidates in the mental component, however this was not analysed because marks for each of the 20 items making up the mental paper are not available.

In Paper 1 questions 1, 2 and 7, the difference in performance between male and female candidates is not statistically significant irrespective of paper choice. Thus, this report analyses question construction and candidate responses to these items to contrast these with those of the other set of questions.

Question Construction

Questions favouring male candidates

Question 5 was a 7-mark question divided into four parts. The question tested candidates' knowledge of algebra and graphs applied to a fictitious scenario. The question presented two mobile phone plans to candidates and the needs of Oliver. They were asked to select which plan is best for Oliver, to convert a plan to an algebraic equation, plot it and determine when it is best to use which plan. Thus, the second part question involved converting information into an abstract

logical representation and is where, according to the Examiners' Report, candidates lost most marks. When it came for graph plotting, candidates did better as most did not use the equation they had designed for the purpose, but opted to use raw data. This could be for different reasons, maybe candidates' lack of comprehension of equations or candidates opting for a 'safer' path rather than using the equation they constructed.

Question 8 was a 5-mark question. This question was not divided in parts. It presented candidates with a fictitious situation in which Xandru has to estimate the time taken for two printers printing at different rates to print a certain number of copies. Like question 5, candidates were expected to convert the information provided into an abstract logical representation. The Examiners' Report notes that this "non-routine" question could also be solved by trial and improvement.

Question 9 was an 8-mark question divided into four parts. The question was about volume and had, unlike questions 5 and 8, no fictitious scenario. The questions could be solved mostly by imagining/drawing a 3x3 cube made up of 1x1 cubes and counting. One part could be solved by simple algebra. The Examiners' Report suggests that candidates performed poorly in this question.

Question 10 was an 8-mark true/false/sometimes true exercise where examples had to be provided for 'sometimes true' cases. It involved basic knowledge of numbers and decimals. The Examiners' Report suggests that candidates performed poorly in the last two items, which were those requiring the use of decimals/fractions to prove why a given statement was not always true (e.g. $100/x < 100$).

Questions with no statistical difference between male and female candidates

Question 1 was divided into two parts and carried a total of 9 marks. It was a rather routine question where part (a) asked candidates to simplify algebraic equations and part (b) asked candidates to calculate the volume of a given shape. The Examiners' Report confirms that candidates did generally well in both parts of this question.

Question 2 was an 11-mark question. The first part of the question (sub-divided into four parts) was designed to assess candidate's use of Microsoft Excel whereby the second part of the question involved the use of a simple ratio. The Examiners' Report notes that "question 2 turned out to be the easiest question in Paper I Section B."

Question 7 was an 8-mark question divided in two parts. The question was about bearings and trigonometric ratios and seems rather routine. Nevertheless, the Examiner's Report notes how candidates had some difficulties with trigonometry, sometimes attributed to calculator errors.

Candidates Responses

The MATSEC Support Unit keeps copies of scripts for the last three candidates obtaining each grade from 1 to 5 for Paper IIA candidates and 4 to 5 for Paper IIB candidates. The performance of these 21 candidates in each of the aforementioned questions was analysed. Achievement in each question was fragmented into parts as per question to try to determine which parts of the questions caused the differences. However, for the available sample, the difference between question marks was not in parallel with that observed for the whole sample. Thus, the sample was concluded to be unrepresentative of the whole group and this part of the analysis was not carried out.

Conclusions

There are four questions involving named fictitious characters in Paper 1 Section B – Stella, Janet, Oliver and Xandru. Two of these are males. A fifth question involves a man who is not named. The two questions involving named males were questions where a statistically significant difference in favour of male candidates was observed. However, it might be over-simplistic to attribute the advantage given to male candidates to this simple fact. Moreover, the near equal prevalence of male and female characters suggests no deliberate 'bias in test content' (Rosser, 1989).

A general observation, however, is that questions 5, 8, 9 and 10 might have been more cognitively demanding than questions 1,2 and 7. Questions 1, 2, and 7 relied on the knowledge and application domains while question 5, 8, 9 and 10 had instances where candidates were expected to tap skills from the reasoning domain. The TIMSS 2011 report describes the three domains as follows:

> *The first domain, knowing, covers the facts, concepts, and procedures students need to know, while the second, applying, focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions. The third domain, reasoning, goes beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multi-step problems. (Mullis et al. (2009)).*

While questions 1, 2 and 7 asked candidates to write answers from existing knowledge and carry out calculations to familiar situations, questions 5, 8 and 10 also asked candidates to make choices, give evidence, and solve unfamiliar problems. This, however, does not account for question 9 which commonly asked candidates to count.

This brief study shows that male candidates outperformed female candidates in 2015 SEC Mathematics. Initial analysis suggests that this difference was pronounced in questions involving higher-order cognitive skills. One questions whether this finding can be sustained through further research, both in Mathematics and other subjects, and at this level and other levels. If so, there would be numerous questions to ask about the causes and effects of this observation.

References

MATSEC Examinations' Board (2015). SEC Mathematics May 2015 Examiners' Report. Retrieved from http://www.um.edu.mt/__data/assets/pdf_file/0003/262830/SECMATHS.pdf

MATSEC Support Unit (2016). MATSEC Examinations Board SEC Examinations 2015: Statistical Report. Retrieved from http://www.um.edu.mt/__data/assets/pdf_file/0004/271273/SECStatisticalReport2015.pdf

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan C. Y., & Preuschoff, C. (2009). TIMSS 2011 Assessment Frameworks. Boston: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Rosser, P. (1989). *The SAT Gender Gap: Identifying the Causes.* Washington: Centre for Women Policy Studies. Retrieved from http://files.eric.ed.gov/fulltext/ED311087.pdf